A Historical Comparison of Resource-Based Theory and Five Schools of Thought Within Industrial Organization Economics: Do We Have a New Theory of the Firm?

Kathleen R. Conner
University of Pennsylvania

A resource-based approach to strategic management focuses on costly-to-copy attributes of the firm as sources of economic rents and, therefore, as the fundamental drivers of performance and competitive advantage. Interest presently exists in whether explicit acknowledgement of the resource-based view may form the kernel of a unifying paradigm for strategy research. This article addresses the degree to which a resource-based view represents a fundamentally different approach from theories used in industrial organization (IO) economics. The central thesis is that, put in formal terms, the resource-based approach is reaching for a theory of the firm. To determine its distinctiveness in comparison to IO, therefore, an appropriate comparison is with other theories of the firm developed within that tradition. Section I summarizes and analyzes five theories that have been significant in the historical evolution of IO. These are neoclassical theory's perfect competition model, Bain-type IO, the Schumpeterian and Chicago responses, and transaction cost theory. The first part of Section II analyzes the resource-based approach in terms of similarities to and differences from these IO-related theories. The conclusion is that resource-based theory both incorporates and rejects at least one major element from each of them; thus resource-based theory reflects a strong IO heritage, but at the same time incorporates fundamental differences from any one of these theories. The second part of Section II analyzes resource-based theory as a new theory of the firm.

A resource-based approach to strategic management focuses on costly-to-copy attributes of the firm as sources of economic rents and, therefore, as the fundamental drivers of performance and competitive advantage (Barney, 1986; Rumelt, 1984, 1987). According to this perspective, a firm's ability to attain and keep prof-

Address all correspondence to Kathleen R. Conner, The Wharton School, University of Pennsylvania, Philadelphia, PA 19104.

Copyright 1991 by the Southern Management Association 0149-2063/91/$2.00.
itable market positions depends on its ability to gain and defend advantageous positions in underlying resources important to production and distribution: as Wernerfelt (1984:171) puts it, "[f]or the firm, resources and products are two sides of the same coin."

A historical review of strategy research suggests that a resource-based perspective long has been central to the field. Influential literature, including, for example, Barnard (1938), Selznick (1957), Sloan (1963), Chandler (1962, 1977), and Rumelt (1974), connects performance with a firm's special competencies in deploying and combining its human, physical, and reputational capital. Indeed, the core notion of strategy as a fit between the internal competencies of the firm and external opportunities [Christensen, Andrews, Bower, Hammermesh, & Porter, 1987] incorporates a resource-based perspective. Rumelt comments on the historical centrality of the resource-based view, as well as its indivisibility from the concept of business strategy, as follows:

In essence, the [strategy] concept is (italics added) that a firm's competitive position is defined by a bundle of unique resources and relationships and that the task of general management is to adjust and renew these resources and relationships as time, competition, and change erode their value. This way of looking at the firm...[is] useful in describing and summarizing the empirical studies of firm behavior that form the core of the business policy literature. (1984: 557-558)

Interest presently exists in whether explicit acknowledgement of the resource-based view may form the kernel of a unifying paradigm for strategy research (see, e.g., Mahoney & Pandian, 1990; Teece, Pisano, & Shuen, 1990). This article addresses the degree to which a resource-based view represents a fundamentally different approach from theories used in industrial organization (IO) economics. The central thesis is that, put in formal terms, the resource-based approach is reaching for a theory of the firm.¹ To determine its distinctiveness in comparison to IO, therefore, an appropriate comparison is with other theories of the firm developed within that tradition. Section I summarizes and analyzes five theories that have been significant in the historical evolution of IO. This discussion reveals, among other things, the central role that an embedded theory of the firm has played in the associated research streams. The first part of Section II analyzes the resource-based approach in terms of similarities to and differences from these IO-related theories. The conclusion is that resource-based theory both incorporates and rejects at least one major element from each of them. The second part of Section II analyzes resource-based theory as a new theory of the firm. Concluding remarks address among other things issues important in operationalizing resource-based theory and the potential contribution of the "new IO's" game-theoretic approach.

Five Theories of the Firm From IO Economics

In the effort to assess the differences and similarities between strategy's resource-based approach and previous theoretical developments in IO, it is important to articulate the type of theory with which the resource-based approach

¹Notes follow this article.
should be compared. The thesis upon which this section is based, in a nutshell, is that the resource-based approach is reaching for a theory of the firm, and thus that predecessor theories of the firm are appropriate standards by which to compare the novelty of this strategy approach.\(^2\)

According to Holmstrom and Tirole's (1989: 65) definition, a theory of the firm must address two central questions: (a) why firms exist (their purpose), and (b) what determines their scale and scope. As Holmstrom and Tirole put it: "[t]he challenge is to offer a genuine trade-off between the benefits and costs of integration. One needs to explain both why firms exist as well as why all transactions are not organized within a single firm."

This section briefly reviews the theory of the firm embedded within five influential schools of thought related to IO economics. The theories treated include neoclassical theory's perfect competition model, Bain-type IO, the Schumpeterian and Chicago responses, and transaction cost economics.

With regard to the question of why firms exist, it is useful to note that each theory discussed here takes as a given that the ultimate purpose of the firm is to maximize profits. Thus these approaches differ not with regard to a firm's fundamental objective (to make money), but rather with regard to the primary means through which firms attempt to realize this objective. Thus a theory of the firm identifies the main problem a firm must solve in order to earn above-normal returns (and, reciprocally, the limits a firm faces in solving this problem). The theories reviewed here are parsimonious in that each concerns a single problem seen as paramountly affecting firm performance.

**Neoclassical Perfect Competition Theory: Firms as Combiners of Inputs**

In the perfect competition model, the firm exists to combine resources to produce an end product. Thus is incorporated the critical concept of the firm as the embodiment of team production. The firm's product is the joint output of multiple inputs working together; the entity we call the "firm" owns the rights to the productive services of the (multiple) inputs (Alchian, 1982; Alchian & Demsetz, 1972). In the neoclassical model, firms produce by teaming two inputs: labor and capital.

Perfect competition theory generally assumes, however, that (a) the "right" input mix can be readily ascertained (thus the firm is normally represented as a cost curve derived by calculus from a production function relating inputs to outputs); (b) the marginal contribution of each input is easily calculated; (c) all parties have perfect and complete information; and (d) resources are completely mobile and divisible and hence flow unimpeded to the highest-valuing use.\(^4\) Firms are identical because perfect information together with a specifiable production function assures that each firm has equal access to product technology; perfect information plus resource mobility and divisibility assures that each firm is able to obtain exactly the right inputs. Thus the individual firm's ambition to maximize profits yields a market equilibrium of zero economic returns to each firm, because they are equally able to team the proper inputs.\(^5\) An input's price equals the input's marginal productive value to the firm.

Because firms are seen as small entities producing single products, scope
(breadth of line) generally is not an issue. Firm size (quantity produced), however, is bounded by the standard assumption of increasing average costs past a production level that is small, relative to the size of the market. Thus limits to firm size are established by technological, and later, managerial, scale factors (see, e.g., Holmstrom & Tirole, 1989:66, for a summary of related research).

**Bain-type IO: Firms as Output-restrainers**

Embedded in Bain-type IO is the view that the firm exists to restrain productive output through exercise of monopoly power or by colluding with other firms. Firms want to restrain output so that market price will be driven up; the successful firm's profit is the difference between an "artificially" high market price and its costs. From a social welfare perspective (the perspective of most economists), above-normal returns thus reflect nefarious firm behavior that occurs at the expense of consumers, in that above-normal earnings are returns to successful direct or indirect price-fixing.

Much attention in this school has been paid to the effects of firm size (absolute or relative) on earnings. The reasoning is that because large firms control substantial proportions of industry output, these firms are believed to have the greatest opportunity and incentive to engage in monopolistic or collusive practices. According to Weiss:

> The standard definition of monopoly power—"the power to set price or exclude competition"—is hardly precise, but it appears to require a persistent market share of well over half of some significant market, with no close challenger. The underlying assumption in this area is that such dominant firms will follow policies that misallocate resources (e.g., set prices well above cost) and redistribute incomes in favor of those in powerful positions. (1975: 184)

A major emphasis has been empirical testing of the "structure-conduct-performance" hypothesis articulated by Bain (1948, 1950, 1951, 1954). In this hypothesis, industry structure (e.g., number of sellers and buyers, product differentiation, barriers to entry, degree of fixed vs. variable costs, vertical integration) determines firm conduct (e.g., pricing, advertising), which in turn determines economic performance (e.g., social allocative efficiency and firm profitability) (Scherer, 1980: 4; Tirole, 1989: 1-2). Because firm conduct is assumed to be determined by industry structure, empirics most often have examined the association between structure and performance (Scherer, 1980: 267). Particular attention has been given to the relationship between firm size or industry concentration (taken as a proxy for the monopoly power of the largest firms in the industry) and profits, and later, between market share and profits (e.g., Gale, 1972; Mann, 1966; Hall & Weiss, 1967; Shepherd, 1972). The empirical results have been, however, less than conclusive, revealing at best a weakly positive association.

Regarding bounds on firm size and scope, Bain-type IO is not satisfied that diseconomies of scale, key limiters in perfect competition theory, are adequate to assure that firms will remain small enough to curtail nefarious activities. Thus the main limitation on size and scope in Bain-type IO is governmental intervention. In addition, collusion itself may act as a brake on firm size: collusive agreements, by
dampening competition, tend to lock-in the status-quo. The motivation for firm expansion in Bain-type IO is to increase monopolization or, alternatively, prevent another firm from gaining monopoly control. For example, vertical integration is seen as a method for extending monopoly power to downstream industries (see, e.g., McKenzie, 1951; Vernon & Graham, 1971), acquiring the firm’s own source of raw material so that the firm is not subject to another firm’s control over price, or erecting entry barriers (see, e.g., Adams & Dirlam, 1964; Comanor, 1967; Wallace, 1937). Similarly, advertising and product differentiation are viewed as ways to erect entry barriers and increase monopoly power (see, e.g., Comanor & Wilson, 1974).

It is useful to note that Bain-type IO encompasses a significantly richer concept of interfirn heterogeneity than does the perfect competition view. Although the perfect competition view predicts that, in equilibrium, no persistent performance differentials will exist, Bain-type IO holds the opposite. For example, as noted by Mann, “industries in which output is produced by a few dominant firms may, in the long run, earn higher rates of return...than...the normal or competitive rate of return” (1966:296). Persistent above-normal returns are based upon long-lasting though limited types of heterogeneity between firms: in Mann’s statement, the central heterogeneity is firm size (“dominance”). In other studies, the central heterogeneity examined is, for example, which side of an entry barrier the firm is on (e.g., degree of product differentiation, as examined by Shepherd [1972]), or differing market shares (e.g., Gale, 1972). Similarly, inherent in the concept of above-average industry returns is a subsidiary concept of firm heterogeneity, either between firms within the industry, or between industry incumbents and those blocked from entry.9

Despite ambivalent empirical results and serious questions concerning Bain-type IO’s theoretical underpinnings,10 the view of the firm embedded in Bain-type IO continues to have appeal.11 To understand why requires an appreciation of this school’s historical context, which centers to a substantial degree on concerns regarding “bigness” in business that first arose in the United States in response to the sudden emergence of large-scale enterprise—especially the pools and trusts—in the last half of the 19th century.

As analyzed by Chandler (1962, 1977), before the coming of large-scale, high-speed rail transportation and telegraph communication networks, developed between 1850 and 1880, the United States had a rural, agrarian economy perhaps most typically characterized by small, family enterprise.12 The coming of these networks led to almost astonishingly rapid revolutions in distribution and production, in which small enterprises were replaced by large.13

During this period of rapid growth, economists (and others) called attention to collusive/monopolistic practices in a wide range of industries.14 Two illustrative examples follow:

A pool in the meat packing industry was organized as early as 1885; and since that date pools of one kind or another have been maintained almost steadily. The pool of 1885 determined the quantity of meat that each member might ship, and by this means succeeded in exercising considerable control over the price of meat. In 1893 a more complete
and effective agreement was entered into. As the result of this agree-
ment the representatives of Swift and Company, Armour and Company,
and Morris and Company held weekly meetings, which were occasion-
ally participated in by representatives of the Cudahy Packing Com-
pany, and Hammond and Company. At these meetings, each of the com-
panies reported on its shipments into designated territories during the
previous week and on the prices received. These reports served as the
basis for the payment of fines for overshipments (40c. per 100 lbs.), and
for the allotment for the ensuing week. In order that full secrecy as to
the arrangements might be maintained and the consequences of legal
proceedings avoided, the parties to this agreement were designated by
certain letters of the alphabet rather than their real names. (Jones, 1922:
10-11)

and

In the tobacco industry an international pool providing for a division of
the field was effected. In September, 1902, the American Tobacco
Company (the American trust) and the Imperial Tobacco Company (a
British combination) entered into an agreement whereby the trade of
the United States, Cuba, Puerto Rico, Hawaii, and the Philippines, was
reserved to the American Tobacco Company, and the trade of Great
Britain to the Imperial Tobacco Company. (Jones, 1922:13)

Similar practices were documented in industries such as sugar, wire-nails, copper,
steel, shipbuilding, and oil (see, e.g., Jones, 1922; Moody, 1904; Ripley, 1905).
Such practices were the impetus for enactment of the Sherman Antitrust Act of
1890 and the Clayton Act of 1914\(^\text{15}\) and lie at the heart of Bain-type IO’s empirical
studies, the precursors of which sought to examine monopolistic practices through
“institutional” case studies of a longitudinal nature (see, e.g., Wallace’s 1937 study
of the aluminum industry).

Bain-type IO received a strong impetus from the Great Depression of 1929-
1941. Reigning economic orthodoxy—both macro- and micro-economic—was
undermined during this period, because it was seen as unable to explain the cause
or to help identify a solution to the terrible, worldwide economic crisis.\(^\text{16}\) Under the
orthodox view, a recession caused by inadequate demand for goods relative to sup-
ply would be solved by the market fairly quickly through reductions in prices. To
many contemporaneous observers, it seemed that the market simply did not work
any more. It appeared to them that prices must not be falling enough to equate de-
mand and supply at high output (and thus employment) levels. Further, it was pre-
sumed that prices \textit{would} fall enough if \textit{competition} reigned, but much of American
industry seemed to be dominated by huge corporations that could keep prices ar-
tificially high, thus resisting the ordinary law of supply and demand in order to
maintain their own profits, at great cost to the American economy. Indeed, Bain’s
(1951, 1954) early studies of his structure-conduct-performance hypothesis ana-
yzed data from the middle of the Great Depression. Thus what is called here Bain-
type IO presented itself when not only the market, but also the existing economic
theory of the (competitive) market seemed woefully inadequate. Of course, as we
moved farther from the decade of the 1930’s, Bain-type IO’s interest shifted to the
consequences of “market power” in an economy not mired in Depression. But the fundamental underlying Bain-type belief was unchanged—that control and deterrence of competition by firms with market power were the central economic forces in our economy, instead of firms competing over making at lower cost a product that consumers preferred. Public policy through the 1970s continued to embrace the Bain-type IO view of the firm, resulting in numerous initiatives to counter industrial concentration.  

*Schumpeter’s Response: A Focus on Dynamics, With Firms as Seekers of New Ways of Competing*

Mason (1951/1957a:91) notes that Schumpeter’s classic discourse on the “process of creative destruction” (1950: chaps. 7 & 8) is his “parting shot at antitrust policy”—in particular, his parting shot at those who connect large enterprise with nefarious means of profit-seeking.  

Schumpeter claims that the social value of various market structures—especially concentrated industries—cannot be assessed by employing a static view of competition, in which the focus is on how price is determined for products and processes that are unchanging. According to Schumpeter

>The essential point to grasp is that in dealing with capitalism we are dealing with an evolutionary process...The fundamental impulse that sets and keeps the capitalist engine in motion comes from the new consumers' goods, the new methods of production or transportation, the new markets, the new forms of industrial organization that capitalist enterprise creates...competition which commands a decisive cost or quality advantage and which strikes not at the margins of the profits and outputs of the existing firms but at their foundations and their very lives. (1950: 82, 83, 84)

Thus, in Schumpeter’s view, the purpose of firms is to seize competitive opportunity by creating or adopting innovations that make rivals’ positions obsolete: “[t]his kind of competition is as much more effective than [price competition over existing products] as a bombardment is in comparison with forcing a door”(1950:84).

Schumpeter, however, ties the scale and scope of the firm seeking radical innovation to the possession of monopoly power. In so doing, he seeks to reverse the presumption of Bain-type IO that monopoly power necessarily has negative welfare implications. Schumpeter points out that investing in radical innovation is inherently risky ("like shooting at a target that is not only indistinct but moving" [1950:88]), given the financial commitments required and the “ever-present threat” of competition, and so

restrictive [monopoly] practices may do much to steady the ship and to alleviate temporary difficulties...restrictions of this type [prices that seem predatory and restrictions of output] are...often unavoidable incidents, of a long-run process of expansion which they protect rather than impede. (1950: 87, 88)

Thus firms that can exercise monopoly power are seen by Schumpeter to have
greater incentive to develop revolutionary innovations. *Ex ante* market power provides the firm with the financial wherewithal to invest in risky innovation; the potential of *ex post* market power provides a powerful incentive to make these investments.\(^{20}\)

Schumpeter’s impact can be seen in Cohen and Levin’s (1989:1060) comment that, within IO, the number of empirical studies attempting to test aspects of Schumpeter’s theory is exceeded only by those examining industry concentration-profitability relationships. The subjects that have received the most attention are the relationship between industry concentration (again, a proxy for monopoly power) and innovation, and firm size and innovation. Cohen and Levin, agreeing with earlier reviews by Markham (1975), Kamien and Schwartz (1982), and Scherer (1980), report that such empirical investigations are beset by measurement and data problems and offer inconclusive results that “are perhaps most accurately described as fragile” (1989:1078). Much less attention has been given to exploring Schumpeter’s vision of the dynamics of competition, with Nelson and Winter’s (1982) innovative treatment representing perhaps the most comprehensive movement in this direction.

*The Chicago Response: A Renaissance of Price Theory, With Firms as Seekers of Production and Distribution Efficiencies*

The Chicago response emerged in part to blunt public policy initiatives arising from Bain-type IO, gaining its initial momentum from an effort to understand the economic rationale of legal stricture on business, and of business practices observed in the real world.\(^{21}\) The scholars of the Chicago tradition, as they saw it, reconfirmed through renewed application of price theory the efficacy of market mechanisms in achieving welfare-enhancing ends.\(^{22}\) The theory of the firm implicit in the Chicago tradition is that firms exist to enhance efficiency in production and distribution. In the Chicago view, however, this objective is, from the firm’s point of view, only second best. As described in Stigler’s seminal article:

> The present paper accepts the hypothesis that oligopolists wish to collude to maximize joint profits....The combined profits of the entire set of firms in an industry are maximized when they act together as a monopolist. (1964/1968d:39)

However, Stigler goes on to point out that effective collusion requires often prohibitively costly monitoring and enforcement, given that each party has an incentive to “chisel” on the agreement. Because of these high costs, the Chicago school holds that effective collusion is not likely to persist (and hence not likely to emerge): thus observed large size and above-normal returns must be due to a firm’s efficiency differential in production and distribution, in comparison to rivals.

The Chicago tradition applies central concepts of neoclassical price theory—particularly those of profit maximizing behavior and competition—while relaxing several critical assumptions of the perfect competition model. As Demsetz notes,

> “[a]ll the conclusions [of the Chicago school] derive from the attempt

*JOURNAL OF MANAGEMENT, VOL. 17, NO. 1, 1991*
of maximizers to overcome certain kinds of costs impediments...That approach was contrary to what was then the general approach in the literature, which was, every time you saw something that was peculiar in terms of the framework of the perfect competition model, to mystically conjure up the word monopoly and stop the analysis right there. (Kitch, 1983: 204)

And Aaron Director comments with respect to the return to efficiency explanations of economic behavior that the Chicago approach "wasn't a new way of looking at the world, only an old way of looking at the world" (Kitch, 1983:201).

A central assumption within the Chicago tradition is that of costly information:

One should hardly have to tell academicians that information is a valuable resource: knowledge is power. And yet it occupies a slum dwelling in the town of economics...I hope to show that some important aspects of economic organization take on a new meaning when they are considered from the viewpoint of the search for information...[J]ust as an analysis of man's shelter and apparel would be somewhat incomplete if cold weather is ignored, so also our understanding of economic life will be incomplete if we do not systematically take account of the cold winds of ignorance. (Stigler, 1961/1986b: 171, 188)

Acknowledgment of information costs has led, for example, to an attack on the Bain-type IO view that advertising is a (contrived) entry barrier erected to create and preserve monopoly power. In the Chicago view, advertising serves a valuable economic function of providing information to consumers, reducing the cost of searching out relevant prices (Stigler, 1961/1968b), and of ascertaining the quality of goods (Nelson, 1974). As Demsetz comments,

What does it mean to say that advertising expenditures and capital outlays constitute barriers to entry? One meaning is that firms must make such expenditures if they are to produce and communicate about the commodities they hope to sell. Such expenditures are no more barriers than are expenditures on labor and material. A second meaning is that existing firms are more efficient in the employment of such inputs than are firms not yet in the industry. If this is so, the existing firms deserve applause, not divestiture. (1975: 173)

The Chicago school has offered efficiency-based explanations for a number of other practices believed within the Bain-type IO tradition to be monopolistic. For example, the degree of vertical integration has been related to the efficiency with which outside suppliers can or cannot supply a firm over the various stages in the product life cycle (Stigler, 1951/1968a). And even the need for governmental intervention in the regulation of "natural monopolies" such as public utilities has been questioned. For example, Demsetz (1968) argues that ex post regulation can be replaced by an ex ante bidding process for rights to own the monopoly: with "competition for the field," there is no reason for prices subsequently charged to consumers to rise above the competitive level.

In the Chicago view, size and scope of the firm are determined by its efficiency.
If the firm can make efficiency gains, it will grow; if not, it will shrink, as competitors erode the firm's source of advantage. As put by McGee:

> We live in a complex and uncertain world. Individuals are not all alike. The teams that make up business firms are not alike, and the effectiveness of firms differs. In the usual textbook analysis of economies of size, costs are determined by a firm's size. In the real world, it is often the other way around: the size of a firm is determined or strongly influenced by the cost level [efficiencies] that it manages to achieve.

(1975:101)

As indicated by this passage, underlying the differential efficiency view is recognition of nonhomogeneous inputs and forces impeding mobility of inputs. At times, however, the Chicago view has seen efficiency gains as resulting primarily from achievement of scale economies. This position emerges most clearly in Chicago's direct attacks on Bain-type IO and its policy prescriptions (see, e.g., McGee, 1975).

An important note regarding the Chicago perspective is the paramount role of entry of new competitors in imposing an efficiency imperative on incumbent firms and on determining long-run earnings potential. Although this view holds that efficiency-based earnings "need not be eliminated soon by competition" (Demsetz, 1973a: 2), in the long run imitative entry will drive a firm's profits to zero. For example, as noted by Stigler, "[i]f the discussion of entry barriers suggests that there are effective permanent obstacles to entry in many industries, it is misleading" (1966:227). This point in turn underscores the general focus of the Chicago school on obtaining efficiency differentials in producing or selling current (rather than new, innovative) products.

**Coase/Williamson Transaction Cost Economics: Firms as Avoiders of the Costs of Market Exchange**

Ronald Coase pioneered the notion that firms and market exchange are alternative methods for coordinating production: "the distinguishing mark of the firm is the supersession of the price mechanism" (1937/1952:334). Therefore, the question to be addressed in understanding the existence of firms is "the basis on which, in practice, this choice between alternatives is effected" (1937/1952:335).

At the core of Coase's analysis is the observation that "the operation of a market costs something and by forming an organization and allowing some authority (an "entrepreneur") to direct the resources, certain marketing costs are saved" (1937/1952: 338). The cost of negotiating contracts for inputs was seen as particularly important:

> [a] factor of production (or the owner thereof) does not have to make a series of contracts with the factors with whom he is co-operating within the firm, as would be necessary, of course, if this co-operation were as a direct result of the working of the price mechanism. For this series of contracts is substituted one.

(1937: 336)
Thus firms exist to avoid (economize on) the costs of conducting the same exchange between autonomous contractors; applying marginal analysis, “a firm will tend to expand until the costs of organizing an extra transaction within the firm become equal to the costs of carrying out the same transaction by means of an exchange on the open market or the costs of organizing in another firm” (Coase, 1937/1952: 341).

Though maintaining Coase’s view that firms exist to avoid the costs of using the market’s price mechanism, Williamson (1975, 1989) expands these insights by analyzing situations in which transaction cost avoidance by firms can be particularly acute. Williamson gives particular attention to exchanges in which opportunistic potential is significant. Such potential exists when three conditions pertain simultaneously: asset specificity, small numbers of potential transactors, and imperfect information. Asset specificity imposes a condition of dependence upon the owner of the specific asset, call it A, because A’s value depends on the presence of another input, call it B, to which the former is specific. Small numbers reinforces this dependence, because A cannot costlessly find a replacement for B, should B’s services be withdrawn. And imperfect information means that (a) complete contingent contracts cannot be written and (b) with incomplete contracts, a priori knowledge of B’s later actions cannot be fully incorporated in determining A’s ex ante hire-price. Thus A cannot nullify the risk of later opportunism by B by ex ante adjustment of the price for A’s services. Using the assumption that within firms “a more nearly joint profit maximizing attitude and result is to be expected” (1975:29), Williamson concludes that firms will exist when opportunistic potential is significant, in order to decrease this potential; when asset specificity, small numbers, and imperfect information are not significant, transactions between autonomous contractors will dominate.

Within Williamson’s framework, size and scope of the firm turn, as they did with Coase’s original specification, on the savings or increases in cost from housing a transaction within the firm, as opposed to obtaining the same input services by market contract. Examples of size and scope issues that have been explored within this framework include the extent of internal hierarchy, vertical integration, and diversification (conglomerateness) (see, e.g., Klein, Crawford, & Alchian, 1978; Ouchi, 1980; Walker & Weber, 1984, 1987; Williamson, 1972, 1975).

Klein and Leffler (1981), and Williamson (1983), introduce a significant extension of Williamson’s original framework. Klein and Leffler argue that existence of opportunistic potential is not sufficient to cause resources to be owned jointly (a firm); for example, in a bilateral situation, if A and B each face the potential for opportunistic “hold-up” from the other of approximately the same magnitude, then these potentials are offsetting and the parties can maintain a stable relationship as autonomous contractors. Further, Klein and Leffler argue that autonomous contractors will take explicit steps toward structuring their relationship so that market relations can continue, in order to avoid diseconomies in management; these steps may include the implicit or explicit posting by each party of “performance bonds,” the value of which will be lost if that party reneges on the agreement. This view has led to new interpretations of business practices previously labeled as mo-
nopolistic, including franchise tying contracts (Klein & Saft, 1985), exclusive territories and resale price maintenance (Klein & Murphy, 1984), and block booking (Kenny & Klein, 1983).

Distinctiveness of Resource-Based Theory

The first part of this section draws from the strategy literature to outline major features of the resource-based approach, comparing these to the IO-related schools of thought discussed above. The second part discusses the resource-based view as a new theory of the firm.

Central Elements of Resource-based Theory: Firms as Seekers of Costly-to-Copy Inputs for Production and Distribution

As in the IO-related theories discussed above, the firm’s ultimate objective in a resource-based approach generally is taken to be above-normal returns (e.g., Barney, 1986; Wernerfelt, 1984). In a resource-based view, obtaining such returns requires either that (a) the firm’s product be distinctive in the eyes of buyers (e.g., the firm’s product must offer to consumers a dissimilar and attractive attribute/price relationship, in comparison to substitutes), or (b) that the firm selling an identical product in comparison to competitors must have a low-cost position.32

Thus in a resource-based view, the critical problem faced by the firm is how to maintain the distinctiveness of its product, or, for identical products, its low cost position, while not investing so much in obtaining this difference as to destroy above-normal returns (see, e.g., Barney, 1986). Distinctiveness in the product offering or low costs are tied directly to distinctiveness in the inputs—resources—used to produce the product, much as the quality and cost of boeuf bourguignon depend on the particular ingredients used and the way in which they are mixed. The neoclassical view of the firm as input-combiner is, therefore, at the heart of resource-based theory, although the resource-based view does not include the accompanying neoclassical assumptions of a freely-available and perfectly specifiable production algorithm, as well as costless resource mobility across uses and infinite input divisibility. Also central to resource-based theory is the view, key to Bain-type IO, that persistent above-normal earnings are possible. A resource-based approach, however, views such earnings as rents accruing to uncopyable assets (Rumelt, 1987), rather than Bain-type IO’s exercise of monopoly (or collusive) output restraints. Resource-based theory recognizes, as argued by Schumpeter, the power of the revolutionary innovation to shift market positions (Prahalad & Hamel, 1990; Rumelt, 1974) yet resource-based theory rejects the necessity of pre-existing monopolistic earnings to support such initiatives and does not find inconsistent the view that less than revolutionary innovations, well protected by “resource barriers” (Wernerfelt, 1974), can yield above-normal returns. Like the Chicago view, and unlike Bain-type IO, resource-based theory sees these returns as resulting primarily from the acumen or luck of the firm in acquiring, combining, and deploying resources (conduct), rather than from the structure of the industry in which the firm finds itself. These and other comparisons (discussed below) of a resource-based view with its IO-related predecessors are summarized in Table 1.

Thus in a resource-based view, the firm is a seeker of unique, or otherwise
Table 1
Comparison of Resource-Based Theory to Five IO-related Predecessors

<table>
<thead>
<tr>
<th></th>
<th>Similarities</th>
<th>Distinctions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Neoclassical</strong></td>
<td>◆ Firm as input-combiner: emphasizes physical production of goods or services</td>
<td>◆ No “given” production algorithm; identification of resources and resource combinations is problematic</td>
</tr>
<tr>
<td></td>
<td></td>
<td>◆ Critical resources may be immobile (not available for purchase, or not easily jettisoned if no longer productive); may be by-products of teamwork</td>
</tr>
<tr>
<td></td>
<td></td>
<td>◆ Firm size and scope are important issues</td>
</tr>
<tr>
<td><strong>Bain-type IO</strong></td>
<td>◆ Firm’s environment (other firms/public policy) poses critical constraints on strategy</td>
<td>◆ Restraints on output through monopolistic or collusive action, or investment in “artificial” entry deterrence, are not primary sources of persistent above-normal returns</td>
</tr>
<tr>
<td></td>
<td>◆ Persistent above-normal returns are possible</td>
<td>◆ The firm (not the industry) is the appropriate unit of analysis for understanding sources of above-normal returns</td>
</tr>
<tr>
<td></td>
<td></td>
<td>◆ The internal organization of firms is a critical variable</td>
</tr>
<tr>
<td></td>
<td></td>
<td>◆ Firms’ behavior may be at least as much the result of conscious choice as it is a foregone conclusion from industry structure</td>
</tr>
<tr>
<td><strong>Schumpeter</strong></td>
<td>◆ Spectacular above-normal returns can result from new ways of competing</td>
<td>◆ Feasibility of new ways of competing does not rest on monopolistic (output-restraining) practices</td>
</tr>
<tr>
<td></td>
<td>◆ Entrepreneurial vision is at the heart of the firm</td>
<td>◆ Imitators are constrained by costly-to-copy resources</td>
</tr>
<tr>
<td></td>
<td>◆ Potential imitators always exist</td>
<td>◆ Exogenous shocks can be critical to “creative destruction”</td>
</tr>
<tr>
<td></td>
<td></td>
<td>◆ Healthy earnings can result from less than “revolutionary” innovation</td>
</tr>
<tr>
<td><strong>Chicago</strong></td>
<td>◆ Firms are production and distribution efficiency-seekers</td>
<td>◆ Focus more on the intermediate (not long) term, so entry need not dissipate above-normal returns in the time span relevant to the firm and its strategic choice problem</td>
</tr>
<tr>
<td></td>
<td>◆ Size and scope of the firm reflect extent to which production and distribution efficiencies are achieved</td>
<td>◆ Efficiency seeking goes beyond current products, extending also to new products</td>
</tr>
<tr>
<td><strong>Coase/Williamson</strong></td>
<td>◆ Asset specificity and small numbers are critical concepts constraining the firm’s strategic options</td>
<td>◆ The heart of the firm centers on deployment and combination of specific inputs rather than on avoidance of opportunism</td>
</tr>
<tr>
<td><strong>Transaction Costs</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Costly-to-copy inputs. Hence primary attention within this framework has been drawn to how managers might discern which inputs are likely (a) to have productive value in excess of their hire-price, as well as (b) to incorporate characteristics that render these inputs costly for rivals to copy. Expressed differently, primary attention has been given to identification of (a) inputs likely to generate rents, and (b) of these, characteristics of the (subset of) inputs to which long-lived rents may accrue. Discussion of these two topics composes the remainder of this subsection.

*Inputs able to generate rents.* In a resource-based view, discerning appropriate inputs is ultimately a matter of entrepreneurial vision and intuition; the creative act underlying such vision is a subject that so far has not been a central focus of re-
source-based theory development. It has long been recognized within the strategy field, however, that important constraints exist that bound the types of inputs able to generate rents. These constraints are of two types—those external and those internal to the firm.

1. External Constraints on Inputs Able to Generate Rents: Underpinning the analytic framework advanced by Porter (1980, 1985) are three sources of external constraints that affect the ability of inputs, when employed in a particular situation, to generate rent. These include (a) conditions of demand relevant to the product, (b) public policy, and (c) competitor action. Regarding (a), the metal lead, for example, is not likely to be a valuable input for forming hot-air balloons, because such balloons are bought ("demanded") in order that they might float. Similarly, investing in ways in which to make non-standard-sized bed linens may not be productive, because consumers gain real benefits from size standardization. Turning to (b), public policy constraints may involve, for example, regulatory provisions, antitrust action, and laws governing establishment of property rights such as patents and copyright. A resource-based perspective is similar to Bain-type IO and the Chicago tradition in that both of these types of external constraints on strategy are recognized. Regarding (c), insofar as substitute products exist, competitor action effectively imposes efficiency requirements on the productivity of inputs and input combinations, even when the firm’s market offering is distinct from these substitutes. If the firm’s costs become too great, driving up its price, sales will be lost to competitors. Resource-based theory here incorporates the Chicago tradition’s emphasis on efficiency seeking in production and distribution in response to competitive threats. Bain-type IO looks instead to the deployment of the firm’s resources to deter competition, to co-opt it through collusion, or to destroy it through below-cost predatory pricing.

It is useful to note that in the strategy literature, Schumpeterian-type revolution has been seen as playing a part in the dynamics of external constraints, which occurs when innovative activity creates "new games" (Barron, 1981) that radically change demand characteristics (through, e.g., introduction of new products), or the efficiencies open to firms (e.g., major process or organizational advances). However, in the resource-based tradition it also has been pointed out that, from the perspective of the firm, radical changes in competitive position also can result from exogenous "luck" factors, such as a change in consumer tastes or regulatory requirements (Rumelt, 1984).

2. Internal Constraints on Inputs Able to Generate Rent: A second set of constraints on what can constitute rent-producing inputs concerns the past history ("resource endowment") of the firm itself. This constraint is reflected in the view that the firm is a "bundle of linked" (italics added) and idiosyncratic resources and resource conversion activities" (Rumelt, 1974: 561). Leaving aside for the moment the issue of idiosyncracy, we must ask what it means for resources to be internally linked, and the relationship of such linkages to above-normal returns.

At least one sense of linkedness is value dependency, or, in other words, asset specificity of the Alchian/Williamson type. If the value of $A$, for example, increases in the presence of $B$, but the value of $C$ is independent of either $A$ or $B$, then
A grows in value when it is teamed with $B$, but $C$ does not. This growth in value is inextricably linked to rent earning; that is, to creation within the firm, through its resource-combining function, of true productive value that would not exist were $A$ kept separate from $B$. Given, however, Barney's (1986) argument regarding the impact of resource costs, it cannot be presumed that rents will accrue from the fact of input specificity alone.

To illustrate, consider firms $X$ and $Y$, each of which would like to acquire one unit of input $A$. Initially suppose that $A$ is potentially specific only to $X$ and $Y$, but not to other firms. Assume that two homogeneous units of $A$ are available. Let $v(X)$ and $v(Y)$ represent the value that $A$ would add to $X$ and $Y$, respectively, and let $v(N)$ represent the value of $A$ to firms for which $A$ is nonspecific. Table 2 represents six possible scenarios. Note that with two homogeneous units of $A$, the price of $A$ is the value of $A$ to the firm(s) putting the next-to-highest value on acquiring $A$ (i.e., the value to the marginal user).

Now consider the same situation when there is only one unit of $A$ available. If $A$ is specific to $X$ and/or $Y$, then the owner of $A$ will have bargaining power and the price of $A$ will fall between the values of $A$ to the two highest-valuing firms (assuming no collusion among buyers). Let $P_m$ be the price of $A$ when $A$ is unique; Table 3 represents six possible scenarios.

Finally, consider the same situation when there are more than two units of $A$.

### Table 2
Rent from Asset $A$, Two Units of $A$ Available

<table>
<thead>
<tr>
<th>Regime</th>
<th>Specific to $X$</th>
<th>Specific to $Y$</th>
<th>Relative Values of $A$ to $X$, $Y$</th>
<th>Price of $A$</th>
<th>Rent to $X$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>No</td>
<td>No</td>
<td>$v(N) = v(X) = v(Y)$</td>
<td>$v(N)$</td>
<td>$v(N) - v(N) = 0$</td>
</tr>
<tr>
<td>2</td>
<td>No</td>
<td>Yes</td>
<td>$v(N) = v(X) &lt; v(Y)$</td>
<td>$v(N) - v(N)$</td>
<td>$0$</td>
</tr>
<tr>
<td>3</td>
<td>Yes</td>
<td>No</td>
<td>$v(X) &gt; v(Y) = v(N)$</td>
<td>$v(X) - v(N)$</td>
<td>$0$</td>
</tr>
<tr>
<td>4</td>
<td>Yes</td>
<td>Yes</td>
<td>$v(X) &gt; v(Y) &gt; v(N)$</td>
<td>$v(Y)$</td>
<td>$v(Y) - v(X) &gt; 0$</td>
</tr>
<tr>
<td>5</td>
<td>Yes</td>
<td>Yes</td>
<td>$v(X) = v(Y) &gt; v(N)$</td>
<td>$v(Y)$</td>
<td>$v(Y) - v(X) = 0$</td>
</tr>
<tr>
<td>6</td>
<td>Yes</td>
<td>Yes</td>
<td>$v(N) &lt; v(X) &lt; v(Y)$</td>
<td>$v(X)$</td>
<td>$v(X) - v(X) = 0$</td>
</tr>
</tbody>
</table>

### Table 3
Rent from Asset $A$, One Unit of $A$ Available

<table>
<thead>
<tr>
<th>Regime</th>
<th>Specific to $X$</th>
<th>Specific to $Y$</th>
<th>Relative Values of $A$ to $X$, $Y$</th>
<th>Price of $A$</th>
<th>Rent to $X$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>No</td>
<td>No</td>
<td>$v(N) = v(X) = v(Y)$</td>
<td>$P_m = v(N)$</td>
<td>$v(N) - v(N) = 0$</td>
</tr>
<tr>
<td>2</td>
<td>No</td>
<td>Yes</td>
<td>$v(N) = v(X) &lt; v(Y)$</td>
<td>$P_m &lt; v(Y)$</td>
<td>$0$</td>
</tr>
<tr>
<td>3</td>
<td>Yes</td>
<td>No</td>
<td>$v(X) &gt; v(Y) = v(N)$</td>
<td>$P_m &lt; v(X)$</td>
<td>$v(X) - P_m &gt; 0$</td>
</tr>
<tr>
<td>4</td>
<td>Yes</td>
<td>Yes</td>
<td>$v(X) &gt; v(Y) &gt; v(N)$</td>
<td>$P_m &lt; v(X)$</td>
<td>$v(X) - P_m &gt; 0$</td>
</tr>
<tr>
<td>5</td>
<td>Yes</td>
<td>Yes</td>
<td>$v(X) = v(Y) &gt; v(N)$</td>
<td>$P_m = v(Y)$</td>
<td>$v(X) - P_m = 0$</td>
</tr>
<tr>
<td>6</td>
<td>Yes</td>
<td>Yes</td>
<td>$v(N) &lt; v(X) &lt; v(Y)$</td>
<td>$P_m &lt; v(Y)$</td>
<td>$0$</td>
</tr>
</tbody>
</table>
The price of $A$ will equal $v(N)$; both $X$ and $Y$ each always will get a unit of $A$. Table 4 represents six possible scenarios.

Tables 2-4 lead to several observations. First, $X$ earns no positive rent when $A$ is not specific to (i.e., linked with) its existing resource base. Second, specificity of $A$ is alone an insufficient condition for generating rent unless $X$'s demand does not affect $A$'s price, as demonstrated in regimes 5 and 6 of Tables 2 and 3. In particular, in regimes 3 and 4 of Tables 2 and 3, rent accrues to $X$ from employment of $A$ only if (a) $A$ is specific to $X$, and (b) $A$ is more specific to $X$ than to the marginal purchaser of $A$, in the sense of $A$ having greater productive value when linked to $X$'s asset base than to the marginal purchaser's. The situation is, however, somewhat different when $A$ is not relatively scarce to $X$ ($X$'s demand does not affect $A$'s price), as shown in Table 4. Perhaps surprisingly, Table 4 reveals that the opportunities for $X$ to earn rent increase if the supply of $A$ is plentiful, and most firms do not find $A$ to be specific. In this situation, $X$ can earn rent even if it is not the highest-valuing acquirer of $A$ (although, all else equal, $X$'s rent will be less than that of a higher-valuing acquirer).

This simple model easily is generalized to a setting in which many firms find $A$ to be specific. If $A$ is specific to firms' asset bases in differing degrees, rent can be earned by any firm for which $A$'s productive value exceeds its price, i.e., any inframarginal purchaser of $A$. However, the greater the degree of $A$'s specificity to a firm, the greater is $A$'s rent potential to that firm. Thus the linkage of an input to the firm's existing asset base—and especially, the relative strength of that linkage as compared to that which competitors would achieve—is central to rents, and hence constitutes a constraint on (and an opportunity for) the input investments of firms in this situation. This resource-based view is different both from Chicago's emphasis on asset acquisition as opening up opportunities for scale economies, and Bain-type I0's view of asset acquisition as a means of gaining market power over price through size or of overawing potential entrants through capital barriers or implicit threats of predation.

There is, of course, another situation in which rents can be earned. Firms can acquire vital inputs before competitors attempt to obtain them; due to lack of competition, the early firm obtains a first-mover advantage manifested as a lower input hire-price, as compared to the price that later-acting competitors must pay

<table>
<thead>
<tr>
<th>Regime</th>
<th>Specific to $X$</th>
<th>Specific to $Y$</th>
<th>Relative Values of $A$ to $X, Y$</th>
<th>Price of $A$</th>
<th>Rent to $X$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>No</td>
<td>No</td>
<td>$v(N) = v(X) = v(Y)$</td>
<td>$v(N)$</td>
<td>$v(N) - v(N) = 0$</td>
</tr>
<tr>
<td>2</td>
<td>No</td>
<td>Yes</td>
<td>$v(N) = v(X) &lt; v(Y)$</td>
<td>$v(N)$</td>
<td>$v(N) - v(N) = 0$</td>
</tr>
<tr>
<td>3</td>
<td>Yes</td>
<td>No</td>
<td>$v(X) &gt; v(Y) = v(N)$</td>
<td>$v(N)$</td>
<td>$v(X) - v(N) &gt; 0$</td>
</tr>
<tr>
<td>4</td>
<td>Yes</td>
<td>Yes</td>
<td>$v(X) &gt; v(Y) &gt; v(N)$</td>
<td>$v(N)$</td>
<td>$v(X) - v(N) &gt; 0$</td>
</tr>
<tr>
<td>5</td>
<td>Yes</td>
<td>Yes</td>
<td>$v(X) = v(Y) &gt; v(N)$</td>
<td>$v(N)$</td>
<td>$v(X) - v(N) &gt; 0$</td>
</tr>
<tr>
<td>6</td>
<td>Yes</td>
<td>Yes</td>
<td>$v(N) &lt; v(X) &lt; v(Y)$</td>
<td>$v(N)$</td>
<td>$v(X) - v(N) &gt; 0$</td>
</tr>
</tbody>
</table>
(Rumelt, 1984: 567; Wernerfelt, 1984: 173). Here, the early firm earns rent on its luck or brilliance (Barney, 1986); notably, the inputs involved need not be specific to the firm's asset base.

With these points in mind, turn to Dierickx and Cool’s intriguing suggestion within the context of the resource-based literature that purchasable assets cannot be sources of long-lived rents, because these assets can be traded in the market: “firms may as well realize the value of their asset bundles through the relevant factor markets instead of deploying them in product markets” (1989a: 1505). The foregoing analysis suggests that when rivals compete over market acquisition of a scarce resource, the rent potential depends upon the input being more specific to the firm (i.e., generating greater productive value) than it is to a firm for which the input's price just equals its value-in-use. Thus firm X that, for example, earns rent because of input A's above-average degree of specificity to its asset base cannot hope to sell A at the market price and receive the equivalent of A’s productive value to X (as in scenarios 3 and 4 of Tables 2, 3, and 4). Hence in this situation a purchasable input generates rent. However, firm Y, to which A is not so specific, may be able to break-even or even gain from selling A, if A's value-in-use to Y is equal to or less than the going price. Even this may not be true, however, if it is in Y’s interest to act strategically in this situation. Y may want to keep A from X if X and Y are rivals and X's acquisition of A would help X significantly and thus put Y at a competitive disadvantage. Because of the bilateral bargaining implicit in this situation, Y will not be able to extract all of X's added rents from A through the negotiated price of A. Furthermore, if X's gain from A is less than Y's loss from X's new competitive advantage over Y (as might be the case if X and Y are not the only firms presently competing in the product market), then Y will not sell A to X even if Y could somehow extract all of X's gain from having A. This analysis can be extended to the situation in which the firm earns rent on a first-mover advantage, even if the early-acquired input is non-specific. If the input was acquired cheaply enough, the purchased asset can generate rent. Further, if the firm continues to require the input, there is no reason to sell it (to realize the first-mover gain) and then re-buy the input at the going price.

Perhaps a better way to interpret Dierickx and Cool’s argument is that inputs that cannot be purchased, such as learning-by-doing and organizational culture, are, on average, likely to be more specific to the firm than purchasable inputs and hence have the potential to be the more significant rent-generators. This view is consistent with the large amount of attention that intangible assets have received in the resource-based literature. For example, Prahalad and Hamel argue vividly for explicit attention to be given to the core competencies of the firm, which involve “the collective learning in the organization, especially how to coordinate diverse production skills and integrate multiple streams of technology” (1990: 82). A focus on non-purchasable, intangible assets also is reflected in, for example, Teece, Pisano, and Shuen’s emphasis on “the mechanisms by which firms learn and accumulate new skills and capabilities” (1990: 11); Winter's (1987) discussion of knowledge and competence as key strategic assets; Itami’s (1987: 12) treatment of “invisible assets,” which include “consumer trust, brand image, control of distribution, corporate culture, and management skill”; and Barney’s (1986) em-
phasis on the firm’s detailed knowledge of its internal capabilities. The value of intangible assets is recognized in the Chicago tradition's focus on information as a factor of production and reputation as a (efficiency-based) barrier to entry and by Williamson (1975) as an area in which opportunistic potential can exist.

*Input characteristics associated with long-lived rents.* First, take up the situation in which rent is generated by a first-mover advantage in a purchasable input (where the input need not be specific to the firm). Persistence of rent in this situation depends centrally on the duration of the contract engaging the input's services and the costs of enforcing the contract. For example, a firm fortunate enough to purchase a mine long ago at a low price can be assured that (so long as the ore supply holds out) its advantage in lower ore costs will persist.43 However, the returns of another firm that long ago obtained a lease for a prime corner restaurant location are in jeopardy once that lease expires. And a third firm that is organized around bringing to market the creations of a uniquely inventive scientist may earn above-normal returns so long as the scientist can be persuaded to stay and not renegotiate his or her contract to capture these rents. Less would be earned should the scientist leave, say, to begin his or her own firm or work for another.

Turn next to consideration of factors contributing to the persistence of rents from inputs that are specific to the existing asset base of the firm. Such factors are well discussed elsewhere and thus are treated only briefly here. Rumelt advances the central concept of *isolating mechanisms*, which are “phenomena that limit the ex post equilibration of rents among individual firms” (1984:567). In resource-based theory, such a firm-level concept appears an appropriate alternative to Bain-type IO’s focus on entry barriers at the industry level. Dierickx and Cool (1989a: 1507-9) advance a useful analysis of mechanisms that render inputs costly to copy. These include time compression diseconomies to imitation, making it more difficult for a newcomer to catch up by simply “throwing money” at acquiring or developing the input; increasing returns to the cumulative magnitude of the stock of the input, so that firms already possessing a substantial stock can add to it at a lower cost than can a newcomer; and a slow decay rate of the input (e.g.,, team embodied knowledge that can be passed on without significant knowledge degradation to successive “generations” of team members). These mechanisms underlie some of the economies and diseconomies of scale important to Chicago theory. Other factors noted by Dierickx and Cool include specificity of the input to *multiple* other inputs in the firm’s asset base (making it difficult for other firms with different asset bases to use the input as effectively) and causal ambiguity (Lippman & Rumelt, 1982) regarding which inputs are causing observed outcomes. In addition, Rumelt (1984) notes that persistent rents can result from legally imposed isolating mechanisms, such as those imposed by regulatory requirements, or patent, copyright, and trademark law;44 isolating mechanisms also can stem from incomplete information on the part of consumers, reflected in, for example, search and switching costs that vary across firms’ offerings.45

In summary, as Table 1 illustrates, a resource-based perspective both incorporates and rejects major elements of each of the IO-related alternatives discussed.
earlier. Thus the resource-based perspective both reflects a strong IO heritage and entails a way of viewing the firm that is different from any one of these IO-related approaches.

**Resource-Based Theory as a New Theory of the Firm**

As noted earlier, the primary requisites of a theory of the firm are to explain the reasons for existence of the firm, and what limits its size and scope. Looking at the question of existence, Coase (1937/1952) and Williamson (1975) pose an important test: for the existence of a firm to make sense in some business activity, the firm, with its internal network of relationships, must outperform the alternative, in which all such relationships are external, arms-length, market transactions. Because resource-based theory is centrally motivated by understanding performance differentials *between* firms, an additional "test" seems appropriate. For resource-based theory, existence needs to be explained in terms of a firm's superiority to two alternative forms of organization: a collection of market contracts *and* other firms. By the latter, the intention is to raise the issue of why a particular firm exists, as opposed to its assets being distributed among other firms.

To begin, consider the existence of a firm in comparison to other firms. As with the previous analysis, firms are seen as possessing heterogeneous asset bases (resource endowments); as put by Rumelt, "a firm's competitive position is defined by a bundle of unique resources and relationships" (1974:557). Thus for any given input, the degree to which it will be specific varies across firms because firms themselves are unique. Therefore, an input may be owned (or developed) by firm X instead of by firm Y because this input may yield greater rent when linked to the asset combination that is X, rather than to that which is Y. In this view, multiple, heterogeneous firms continue to exist because the assets with which they will come to be mated are themselves heterogeneous, each making a better fit with (more specific to) some firms than with others. And it is seen to be impractical for any one firm to maintain within itself the spectrum of heterogeneities found in the several firms, as the example of organization culture helps demonstrate.

In contrast, relatively little attention has been given thus far in the resource-based literature to the question of why the firm exists, in comparison to the same business activity being carried on by autonomous input owners via contingent claims or spot contracts. It is useful to examine whether resource-based theory is identical to transaction cost theory in analyzing the question of the firm's existence versus a collection of market contracts. As discussed above, transaction cost theory (Williamson, 1975) emphasizes the existence of firms as a means to lessen the opportunistic potential that arises when team-specific investments are made. Is mitigation of opportunistic potential also the heart of the firm in resource-based theory?

A striking characteristic of much of the resource-based literature is (a) acknowledgment of the importance of asset specificity (and attendant small numbers) but (b) little discussion of avoidance of opportunistic potential as the central activity or purpose of the firm. Put differently, instead of viewing the firm as an "avoider of a negative," the resource-based literature tends to see the firm as the "creator of a positive," as creator of unique productive value (see, e.g., Prahalad &
Hamel's (1990) discussion of core competencies and core products). The literature's direction suggests a need to ascertain whether a case can be made that firms (as compared to collections of market contracts) exist for reasons primarily related to "creating positives," with or without opportunistic considerations. This the issue addressed below.

To encapsulate the argument to follow, the case made here is that, under certain circumstances, firms have advantage over market relationships in the joint activity of creating and redeploying specific capital. Further, the advantage of firms in the creation-redeployment combination need not stem from an opportunism-control advantage. Instead, it is argued that the firm has advantage over a collection of market transactions in those situations where redeployment inside the firm is more efficient and, perhaps more important, qualitatively more productive because of the opportunity to benefit from asset interdependencies within the firm.

First looking at the question of firm scope, consider an existing firm that undertakes a research project to create a product embodying advances in technology to be developed in the course of the project. The project may be undertaken (a) strictly or partly by an in-house team, or (b) wholly by an outside team on a contract basis.

If the research project has the potential to be closely linked with (specific to) the firm's existing operations, option (a) will be superior to (b) because an in-house team is likely to produce technological knowledge, skill, or routines that fit better with the firm's current activities (i.e., to produce new assets highly specific to the firm's existing assets). In addition, the value of the research project may come as much from the process of doing the research as from the end product. The full know-how gained from the research process is not likely to be transmittable as a discrete lump; rather, the research investment may generate greater gains when the possessor of that know-how is continuously and intimately engaged in applying the know-how to the manufacture of the new or existing products, the design and manufacture of other products, or the undertaking of follow-on research. That is, when the new project is related closely to the firm's other activities, the greater gains are more likely to be generated when the know-how possessor is an employee. Furthermore, an employee familiar with the firm's culture, routines, etc. should be better able than an outsider to transmit/translate such know-how in a way that aids implementation elsewhere in the firm. Put another way, proceeding by option (a) not only increases the likelihood of generating redeployable resources, but also makes more efficient the actual redeployment.

Note that gains from in-house activity are likely to arise whether or not there is a conscious orientation of the research in ways most likely to generate value-producing redeployment, or a conscious redeployment of the new "resources" by the firm to its existing base, though the size of the gains may well be increased by such strategies. Rather, the existence of gains is likely to result automatically from the underlying linkedness of the firm's resources and resource conversion activities; in our case, the linkedness in orientation and redeployment arises from the inside researcher's broader knowledge of and involvement in the activities of the firm.

It follows from this analysis that the scale and scope of the firm, in comparison to the alternative of a collection of market transactions, depends critically on the
degree to which new undertakings actually are specific to the firm’s existing asset base. It is such “relatedness” that provides opportunity for the gains from generating new, redeployable resources and from redeploying them, and consequently provides opportunity for earnings growth. Thus limits to integration come from a lack of specificity, and also may be tied to diseconomies of scale or scope in management and the value that may be gained from obtaining the new ideas and perspectives of outsiders. Hence a hybrid form of integration, such as a joint venture, which entails a team composed partly of outsiders, may offer the benefits of exposure to outside capabilities, but also can be expected to involve the costs of results that are less specific to and harder to redeploy within the firm.

So far, we have taken the firm’s “base,” to which resources may be redeployed, as a given, in order to develop a theory of the scope of the firm. Now turning to the question of “Why does the firm exist?” we ask, Why are the firm’s activities conducted as a firm instead of as a collection of market transactions? For this part of the analysis, consider two related activities, S and T. We are asking, under what circumstances should S and T be owned in common rather than existing as independently owned operations contracting with one another? Our purpose here is to establish that resource-based theory provides a reason for S and T to be owned in common, even if there is no opportunism.

We must be careful to define the concept of ownership. For our purposes, the only component of ownership of interest is the operational control of the entity’s resources, rather than the legal title to these resources. In particular, if the separate legal owners of S and T contract with person R for R to be Chairman and CEO of both S and T, then we may treat S and T as “owned” in common (i.e., we may view S and T as merged into a single firm). The only difference, for our purposes, between “operational” and “legal” merger of S and T is that, in the former, either of the two owners could renege on the contract, appoint his own CEO for his company, and refuse to continue performance of the contract. We are inquiring, however, whether the resource-based view contains a theory of the firm’s existence that does not depend on the presence of such potential opportunistic behavior. Therefore, we can ignore the possibility of this sort of contract-breaching opportunism, and inquire whether there are other, resource-based reasons why the legal owners of S and T might favor common ownership.

For our purposes, we may assume that firms may develop, to varying degrees, a culture and organizational routines that affect behavior throughout the firm. In the situation outlined here, we can think of the genesis of such firm-level influences, should they exist, as residing in the initial fact of a common CEO; over time, the initial culture and routines can be expected to evolve in response to learning, changed circumstances, etc. Thus firms with different CEO’s will begin with or develop different firm-level cultures and routines. Put another way, they start as or become heterogeneous entities.

Use the notation ST if S and T are owned in common and S + T if they are owned separately and are bound only by a contractual tie. Based on the above, suppose that ST has a common CEO and a firm-level culture, routines, etc., whereas S + T has two CEO’s and two distinct firm-level cultures, etc.

Consider two situations. First, as discussed above, assume that a research proj-
ect is undertaken to create a product using advances in technology to be developed during the course of the project. Assume that under ST, the project is done in-house, whereas under S + T a research team is formed, by contract between S and T, that includes members from both S and T. ST, but not S + T, has the potential to derive the full benefits, as discussed above, from the research project. S + T, as compared to ST, will find it more difficult to orient the research so as to produce redeploable knowledge, skills, and routines, because S + T must try to orient in two different directions, one for S and another for T. Similarly, the actual redeployment will be more difficult for S + T than for ST, because in ST the gains need be redeployed within the context of only one set of firm-level routines, culture, etc., instead of trying to integrate gains into two, distinct such sets. As the second situation, consider a manufacturing advance discovered by the staff in the S operation. Assume that, in principle, this advance could be very useful to the operation of T. The transmission of this information presents less of a problem in ST than in S + T, because in S + T the information must in essence be transplanted from one environment of firm-level routines, culture, etc. to a different such environment. Not only will this be more costly and time consuming in S + T than in ST, but, qualitatively, the odds are higher that the transplant will not "take."

In both situations, provided that ST has developed a firm-level culture, routines, etc. that compose positive linking mechanisms, the firm ST is seen to have real, resource-based advantages over the contractual S + T. The resource-based view thus implies a theory of the firm's existence that turns on advantages (over the market contract) in inter-component knowledge transplantation and in the creation-redeployment of specific assets. The resource-based view's implicit theory of firm scope (again relative to market contracts) also turns on advantages in creation-redeployment of specific assets, with the magnitude of these advantages depending on the strength of firm-level linkages. Lastly, the resource-based perspective's theory of why distinct firms exist turns on the notion that heterogeneous firms exist, in part, to take advantage, as asset receptors, of potentially valuable idiosyncratic assets to which the firms are made specific through their own idiosyncratic heterogeneity. This theory of the firm does not depend on the presence of opportunism problems.

Thus we arrive at a fundamental difference between resource-based and transaction cost theory. Transaction cost theory assumes that the same productive activity can be carried on either within the firm or by a collection of autonomous contractors: that is, except for problems of opportunism, the same inputs can be used equally productively in a firm or a market context. Resource-based theory, on the other hand, draws from the vision of the firm as a unique combination of inputs to question this assumption. In a resource-based view of the firm, team-specific assets within the firm will be more specific to other teams inside the firm than to teams outside the firm, and hence more productive.

This view of the reason for the firm's existence parallels that reflected in the resource-based literature (e.g., Teece, Pisano, & Shuen, 1990; Wernerfelt, 1984) and makes clear why Penrose's (1959) theory of firm growth is believed central to a resource-based perspective [for a review of its impact, see Mahoney & Pandian, 1990]. It also may be at the heart of the large stream of strategy literature con-
cerned with the effect of relatedness on diversification (see, e.g., Montgomery, 1982; Montgomery & Singh, 1984; Rumelt, 1974; Wernerfelt & Montgomery, 1986), and sheds light on why this literature is so central to strategic theory.

Thus it appears that the theory of the firm embedded in the resource-based approach may be substantially different from that in transaction cost theory, despite the fact that input specificity and small numbers figure prominently in each. Resource-based theory focuses on the firm as input-combiner (neoclassical theory) and as efficiency seeker in production and distribution (Chicago), the success of which both depend upon the environment in which it operates (Bain-type IO) and also how it shapes that environment. Thus the resource-based approach embraces the (positive) value-creating potential of the firm as at the heart of theory of the firm, rather than avoidance of the (negative) effect of opportunism. It is important to note, however, that the resource-based theory of the firm does not imply that opportunistic potential does not exist, nor does it rule out integration decisions based on reducing opportunistic potential. Rather, the resource-based view presents an alternative rationale for existence of the firm and its scope. Empirically, there is no reason why it might not be observed that some integration decisions seem to turn on specific asset creation/redeployment issues, and others on reducing opportunistic potential. A fruitful area for further theory development may be exploration of situations in which integration decisions can be expected to depend on one or the other of these considerations, or a combination.

Conclusion

The aim of this article has been to assess the distinctiveness of strategy's resource-based theory, as compared to major streams of antecedent theory related to IO economics. These include the neoclassical perfect competition model, Bain-type IO, the Schumpeterian and Chicago responses to Bain-type IO, and Coase/Williamson transaction cost economics. A central thesis of the paper is that the resource-based approach is reaching for a theory of the firm; therefore, an appropriate comparison is to the theories of the firm embedded within these research streams.

The analysis indicates that a resource-based approach both incorporates and rejects at least one central feature of each of these IO-related theories. For example, the neoclassical view of the firm as input-combiner also is at the heart of resource-based theory, as is the Chicago view of the firm as efficiency seeker in production and distribution. A resource-based view, however, rejects neoclassical theory's assumptions of perfect information, resource mobility, and divisibility. And the Chicago focus on the long term, in which entry dissipates above-normal earnings, is not shared by resource-based theory, which instead emphasizes analysis of the strategic problem of the firm in the short and intermediate runs of greater interest to the firm, in which inputs may remain costly to copy. In fact, in terms of long-lived earnings, a resource-based view finds greater similarity to Bain-type IO, which holds that above-normal returns can persist in the long run. However, a resource-based perspective is at odds with the Bain-type view that such earnings reflect exercise of monopoly power or collusion, coupled with entry deterrence and predation. A resource-based approach instead emphasizes above-normal
earnings as rents to costly-to-copy productive assets. Further, it sees close-in diversification and vertical integration as the fuller use and realization of gain from unique assets, rather than as a means to the Bain-type goals of "extension of monopoly" or capital barriers to entry. A resource-based view embraces Schumpeter's notion of a dynamic process of "creative destruction," in which firms can make stunning gains in (or experience equally stunning losses of) competitive position; however, a resource-based view also holds that above-normal earnings, in which inputs remain costly to copy, may come from much less spectacular innovation as well. As in transaction cost economics, asset specificity and small numbers are critical to a resource-based view, for these lie at the heart of what renders inputs difficult to copy and hence are sources of above-normal returns. But unlike transaction cost theory, resource-based theory as theory of the firm does not depend on the presence of opportunistic potential. Thus a resource-based approach both reflects a strong cumulative IO heritage and is at the same time unique in that it incorporates a major departure from each of the five IO theories analyzed.

Significant attention in the article also has been given to resource-based theory as theory of the firm. Two "tests" of whether resource-based theory constitutes a theory of the firm have been posed: explanation of why the firm exists (and the limits to its organization) in comparison to (a) a collection of arms-length contracts between autonomous input owners, and (b) other firms. It is argued that, in comparison to its IO predecessors, resource-based theory comprises a new theory of the firm. Special attention is given to comparison with transaction cost theory because the latter (a) explicitly addresses the firm versus market question, and (b) emphasizes asset specificity and small numbers. The analysis suggests that a resource-based theory of the firm is unique in that it need not incorporate opportunistic potential to meet the two tests outlined above.

The contributions that resource-based theory and empirics ultimately will make depend to great measure on how the approach is operationalized. Four issues in this regard are identified below.

First, resource-based theory holds that performance differentials between firms depend to significant measure on possession of unique inputs and capabilities. Thus in the best of all worlds, resource-based theory would be aimed at generalizing about uniqueness, which clearly is impossible. Thus, for example, empirical work faces the quandary noted by Hatten and Hatten (1987:333): "[b]ecause the strategist is attempting to difference his company from its competitors, the critical information is not similarities, but differences." This issue needs to be addressed in further development of resource-based theory.

A second and related issue concerns the level at which inputs (resources) are defined. Dierickx and Cool (1989a) distinguish two levels: resources such as reputation and dealer loyalty (called stocks), which are created from investment in more elementary resources, such as team capabilities in quality control (called flows). This division suggests that, within firms, "hierarchies" of resources may exist: elementary resources such as individuals' capabilities may contribute to creation of another, more aggregate level of resources such as organization culture, which may contribute to still another, even more aggregate level such as company reputation. Explicit attention needs to be given to understanding the levels of resources that
may exist within firms and to the potential contribution of each to performance differentials. Recognizing such levels appears especially important in preventing resource-based theory from becoming tautological: at some level, everything in the firm becomes a resource and hence resources lose explanatory power.

A related third point concerns acceptable empirical proxies for firm resources. In particular, to what extent should outcomes of application of firm resources be used as proxies for the underlying resources? For example, it might be tempting, considering data problems, to use “number of new products launched” as a proxy for a firm’s resources related to R&D capabilities. It is clear, however, that (a) firms measuring identically on launches may have entirely different components of R&D capability, or (b) instead of reflecting R&D capability, launches may indicate management proclivities regarding when a product is deemed ready for market or plain luck in development time. It appears that reliable measures of underlying resources pose a heavy data burden on empirics.

The fourth and final comment concerns the relationship between the resource-based perspective and the “new IO’s” game theoretic approach (for reviews see, e.g., Fisher, 1989; Shapiro, 1989; Tirole, 1988). It is apparent that a resource-based approach views a firm’s performance as resulting from the the simultaneous interaction of at least three forces: the firm’s own asset base, the asset bases of competitors, and constraints emanating from the broader industry and public policy environment. Although further development of the resource-based approach will benefit from employment of a variety of research methods, developing the theoretical implications of such complex interactions is an area in which resource-based theory may gain from application of the new IO’s game-theoretic techniques.

Notes

1 Such theory addresses why the firm exists, and what determines its scale and scope [Holmstrom & Tirole, 1989].

2 The term “reaching” is used here to indicate that explicit resource-based theory is still under development. Examples of work toward or affecting this end include Rumelt (1984, 1987, 1989), Lippman and Rumelt (1982); Wernerfelt (1984); Barney (1986, 1989); Dierickx and Cool (1989a, 1989b); Mahoney and Pandian (1990); and Teece, Pisano and Shuen (1990). Connection of resource-based theory with theory of the firm is consistent with Rumelt’s [1984] treatment.

3 Neoclassical economics contains two models related to IO: perfect competition and monopoly. Because monopoly is a prime motivator of Bain-type IO, it is discussed there.

4 Histories of the development of the perfect competition model can be found in McNulty (1968) and Stigler (1957). Stigler credits Edgeworth (1881) with the first rigorous definition of this model, with its complete statement achieved by Knight (1921).

5 As Singer (1981:2) notes, “[t]he assumptions of the pure competition model reflect the basic structure of an undeveloped agrarian economy: standardized products, numerous firms in markets, each firm with a small share and unable by its action alone to exert significant influence over price, no barriers to entry, and output carried to the point where each seller’s marginal cost equals the going market price.”

6 Historically, a number of influential IO scholars disagreed with central underpinnings of Bain-type IO as it developed in practice. Some areas of divergence touch upon issues close to a resource-based theory of the firm. These include, for example, the relationship of internal firm processes to performance, variety of competitive response to similar external circumstances, and competition as extending beyond monopolization incentives. Several such views are summarized below.

E.S. Mason, together with Chamberlin (1933), is credited with initiating the IO field (for a history, see, e.g., Grether [1970]). In what is generally acknowledged as Bain-type IO’s seminal article, Mason posits the objective of understanding how large firms set prices, (i.e., understanding pricing in a not-perfectly-competitive economy). He writes that this understanding can only be achieved through simultaneous consideration of both the
internal resources of firms and their external environment, that is, of “(a) the influence of the organization of a firm on the character of the firm's reaction to given market situations; and of (b) elements of market structure which include many more things than numbers and product differentiation” (1939: 61). He goes on to say that “firms are not, regardless of what economic theory may suppose, undifferentiated profit-maximizing agencies which react to given market situations in ways which are independent of their organization...An analysis of the relationship between organizational and market differences and the character of the price response is the central problem of price analysis. The relation of size to price policy is merely one part of the problem which, taken out of its setting, is not very amenable to fruitful discussion” (1939: 62, 75). Although Mason reveals in this article his particular interest in the role that market structure (broadly defined) plays in price-setting, he makes it clear that such understanding will only clarify a part of the price-setting process. This view is reflected in his later criticism of Bain-type IO, which he regards as seeking overly simple explanations of complex phenomena (see, e.g., Mason, 1949, 1957c, 1966). For example: “given 50 per cent of sales in the hands of four firms, can it be assumed for the purpose of predicting price and quantity response to some outside disturbance that the firms in question will act in a way sufficiently similar...to the action of one firm...to justify the introduction of such a simplifying assumption? Or...can we usefully assume, given a concentration ratio of less than 50 per cent, that the market in question will behave sufficiently like the Marshallian competitive market...to justify treating it as such? The answer to both these questions must, I think, be no. Nor can we conclude from an inspection of concentration ratios that antitrust action is or is not justified.” (1957c: 35).

J.M. Clark argues for (a) a fuller conception of competition, which would dispense with perfect competition as the standard for “competitiveness” and that would include dynamics, and (b) a recognition of the critical role diversity (uniqueness) plays in firm performance and market structure. Regarding (a), Clark (1940) introduces the notion of “workable competition” as a norm for policy decisions that recognizes market frictions driven, at least in part, by persistent differences between firms. Clark holds that normative policy judgments should rest on the more solid ground of a dynamic conception of competition that includes “initiatory action by a firm, responses by those with whom it deals, and responses to these responses by rival firms” (1955: 457). Clark encourages the view that earnings need not be a manifestation of monopoly: “the dynamic system is not one of elimination of profits, but one of erosion and re-creation, both of which are jointly essential. For the economy as a whole, this process implies the creation, reduction, and re-creation of differential rewards in different industries...” (1955: 454). And, with regard to IO's avoidance of many forms of firm-level uniqueness: “healthy degrees of diversity are an indispensable condition of the persistence of effective competitive forces in modern industry and trade. Theory, qua theory, contains a built-in predisposition against diversity, or qualitative innovation. Here diversity spells adversity, complicating theories and spoiling their precise determinateness” (1958: 475).

Heflebower (1954) outlines a theory of industrial markets and prices in which stability can exist without collusion. Central to his thesis is the concept of “market position,” reflecting the firm's internal resource structure: “each participant acquires not merely a place in the market but also a position which is a major long-term attribute. By market position is meant more than what the firm sells and to whom and by what means, and more than its share of the volume. Instead, market position is that composite of attributes which governs the ability of the firm to compete. Obviously, the relative importance of elements in this composite differs among product areas and even among firms participating in the same product area...market position becomes a means of long-term profit maximization under conditions of uncertainty. Consequently, the concept of market position is at the heart of the problem of the stability of industrial markets under given factor prices, industry demand, and basic technology because it governs the feasibility and effectiveness of various forms of rivalry” (1954: 125, 126). Heflebower urges that market structure be regarded and analyzed as endogenous. For example: “the structure and operation of markets have a strong tendency to reflect the market characteristic of the product, the feasible conditions of its production, and the organization of adjacent industries.”

Other early studies casting doubt on the importance of monopolistic/collusive practices to understanding firm behavior include Nutter (1954), Robertson (1954), Harberger (1954), and Galbraith's (1954) theory of countervailing power of buyers and suppliers to check monopoly; all of these (including Heflebower [1954]) were presented at the 1953 meeting of the American Economic Association. In addition, Cyert and March relate the firm's internal organization structure to its pricing decisions, holding that “the firm's perception of the market and the firm's perception of its capabilities for action are both affected by its own organizational structure” (1955: 130), a view subsequently developed more extensively (Cyert & March, 1963). And Lanzillotti, based on interview data with officials from 20 companies, paints a remarkably vivid picture of how firms set prices that has little to do with traditional concerns of Bain-type IO. For example, “the findings show that one company will prefer stability, another will seek to expand its market share, or to engage in continuous discovery and pre-emption of new fields, while others will be content to meet competition, to satisfy a set target, or to aim at combinations and variations of these goals” (1958: 939).

A final example, perhaps better known among strategy scholars, is Penrose's theory of firm growth (1955, 1959), which views size and scope as a function of a firm's internal process of development and which concludes in response to monopoly-power views that “size is but a by-product of the process of growth, that there is no 'optimum,' or even most profitable size of firm” (1959: 1, 2). Penrose bases her theory on analysis of the pro-
ductive capabilities of internal firm resources, arguing that "expansion of firms is largely based on opportunities to use their existing productive resources more efficiently than they are being used," further noting that firms attain a unique character by virtue of the heterogenous resources of which they are composed (1959: 88, 75).

See, for example, reviews by Demsetz (1973b), Weiss (1975) and Scherer (1980).

Gale's paper is particularly interesting as it explicitly introduces firm resource heterogeneity into the Bain-type IO framework. Gale sought to determine the relationship between market share and firm profitability, assuming that "market shares in both homogenous and differentiated oligopolies are determined by past and present rent-yielding intangible assets (unit costs and market advantages) possessed by the firm" (1972: 413). Gale brings this concept into the Bain-type IO framework as follows: "In an oligopoly environment the large-share firms tend to recognize their mutual interdependence and attempt to coordinate their market conduct so as to raise the industry profitability above what it would be in the absence of such coordination. While the firms have a common interest in increasing the industry rate of return, their interests conflict on the question of how this increased profitability should be allocated to the rival oligopolists. A firm with a large market share is in a position to bargain for a pattern of industry conduct which will reflect its own interests perhaps at the expense of smaller share firms" (1972: 413).

Of course, at its heart Bain-type IO holds that competition is effectively prevented by the monopolistic/collusive practices enabled by large size. Once established, persistence of firm heterogeneity is logically consistent with elimination of competition.

See, for example, Goldschmid, Mann, and Weston (1975).

For example, "[t]he Federal Trade Commission is conducting an unprecedented investigation into alleged conspiracies among Japanese car makers and parts suppliers operating in the U.S. to squeeze out U.S. suppliers. The Justice Department, after months of internal study, is on the verge of giving government lawyers expanded authority to prosecute foreign companies in U.S. courts for actions committed abroad. "We will not stop at the shoreline in our pursuit of harmful monopolistic practices," Attorney General Dick Thornburgh declared in a recent speech" (Wall Street Journal, October 31, 1990: A16).

"Before 1850...[i]ndustrial enterprises were very small, in comparison to those of today. And they were usually family affairs. The two or three men responsible for the destiny of a single enterprise handled all its basic activities—economic and administrative, operational and entrepreneurial" (Chandler, 1962: 19).

"Rarely in the history of the world has an institution [large-scale business enterprise] grown to be so important and so pervasive in so short a period of time." For example, "in the 1840's the traditional mercantile firm, operating much as it had for half a millennium, still marketed and distributed the nation's goods. Within a generation (italics added) it was replaced in the sale of agricultural commodities and consumer goods by modern forms of [mass] marketing enterprises" (Chandler, 1977: 4, 289).

"Chandler notes that manufacturing enterprises met the opportunities for growth in two ways: by expanding and integrating through creating their own marketing function, or by banding together in horizontal combination (1962: 25). Both strategies led to previously unknown size relative to the market. JB. Clark, summarizing popular sentiment at the time, states: "American industry has recently gone through a rapid and startling evolution. Consolidations of capital, which look like monopolies, have come, apparently to stay. Just as we were beginning to understand an economic system in which competition ruled, the system transformed itself into one which is seemingly based on the repression of the competitive process" (1901: 1).

"Lilienthal, in a defense of big business, nonetheless says this about the Sherman and Clayton Acts: "[t]hose laws were enacted many years ago as the creative Magna Carta of economic freedom for an America emerging from an agricultural economy. There are few greater legislative achievements than these laws that struck at the monopoly practices of the trusts of the nineties and the early years of this century, with their price agreements, pools, rebates and the ruthless slaughter of industrial newcomers" (1953: 33-4).

"Some economists had what amounted to a crisis of faith. For example, according to Milton Friedman, "in 1934...I would say that close to a majority of the social scientists and the students in the social sciences at the University of Chicago were either members of the Communist party or very close to it. That was the environment in which Frank Knight gave a series of lectures under the title, 'Why I Am a Communist, by an Ex-Liberal.' It was an environment in which the general intellectual atmosphere was strongly prosocialist. It was strongly in favor of government going all the way to take over the whole economy" (Kitch, 1983: 178-9).

"The Bain-type IO view has deep popular roots. For example, Blake writes that "opposition to the growth of economic concentration, with its real or supposed evils, is one of a few basic tenets about which there has been a wide and continuing consensus in the Congress at least since the passage of the Sherman Act nearly eighty-five years ago...Congress [has a] deeply held belief that most forms of economic power should be dispersed" (1974: 341). Mason earlier comments that ours is "a society in which size is popularly considered a menace" (1939: 63). Interestingly, Dewey holds that antitrust is less of a reaction to inefficient resource allocation than to a fear of concentration of discretionary authority (1975: 13). An early example supporting Dewey's view comes from the Governor of Massachusetts' 1840 message to the Massachusetts state legislature, in which he said of the "large" corporations then existing:

These societies are one of the vices of our time. They encourage speculation and fraud, mobilize landed property, overturn matrimonial arrangements, escape publicity in the transfer of real property, diminish
the sense of individual responsibility, create property in mortmain, prevent all penal remedies, are lacking in moral sense, and constitute finally a grave social peril by concentrating too much power in the hands of certain of our citizens. [quoted in Stocking, 1959].

For example, in 1968 the White House Task Force on Antitrust Policy proposed the "Concentrated Industries Act." Under this bill, the U.S. Attorney General would identify "oligopoly industries," and if the firms did not voluntarily take steps to reduce their industry's concentration, a decree would be entered under which a reduction of concentration [would be achieved] such that the market share of each oligopoly firm in such oligopoly industry does not exceed 12% (italics added). Such decree may include provisions requiring a party (i) to modify its contractual relationships and/or methods of distribution; (ii) to grant licenses (which may, in the discretion of the court, provide for payment of royalties) under and/or dispose of any patents, technical information, copyrights and/or trademarks; and (iii) to divest itself of assets, whether or not such assets are used in an oligopoly industry, including tangible assets, cash, stock or securities (including securities in existing firms or firms to be informed), accounts receivable and such other obligations as are appropriate for the conduct of business. (Goldschmidt, Mann, & Weston, 1975: 449-56)

In 1972 Senator PA. Hart introduced the "Industrial Reorganization Act," which proposed to make possession of monopoly power unlawful. Particularly noteworthy was the following provision:

There shall be a rebuttable presumption (italics added) that monopoly power is possessed
(1) by any corporation if the average rate of return on net worth after taxes is in excess of 15 percent (italics added) over a period of five consecutive years out of the most recent seven years preceding the filing of a complaint, or
(2) if there has been no substantial price competition among two or more corporations...for a period of three consecutive years out of the the most recent five years...
(3) if any four or fewer corporations account for 50 percent (or more) of sales...in any year out of the most recent three years preceding the filing of the complaint. (Goldschmidt, Mann, & Weston, 1975: 445-48)

These legislative proposals, ultimately not enacted, took the Bain-type IO view of the firm to its logical conclusion—disallowing market share or above-normal returns over a certain threshold.

Mason comments that Schumpeter "regarded antitrust policy in practice as a relatively senseless harassment of large firms by people with little understanding of the historical and potential contribution of large-scale enterprise to effective competition" (1953/1957b: 372).

Mason comments that Schumpeter's view "represent[s] one of the most effective as well as most drastic critiques extant concerning traditional pattern of antitrust thought. The critique is drastic and effective because it plausibly undermines the two main pillars of the traditional ideology: first, that market power is the proper object of attack since power means the ability to exploit; and second, that the preservation of competition, meaning the exclusion of positions of market power, will assure the efficient use of resources" (1951/1957a: 91). Mason's view is not contrary to that of other authors commenting on the impact of Schumpeter's vision, including Scherer (1980) and Cohen and Levin (1989).

A history of the evolution of the Chicago tradition can be found in Kitch (1983), which contains a transcript of an extraordinary meeting of scholars at the center of development of this tradition.

Stigler (1968c: 1), for example, treats Chicago-type IO as an applied branch of price theory. On the macroeconomic side, the Chicago tradition amounted to a similar counter-reformation, reviving and strengthening the older monetary theory in opposition to Keynes and his disciples (see, e.g., Friedman, 1956).

This position may have been taken because it provides an alternative efficiency explanation for size that relies on standard price theory alone.

The Chicago tradition also is noteworthy in its questioning of the power of government intervention to increase welfare. Although Bain-type IO adheres to the view that government, if summoned, can "fix" the "problem," the Chicago view is that many practices previously alluded to as monopolistic or collusive are, in fact, outcomes of firms trying to maximize productive efficiency and that even if elements of monopoly are involved, the ability of governmental intervention to diminish these (and not make them worse) must be questioned. For example, "[d]ivestiture because of size or concentration runs the risk of incurring social costs that exceed the possible gains from reducing any control over price that might be enjoyed by large firms" (Demsetz, 1973b).

Another cost of using the market mechanism cited by Coase is discovering what the relevant input prices are (1937/1952: 336).

Coase went beyond the 1937 article in Coase (1960), an influential article establishing the importance of transactions costs, interacting with the regime of property rights, in determining the pattern of activity (and non-activity) in many areas other than the organization of the firm.

Alchian offers a particularly penetrating discussion of asset specificity (what he calls "interspecific resources"). "An interspecific input is one whose value depends on—i.e., 'is specific to'—the behavior of some other particular resource or some activity. The investment cost of an interspecific resource that is non-salvageable if the other resource, to which it is specifically dependent, disappears is called the interspecific quasi-rent.

JOURNAL OF MANAGEMENT, VOL. 17, NO. 1, 1991
For example, if I invest in a machine, A, whose value depends on services rendered by some particular other resource, say B, then the loss of A's investment value if B's services are withdrawn is a non-salvageable quasi-rent value of A and it is specific to the presence of B." (1982: 14-15)

28 Here, referring to Williamson's (1975) conditions of bounded rationality and information impactedness.

29"Firms" here mean assets owned in common, as opposed to transactions between autonomous contractors.

A  Chalian provides the following definition: "a particular set of contracts involving owners of private property rights in intespecific inputs for teamwork is typically a firm. Physically the firm is the set of resources subjected to that set of contracts." A Chalian goes on to conclude, not inconsistent with Williamson's argument, that there are no "purely rental" firms: the "firm would not exist, if all resources were non-specialized and separately owned." (1982: 16, 24)

30 The idea here is that if A threatens to engage in opportunism, costing B an amount X, A's threat is not credible if A also stands to lose X if the relationship is terminated.

31Williamson notes that he "regards the business firm as a governance structure [for transactions] rather than as a production function" (1989: 136). Klein and Leffler (1981), by incorporating diseconomies to management, implicitly go beyond the notion of the firm as decider of which assets to own and which to "rent." Klein and Leffler add to the transaction cost perspective the notion (from the Chicago tradition) that firms, in addition to being avoiders of transaction costs, are producers of end-products, and seek to economize on production costs.

32 These two avenues to above-normal returns reflect Porter's (1980) generic strategies. Given the real-world pervasiveness of product differentiation, condition (a) may be the most empirically relevant.

33 "Costly to copy" means that rivals would have to pay more to obtain, duplicate, or substitute for, an input than did the originating firm. "Unique" assets are those that for analytic purposes can be regarded as infinitely costly to copy.

34 For example, Dierickx and Cool remark that "[w]e agree with Barney that it may be logically impossible to formulate a set of rules to systematically create competitive advantage" (1989b: 1514).

35 Asset specificity is discussed above in note 27. For example, asset specificity as linkedness emerges in Porter's (1985) discussion of the value chain, particularly in the discussion of the relationship between elements of the value chain.

36 Firms other than the highest-valuing firm do not succeed in acquiring A; they get zero rent from A, therefore. For convenience of presentation, if A is specific to both X and Y but υ(X) = υ(Y), or if A is specific to neither, assume that X gets A.

37 That is, unless A is not relatively scarce, from X's point of view. This is the situation depicted in Table 4.

38 If only one unit of A is available, there is a bilateral bargaining situation in which it makes no sense to speak of a marginal purchaser.

39 This discussion concerns potential rents, abstracting for now from the effect of opportunistic potential.

40 In a related vein, Wernerfelt discusses competition over creating a resource position barrier: "If everyone goes for the potentially attractive resources and only a few can 'win' in each, firms will lose unless they pick their fights well. So firms need to find those resources which can sustain a resource position barrier, but in which no one currently has one, and where they have a good chance of being among the few who succeed in building one. They have to look at resources which combine well with what they already have and in which they are likely to face only a few competitive acquirers" (italics added) (1984: 174-5).

41 The model in Tables 2-4 implicitly assumes perfect information on the part of the potential buyers, X and Y.

42 It can be argued that in many situations X and Y will not know the value of A ex ante, and in fact may never know it, given uncertain imitationability. (Lippman & Rumelt, 1982). In this case, the firms' valuation of the input becomes an expected value, and the risk profile of the firm affects its reservation price for A, and thus the expected rent. But, perhaps more importantly, the owner of A may also have imperfect information, one source of which may be who the potential buyers are (analogous to Coase's [1937/1952: 336] "discovering what the relevant prices are" as one of the costs of using market transactions). Firms may also earn rent on the seller's lack of knowledge. For example, in the tables, the uninformed seller could think that regime 3 applies instead of regime 4.

43 This may occur because, e.g., X is in a better position to use A, or through luck (Barney, 1986).

44 Assuming that no cheap substitute for or superior technology not needing the ore is found. Rent from early-obtained ownership of the mine could be dissipated if, for example, federal legislation was proposed that outlawed excavating in the area where the mine is located, and the firm invests in battling this proposal (i.e., attempts to "enforce" its contract ownership rights through political action).

45 "In a related vein, Demsetz, in the Chicago tradition, argues that "the problem of defining ownership is precisely that of creating properly scaled legal barriers to entry" (1982: 52).

46 Also in the Chicago tradition, Demsetz writes that "information costs are the more fundamental barrier to entry. These costs, not necessarily in identical amounts, constitute hurdles to all who would (and have) enter(ed) the industry. Complete knowledge about products and firms would make brand loyalty useless from both consumer and seller viewpoints...In the presence of such costs, consumers will find it useful to rely on a firm's experience and reputation, or more correctly, on its history or on the fact that it has made sizable investments specific to this industry...If new firms are to sell equal quantities at equal prices, they may need to incur higher costs of persuading and communicating than presently is required of older firms (1982: 50)."
In the neoclassical tradition, the firm was treated as little more than a nexus of contracts among input suppliers, and the issue of existence in comparison to the market did not arise. Bain-type IO also was largely silent on this issue.

A firm's original resource endowment can be seen as capabilities of the entrepreneur(s) who founded the firm. Thus because individuals are unique, even tiny firms also are unique.

This view is only superficially consistent with Bain-type IO, in which the key "assets" of interest are market power—to be able to raise price above the competitive level—and active or passive barriers to entry—so as to be able to keep prices high.

Doing the work in-house also may help to develop "absorptive capacity" to make use of information spillovers from other firms, as discussed by Cohen and Levinthal (1989).

It is not sufficient to object that the same result can be achieved by a long-term contract with, say, outside scientists who possess this know-how and who work full-time for and within the firm, because such a long-term contract is tantamount to employment within the firm, for our purposes.

This point appears central to Prahalad and Hamel's (1990) view that the critical activities of the firm are creation of a set of core competencies—which involve the coordination of diverse production skills and integration of multiple streams of technology—and core products that embody these competencies (often intermediate products). Prahalad and Hamel argue that these functions should not be contracted out even if, for example, a supplier can produce a core product more cheaply than can the firm, because production and improvement of core products extends and enhances the underlying core competence and suggests new avenues for its redeployment.

This concept of ownership is not inconsistent with that of Williamson (1975), and is broad enough to apply to the situation when "resources" are people (who cannot, of course, be "owned" by a firm).

For a discussion of factors of such firm-wide linkages, see Phillips (1990).

References


JOURNAL OF MANAGEMENT, VOL. 17, NO. 1, 1991


Mahoney, J., & Pandian, J.R. 1990. The resource-based view within the conversation of strategic management. Unpublished manuscript, Department of Business Administration, University of Illinois.


JOURNAL OF MANAGEMENT, VOL. 17, NO. 1, 1991
A HISTORICAL COMPARISON


JOURNAL OF MANAGEMENT, VOL. 17, NO. 1, 1991