A Resource-Based Theory of the Firm: Knowledge versus Opportunism

Kathleen R. Conner; C. K. Prahalad


Stable URL:
http://links.jstor.org/sici?sici=1047-7039%28199609%2F10%297%3A5%3C477%3AARTOTF%3E2.0.CO%3B2-G

Your use of the JSTOR archive indicates your acceptance of JSTOR’s Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR’s Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

Organization Science is published by INFORMS. Please contact the publisher for further permissions regarding the use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/informs.html.

Organization Science
©1996 INFORMS

JSTOR and the JSTOR logo are trademarks of JSTOR, and are Registered in the U.S. Patent and Trademark Office. For more information on JSTOR contact jstor-info@umich.edu.

©2003 JSTOR
A Resource-based Theory of the Firm: Knowledge Versus Opportunism

Kathleen R. Conner • C. K. Prahalad
University of Michigan Business School, University of Michigan, Ann Arbor, MI 48109

Abstract
This paper develops a resource-based—knowledge-based—
theory of the firm. Its thesis is that the organizational mode
through which individuals cooperate affects the knowledge
they apply to business activity. We focus on the polar cases
of organization within a firm as compared to market contract-
ing. There will be a difference in the knowledge that is
brought to bear, and hence in joint productivity, under the
two options. Thus, as compared to opportunism-based, trans-
action-cost theory, we advance a separate (yet complementary)
answer to the question: why do firms exist? Our aim is
to develop an empirically relevant and complementary theory
of why firms are formed: a theory based on irreducible
knowledge differences between individuals rather than the
threat of purposeful cheating or withholding of information.
We assume limited cognitive abilities on the part of individu-
als (bounded rationality), and assume that opportunistic be-
behavior will not occur. The latter allows us to determine
whether resource-based theory has independent force, as
compared to the opportunism-based, transaction-cost ap-
proach. The paper predicts choice of organizational mode,
identifying whether firm organization or market contracting
will result in the more valuable knowledge being applied to
business activity. The resource-based predictions of organiza-
tional mode are compared and contrasted with correspond-
ing opportunism-based, transaction-cost ones. A principal
point is that knowledge-based considerations can outweigh
opportunism-related ones. The paper also establishes the
relation of a theory of the firm to a theory of performance
differences between competing firms.
(Theory of the Firm; Strategy; Resource-based; Transac-
ton Costs)

This paper develops a resource-based theory of the
firm. Its thesis is that the organizational mode through
which individuals cooperate affects the knowledge they
apply to business activity. For example, $Y$ and $Z$ might
face a choice between working together as employees
in a firm or completing the same task as independent
contractors. There will be a difference in the knowledge
that is brought to bear, and hence in their joint
productivity, under the two options. This conclusion
depends on the straightforward assumption that $Y$ and
$Z$ each possess experience, insights, or skills that are to
some extent different from those of the other. Tacit
knowledge, which can be learned only through personal
experience (Polanyi 1962, Nelson and Winter
1982, Nonaka 1994), is an example of know-how that is
difficult to transfer ex ante. The difference in the
knowledge that is brought to bear, once anticipated,
affects the choice of organizational mode itself. Thus,
as compared to opportunism-based, transaction-cost
theory, we advance a separate (yet complementary)
answer to the question: why do firms exist?

As the literature makes increasingly clear, a knowl-
edge-based view is the essence of the resource-based
perspective. The central theme emerging in the strateg-
ic management resource-based literature is that pri-
ately held knowledge is a basic source of advantage in
competition. The resource-based view generally ad-
dresses performance differences between firms using
asymmetries in knowledge (and in associated compet-
tencies or capabilities; see, e.g., Amit and Schoemaker
1993, Barney 1991, Chen 1996, Henderson and
Cockburn 1994, Peteraf 1993, Prahalad and Hamel
1990, Robins and Wiersema 1995, Schoemaker and
the firm thus entails a knowledge-based perspective.¹

Using the standard established by Coase (1937), a
theory of the firm generally addresses two issues: why
firms exist, and what determines their scale and scope
(Holmstrom and Tirole 1989). Of the two, the reason
for the firm's existence is logically prior and is the
primary subject of this paper. In the literature, it is
tantamount to asking why a firm exists instead of the
alternative of market contracting between the same
individuals.

Since received theory concentrates centrally (and
virtually universally) on the effect of opportunistic be-
behavior (see, e.g., Alchian and Demsetz 1972; William-
son 1975, 1985a; Klein et al., 1978; Grossman and Hart
1986; Kreps 1990; Milgrom and Roberts 1990), we stay
away from opportunism in order to determine whether
a resource-based theory of the firm has independent force. We do not claim, however, that resource-based theory constitutes the exclusive explanation of firms. Rather, the approach presented here accepts the validity—and, indeed, the remarkable intellectual achievement—of the opportunist-based view in explaining some of the motivation for firm organization. Our aim is to identify additional, equally valid, empirically relevant causes. In particular, we argue that firms can exist because of knowledge-based transaction costs that are independent of the opportunistic considerations explored by Williamson. Recognition of knowledge-based transaction costs leads, we believe, to a fuller realization of Coase’s (1937) original insights as to the reasons for firm organization.

We focus on basic choices: firm organization versus market contracting. We thus concentrate on polar organizational modes rather than complex or “hybrid” arrangements, such as joint ventures, alliances, or complicated corporate structures. Although the latter may be of great practical and scholarly interest, analysis of them is facilitated by studying the polar modes, since the more complicated structures can be seen as permutations of the polar ones. A joint venture, for example, may consist of blending aspects of firm organization and market contracting, while a geographically extended corporate structure, in contrast, may attempt to gain the basic advantages of firm organization repeatedly, through linking multiple layers or units. The polar cases are “basic particles” from which more elaborate arrangements are constructed. In addition, in this study we are concerned with human, rather than physical, capital. We also are concerned with joint production between cooperating individuals, rather than simple exchange of discrete, separable products. The analysis is presented, for simplicity, in terms of the choices of natural persons. However, it also applies to organizational entities, such as firms or organizational subunits, considering operating as one unit versus transacting autonomously.

Section 1 motivates the study by asking: why should strategic management and organizational scholars be interested in primary development of a theory of the firm? We contend that a theory of performance differences between firms requires, as an essential component, a theory of when advantage can be gained by organizing activity within a company instead of through arms-length markets. A logical nexus exists, therefore, between a theory of performance differences between firms and a theory of the firm itself, in that the former necessarily incorporates the latter.

Section 2 identifies key concepts used in this study. These concern the nature of firm organization and individual behavior. We define the firm in a way consistent with Coase (1937), Simon (1951), Williamson (1975, 1985a), and Milgrom and Roberts (1988, 1990), among others. Firms are distinguished from markets based on an authority (employer-employee) relationship in the former, as compared to autonomous parties contracting in the latter. As to individual behavior, we assume bounded rationality (Simon 1957). This has an important corollary: cognitive limitations prohibit individuals from possessing identical stocks of knowledge. Individuals additionally are taken to behave truthfully. This insures that we avoid opportunism-related reasons for the firm, allowing us to focus on other, knowledge-based considerations. In contrast to Williamson (1985a), we maintain that truthful behavior does not rule out market contracting frictions that can give rise to firm organization.

Section 3 presents a resource-based theory of the existence of the firm. Organizational mode affects the knowledge applied to business activity in two ways. These are (1) how the parties’ existing knowledge is blended and used, and (2) how new learning or developments occurring during the course of the work are taken into account. We call these, respectively, the “knowledge-substitution” and “flexibility” effects of organizational mode. As to knowledge-substitution, an employee sometimes acts according to the manager’s direction rather than in conformity with what the employee otherwise would do. The manager’s wisdom in such situations is substituted for the employee’s. If the employee cannot absorb the manager’s wisdom before the employee profitably can apply it, the employee may opt to be directed by the manager and hence to be employed in a firm. In contrast, under market contracting, each person uses its own judgment to decide the specific responsibilities and duties that it will agree to carry out. Indeed, the very notion of an autonomous market contractor means that the individual “retains the right or power of self-government” (Merriam-Webster’s, 1994), or put differently, “works for itself.” An individual will favor knowledge-substitution—and hence, all else equal, a firm—when the manager’s understanding (present or future) is believed to be of superior value compared to corresponding elements of the employee’s.

The flexibility effect concerns the cost of altering an individual’s responsibilities or duties on an ongoing basis, in order to respond to new learning or other developments arising during the course of the work.
The cost of achieving flexibility under the two organizational modes differs, due to a fundamental distinction between firms and markets. As analyzed by Simon (1951), an employment contract need not be renegotiated when an employee undertakes a new or modified duty. In contrast, market contracts require renegotiation if responsibilities and duties are to be changed from the set originally agreed to. It is possible for either organizational mode to provide the cheaper flexibility, depending on the cost of operating the firm versus renegotiating the market contract. These costs need not depend on opportunistic potential.

Section 4 predicts choice of organizational mode from the resource-based perspective. It identifies whether firm organization or market contracting will result in the more valuable knowledge being applied to business activity, based on the combination of knowledge-substitution and flexibility effects and independent of opportunism.

Section 5 compares resource-based predictions of organizational mode with corresponding opportunism-related ones. The forecasts are identical in some instances, and opposite in others. A principal point is that knowledge-based considerations can outweigh opportunism-related ones. For example, we predict firm organization in some situations in which opportunistic potential is too low for opportunism-based theory to do so. In others, we anticipate market contracting even though a firm would be chosen if opportunism-related factors were the only ones taken into account.

Resource-based limits to firm size and scope are addressed in Section 6. Since full treatment of this question is not possible in this paper, the aim is to outline an approach through which it might be addressed later. Concluding remarks follow in Section 7.

1. Why Should We Be Interested in a Theory of the Firm? 8

In comparison to economists, less attention has been devoted by strategic management and organizational scholars to explicit development of a theory of the firm. 8 Some may have regarded theory of the firm as an essentially completed topic, given the works of Coase (1937), Williamson (1975, 1985a), and others. However, numerous fundamental issues remain unresolved and debated (see, e.g., Cheung 1983; Demsetz 1988; Dow 1987; Englander 1988; Kay 1992; Masten 1988; Milgrom and Roberts 1988, 1990; Nelson and Winter 1982; Postrel and Rumelt 1992; Williamson 1991a, 1991b, 1992, 1994).

Alternatively, primary development of a theory of the firm may not have been seen as directly germane to strategic management or organizational research. In particular, examining why a firm exists as compared to a collection of market contracts, may not have been viewed as immediately related to the question of performance differences between firms. 9 Our aim in this section is to identify a theory of the firm as a necessary building-block for addressing this broader issue.

Consider as an example two competing firms, Red and Blue. Red employs two persons, Y and Z, and Blue, two others, S and T. As illustrated in Figure 1, although there are only two firms, each nonetheless competes against two other players. Red contends not only against the other firm, Blue, but also against the disaggregation of Red’s assets, Y and Z, into a market contracting arrangement. A parallel situation pertains to Blue. Direct competition between Red and Blue, portrayed within the dotted-line box in Figure 1, implicitly entails each already having had to “beat the market,” in the sense that the employees of each firm deem firm organization to be superior to the alternative of market contracts. Indeed, disaggregation of the firms into autonomous relationships is the “stealth competitor” in any observable market contest. 10

A single firm’s productivity can be divided, for analytic purposes, into two components, as in Figure 2. Using Red as the example, these are: (1) the productivity that would be realized if the efforts of Y and Z were coordinated by means of a market contract; plus (2) the extra margin of productivity created (or lost) by applying firm organization to what Y and Z otherwise would achieve. Component 1 in Figure 2 can be thought of as Red’s “base” level of productivity, i.e., the level that would be produced if the constituent assets, Y and Z, were completely self-directing. Component 2, on the other hand, is a change in the base level, which can be positive or negative. It is generated by altering the relationship between Y and Z, so that autonomous dealings are replaced by those occurring in a firm. 11
Competition between Red and Blue can be analyzed as a contest over their relative standing as to the Figure 2 components of individual firm productivity. As illustrated in Figure 3, Red and Blue contend over the base level of productivity that each achieves, as well as over the strength of benefits each provides because its employees are managed within a firm. Competition between firms can be divided into the two components shown in Figure 3.12

The connection between a theory of the firm and a theory of performance differences between firms also is illustrated by Figure 3. Specifically, one source of differential performance between firms is the degree to which each implements the productivity benefits arising from firm organization itself—Component 2 in Figure 3. Isolating the elements creating a difference in the productivity of a firm as compared to market contracting (i.e., a theory of the firm), is identical to identifying a subset of the factors determining why some firms outperform others. For example, if selecting firm organization over market contracting involves certain knowledge-related factors, then the comparative performance of firms will be affected by the relative strength of their implementation of these same factors. Our conclusion: a theory of performance differences between firms necessarily implies and incorporates a theory of the firm itself.13

We therefore have ample reason to be concerned with primary development of a theory of the firm. It is an essential building-block for the more general theory to which resource-based analysis aspires. By developing a resource-based theory of the firm, we simultaneously take an important step toward the more complex task of creating a resource-based theory of performance differences between firms. In addition, we gain insight as to opportunism-independent, knowledge-based reasons for vertical integration itself.14

2. Definitions and Assumptions

A theory of the firm depends on the definitions and assumptions that form its basis. This section concerns two such aspects: firm organization, and key parameters of individual behavior.

2A. Firm Organization

Establishing how the organizational modes differ is essential, since any extra rents to firm organization are returns to the difference between—characteristics that distinguish—firms and markets.15 As reviewed in the appendix, defining firm organization has been problematic within the theory-of-the-firm literature (see, e.g., Masten 1988).

Our intent is to employ a firm-market distinction parallel to that generally used in opportunism-based, transaction-cost theory. Firm organization is distinguished from market contracting based on existence of authority in the former (i.e., employer over employee), as compared to the parties acting autonomously in the latter. Authority, or the possession of control-rights, allows one person to direct—manage—the actions of another. Therefore, the firm represents a different
The specific conception of firm organization and market contracting used in this study is diagrammed in Figure 4. It has been chosen for its simplicity and parsimony, in order to bare essential features of the resource-based argument. The "raw materials" (the individuals, Y and Z) are identical under the two organizational modes, as is traditional in theory-of-the-firm analyses.

Under market contracting, Y and Z each can be thought of as directors of their own, one-person firms (Demsetz 1988). They establish a contract, herein called a market contract, which sets out obligations of, and compensation due to, both. Once market contract terms are established, Y and Z are obliged to fulfill them. Under firm organization, on the other hand, Z is the manager, and Y, the employee. It may be typical to think of Z also as the legal owner of the firm, in the sense that Z is the sole shareholder if the firm is a corporation, or the owner if the firm is a sole proprietorship. Thus it may be typical to think of Z as the "residual claimant" of the firm. For our purposes, however, it is the power to give direction, with which Y is required to comply, rather than legal ownership per se, that is important. Thus it is not necessary that Y be paid a wage by the firm, with Z, as residual claimant, being entitled to everything that remains after paying costs. As Demsetz (1988) puts it,

[t]he direction of some by others catches the spirit of managed coordination [within a multiperson firm] (p. 156).

Under firm organization, Y enters into an employment contract with Z, as compared to a market contract between them. In a seminal article, Simon (1951) addresses the nature of "[t]he authority relationship that exists between an employer and an employee, a relationship created by the employment contract" (p. 293).

We will call our employer B (for "boss"), and our employee W (for "worker"). The collection of specific activities that W performs on the job (typing and filing certain letters, laying bricks, or what not) we will call his behavior. We will consider the set of all possible behavior patterns of W and we will let x designate an element of this set. We will say that B exercises authority over W if W permits B to select x. That is, W accepts authority when his behavior is determined by B's decision. In general, W will accept authority only if x0, the x chosen by B, is restricted to some given subset (W's "area of acceptance") of all the possible values (Simon 1951, p. 294).

A central point is that, under an employment contract, the employer receives from the employee the privilege of postposing, until some time after the contract is made, the selection of [the employee's behavior,] x (Simon 1951, p. 295, our italics).

Thus the authority an employer exercises as to an employee, i.e., Z's right to "manage" Y, stems from Y giving Z the right to postpone final specification of the acts required of Y until after an employment contract is established between them.17 In Simon's model, a market contract, in contrast, does not involve Y giving Z the right to specify Y's particular duties postcontractually or continuously.18 Instead, in a market contract, Y and Z endeavor to lay out in the contract itself, the particular acts required of each.

The [employment] contract differs fundamentally from a . . . [market] contract. . . . In the . . . [market] contract each party promises a specific consideration in return for the consideration promised by the other. The buyer (like B) promises to pay a stated sum of money; but the seller (unlike W [in the context of an employment contract]) promises in return a specified quantity of a . . . specified commodity (Simon 1951, p. 294).

Of course, due to bounded rationality, the laying-out of these requirements is never complete. But the essence of Simon's point is that some, and perhaps an extended, itemization of specific acts required of Y and Z is contained within a market contract, whereas in an employment contract, such itemization is replaced by Z being given the right to select the particulars of Y's duties postcontractually (within prescribed limits).

Williamson's (1975) view of firm organization does not appear to be contrary to that of Coase and Simon.

Among the most widespread characteristics of organizations is the prevalence of authoritative allocation. Virtually univer-
sally, in organizations of any size, decisions are made by some individuals and carried out by others. The fields in which an authority is valid may be limited; and the recipient of orders at one level may have his own field for authority. But within these limits, the giving and taking of orders, having someone tell someone else what to do, is an essential part of the mechanism by which organizations function (p. 63).19

Williamson (1994) later makes the point even more strongly.

I argue that the main instrument to which firms have access that markets do not is flat…. The exercise of fiat through hierarchy is…a very conscious, deliberate, and purposeful way of accomplishing coordination (p. 324).

A similar view characterizes the more recent, incomplete-contracting school.20

This paper incorporates an authority-based—i.e., managerially-based—definition of the firm, consistent with Coase, Simon, and Williamson. Accordingly, the problem of choosing an organizational mode is equivalent to asking when Z’s right to postpone selection of specific acts to be performed by Y (and Z) results in more profitable operation of the enterprise, than an ex ante specification of particular duties to be performed by Y and Z in a market contract, followed by potential renegotiation.

2B. Behavioral Assumptions

We assume bounded rationality on the part of economic actors. This leads to an important corollary: cognitive limitations imply that no two individuals possess identical stocks of knowledge. However, unlike the opportunism-based approach, we focus on honest dealings between individuals.

As to bounded rationality, we suppose that individuals are "intendedly rational, but only limitedly so" (Simon 1957, p. xxiv). As a result, they possess finite cognitive abilities. Williamson (1975) characterizes these as follows:

Bounded rationality involves neurophysiological limits on the one hand and language limits on the other. The physical limits take the form of rate and storage limits on the powers of individuals to receive, store, retrieve, and process information without error. . . . Language limits refer to the inability of individuals to articulate their knowledge or feelings by use of words, numbers, or graphics in ways which permit them to be understood by others. Despite their best efforts, parties may find that language fails them (possibly because they do not possess the requisite vocabulary or the necessary vocabulary has not been devised), and they resort to other means of communication instead. Demonstrations, learning-by-doing, and the like may be the only means of achieving understanding when such language difficulties develop (pp. 21, 22).

Thus tacit knowledge—knowledge that can be acquired only through personal experience (e.g., Polanyi 1962; Nelson and Winter 1982; Nonaka 1994)—is a result of bounded rationality.

A corollary of bounded rationality is that no two individuals possess identical stocks of knowledge, because cognitive limitations prohibit one person, such as Y, from absorbing the entire accumulated knowledge and skills of another, such as Z (and vice-versa). Thus each individual possesses experience, insights, or skills that are to some extent different from that of another.21

As a second principal assumption, we invoke what Williamson (1975) calls “stewardship” behavior.

Opportunism is to be distinguished from…stewardship behavior…. [S]tewardship behavior involves a trust relation in which the word of a party can be taken as his bond…. Opportunistic behavior differs… because it involves making “false or empty, that is, self-disbelieved, threats and promises” in the expectation that individual advantage will thereby be realized… (p. 26).

In this study, each player (Y and Z) is taken to tell the truth to the extent that it is known to that person (i.e., within the constraints of bounded rationality). Further, each knows that both will live up to whatever is agreed. However, Y and Z nonetheless act in their own self-interest (again, within the constraints of bounded rationality). Thus each tries to maximize its own returns as well as it can within cognitive limitations, and without acting “in a self-disbelieved way” (Williamson 1975, p. 27).22

Williamson (1985a) claims that firms will not exist without opportunism, tantamount to arguing that a theory of the firm not based on opportunism is impossible.

I… insist that bounded rationality notwithstanding, [market] contracting would be ubiquitous in the face of nonopportunism…. [E]x post contracting problems… [would be] annihilated by recourse to a “general clause” whereby parties to a contract promise to disclose all relevant information candidly and to behave in a cooperative fashion during contract execution and at contract renewal intervals (p. 66).

Continuing, Williamson holds that

[agents] who value decision participation will… make this clear in the contracts they reach. All adaptations for which net gains can be projected will thereafter be realized without resistance within a community of nonopportunists. Should the nexus of contracts need to be expanded or otherwise altered—for insurance purposes, for example—this will come about by displaying the relevant data in a fully objective way. Reversals of decision roles, due to aging, learning, or the like, will
simply come about whenever net gains are in prospect, the
disposition of these gains being distributed according to the
gainsharing rule negotiated at the outset (p. 66, footnote 1).

With market contracting frictions thus minimized, Williamson maintains that honest behavior will result in market contracting, since transaction costs would be negligible. We disagree with the proposition that honest behavior necessarily results in market contracting. As discussed below, contractual frictions of the sort that can produce firm organization, can occur even when opportunism is held constant at zero.

The key is that some of each person’s knowledge necessarily remains private, as established by the bounded-rationality corollary. Honest persons therefore may disagree about the best course of joint (or even individual) action, or the division of gains. For example, a truthful individual may believe that it has identified a better way to proceed with a joint task, and may argue for a concomitant change in approach that necessitates modifications by the entire group. The person’s “discovery” may produce lengthy and costly negotiation, which includes efforts to convey to the others both the originator’s analysis and the knowledge on which it is based. Because of irreducible differences in the knowledge possessed by the involved individuals, adoption of the innovation may not be automatic. Indeed, objections may never be surmounted, despite the originator’s best (i.e., lengthy and costly) efforts.

Thus we disagree with Williamson that “[a]ll adaptations for which net gains can be projected will . . . be realized without resistance within a community of nonopportunists.” In our example, while the originator projects net gains from the innovation, those with whom it cooperates may believe, even after exhaustive discussion, that no gains, smaller gains, or even losses will result. Thus an adaptation may not “be realized without resistance” under market contracting (or, indeed, may never be realized), even though net gains are projected (by at least the originator) and all are nonopportunists. The parties may have different expectations as to the nature of future gains (or losses), even after each, acting honestly, does its best to explain its reasoning to the others and to understand the alternative positions. Irreducible differences in the individuals’ knowledge can lead them to make different judgments or expect different outcomes. Moreover, truthfulness in expressing an opinion or providing data need not erase all differences between the individuals’ knowledge and judgments. Honest behavior, including the fullest possible disclosure of relevant data, does not guarantee that an “objective truth” will emerge that is easily recognized by all parties, as Williamson implies. The limits on information processing and knowledge articulation that Williamson describes (in an earlier-quoted passage) as characteristic of bounded rationality, are the very factors that can prohibit honest individuals from reaching easy agreement. Similarly, it cannot be presumed that “[r]eversals of decision roles, due to aging, learning, or the like, will simply come about whenever net gains are in prospect . . . ,” as Williamson maintains. Because individuals possess different knowledge, honest persons may disagree as to the best allocation of individual responsibilities, or whether a particular arrangement of decision roles has the potential to generate net gains.

We conclude that honesty does not rule out intense disagreement or haggling. Each party, acting truthfully, may have a different view of the factors (or their relative future importance) that should be taken into account in designing present and future courses of action based on predictions of uncertain, present or future, exogenous or endogenous, realities. For example, individuals’ views of the best course of action may depend on their views of what a business rival will do, which presidential candidate or political party may win the next election, or what policies the new president or government will end up adopting (and how vigorously they will be implemented). The latter may include anticipating what particular political compromises will be struck (e.g., tax rates and mix, antitrust views, historical attitudes towards particular industries, dependence on the support of particular interest groups, etc.). It is a commonplace that, even after (interminable) discussion, honest people still can disagree about what is going to happen. As a consequence, truthful individuals honestly may disagree about the best present and future course of action for their business activities. Or, the parties may possess different mind-sets generally. Discord fundamentally derives from personal knowledge that cannot be communicated fully to others at the time of the disagreement. Because each individual possesses different knowledge, even truthful people may not reach agreement, nor can we be certain of the form (e.g., “rule-governed”) that concord will take if it can be reached. Expressed in terms of the tenets of the incomplete contracting school, there may be no contract interval short enough so that the parties are in complete agreement as to the preferred course of action. Because of irreducible differences in perspectives, experience, or skills, even the tiniest contract period may still involve substantive disagreement and hence negotiation and friction between truthful parties.
Honest behavior, therefore, need not do away with firm organization. More specifically, honesty does not imply absence of transaction costs, i.e., of the costs—frictions—of using market contracting (Coase 1937). Truthfulness does not guarantee automatic agreement, easy recognition of a present or future "common interest," or absence of haggling. Instead, these exist only if we eschew bounded rationality itself. Contrary to Williamson, frictions between economic actors can occur without opportunism, because of inevitable, irreducible differences in their knowledge. Expressed in another way, even in the absence of opportunism, knowledge-based transaction costs can exist. Further, knowledge variances between individuals generally will persist even as the contract interval shrinks, despite the individuals' best efforts to the contrary. These differences, as important to our view as opportunism is to Williamson's perspective, pave the way for a knowledge-based theory of the firm.


Our thesis is that the organizational mode through which individuals cooperate affects the knowledge they apply to business activity. The difference in the knowledge that is brought to bear under the two organizational modes, once anticipated, impacts the choice of mode itself. As we demonstrate below, no single organizational mode is always superior. Our objective is to establish a generalizable, empirically relevant relation between an organizational mode and the situations for which it is optimal, holding the probability of opportunistic behavior constant (at zero).

Organizational mode influences the knowledge applied to business activity in two ways. These concern (1) how the parties' starting knowledge endowments are blended and used, and (2) how learning or developments occurring during the course of the work are taken into account. Organizational mode thus affects both (1) the way in which static (i.e., presently possessed) knowledge is employed, and (2) the dynamics of future knowledge acquisition and response to new developments. We refer to (1) and (2), respectively, as the "knowledge-substitution" and "flexibility" effects of organizational mode. In the theory developed below, an individual chooses between firm organization and market contracting according to which causes the individual to apply the more valuable understanding (net of applicable costs) to business activity, based on the combination of knowledge-substitution and flexibility effects and independent of opportunism.

3A. The Knowledge-substitution Effect

The knowledge-substitution effect concerns how presently held knowledge is applied to the activity. Under either organizational mode, Y obviously has access to its own knowledge. One way for organizational mode to affect the understanding Y applies to the activity, is to influence how Y uses Z's knowledge. Z can make its insights available to Y in many ways. A primary avenue is by giving directions (through Z's role as manager under firm organization) or making suggestions (as Y's contractual partner under market contracting). The directions or suggestions are outcomes of Z applying its understanding to the particular problem or issue at hand.

Masten points out a fundamental, practical implication of the authority relationship in a firm, as compared to market contracting.

The overriding consideration expressed by the courts in...[disputes over the nature of a particular transaction] is the control exercised by an employer and, especially, whether the latter is concerned with the manner in which the work is performed and not solely with its outcome. In evaluating [whether a] transaction [occurs within a firm or as part of a market contracting arrangement], "the first, and seminal, inquiry is whether the alleged employer...has the right to control...the details of the alleged employee's work"... whereas, "an 'independent' contractor is generally defined as one who in rendering services exercises an independent employment or occupation and represent his employer only as to the results of his work and not as to the means whereby it is to be done"... (Masten 1988, p. 186, citations omitted).

Under market contracting, Y agrees to provide a service for a price. It is up to Y to decide how to produce the contracted-for service. In contrast, under firm organization, Z has the right to "control the details" of Y's performance, as Masten indicates. Z can direct Y as to the individual steps Y will take in creating a desired outcome. Indeed, exercising control as to indi-
vidual work elements is an essential and legally-recognized aspect of Z “managing” Y under firm organization, as compared to independent contracting, consistent with our earlier discussion in Section 2A.

Remembering the focus on polar cases, if the market relationship is not to cross the boundary of becoming an authority-based (employment) one, then Y’s own knowledge must be the final guide to its behavior. While Z, as market contracting partner, may make recommendations to Y during contract negotiations or afterward, Y will implement only those steps or actions that it comes to understand and agrees to perform, likely as part of a negotiated package. For as an autonomous contractor, Y retains “the right or power of self-government” [Merriam-Webster’s 1994], or otherwise “works for itself.” Z’s knowledge thus will be used by Y if Y is able, either during contract negotiations or afterwards, to internalize enough of this knowledge so as to agree to act in a way different than Y would without it.

In contrast, under firm organization, Y can be called upon to act on the basis of Z’s knowledge prior to Y internalizing it (or even if Y never does so). This occurs because the employment contract enables Z to give directions to Y. Hence Z can require Y to act (within the employment contract’s boundaries) according to Z’s judgment, rather than in conformity with Y’s own. Knowledge-substitution takes place when Z tells Y to do something that Y, relying on its own knowledge and that which it has managed to absorb from Z, otherwise would not do.30

The knowledge-substitution effect thus involves the relationship between taking an action and internalizing the wisdom upon which it is based. An important question is why Y ever will favor a firm on knowledge-substitution grounds, if firm organization causes Y to take actions that Y otherwise would not. The answer is that knowledge-substitution is a fundamental response to cognitive limitations, having the effect of economizing on them. Knowledge-substitution expands Y’s productive capability without requiring fully concomitant knowledge absorption by Y. In effect, Z’s knowledge becomes blended with Y’s to a greater extent than would occur if Y had to internalize the relevant portions of Z’s knowledge in their entirety. A primary effect of firm organization—of the authority relationship—is to cause an individual to use the knowledge of another before the former fully understands or agrees with it. Conversely, a main effect of market contracting—an autonomous relationship—is to oblige knowledge to be internalized before the individual agrees to modify its actions on the basis of that knowledge.31

Thus for knowledge-substitution to be beneficial, there must be valuable aspects of Z’s knowledge that Y cannot use autonomously until Y fully understands and absorbs them. An example is Z’s tacit knowledge, which can be internalized by Y only through Y’s personal experience, i.e., only after Y has had practice developing or applying the knowledge. As an autonomous contractor, Y may doubt the value of a particular action suggested by Z, and hence not undertake it, because of Y’s lack of experience with—and consequent lack of personal knowledge about—such a step. Under firm organization, on the other hand, Z simply can tell Y to take the action (thereby substituting Z’s knowledge for corresponding elements of Y’s).

Y’s anticipation of knowledge-absorption difficulties can cause it to favor a firm, because this mode is the organizational mechanism through which Y allows Z’s judgment to dominate corresponding elements of Y’s own. Conversely, market contracting is the organizational mechanism through which Y retains the right to exercise its own judgment, even at the cost of rejecting Z’s. This point is related to Rumelt’s (1995, p. 124) conclusion that “[w]ithin the firm . . . it is possible to generate more and richer coordinative activity than can be accomplished in markets . . . it is the gains to greater coordination that rationalize the firm.” It also is consistent with Demsetz (1988, p. 172) insight that “[d]irection substitutes for education (that is, for the transfer of the knowledge itself).” It also underscores Cohen and Levinthal’s (1990) conclusion that sufficient capacity for new-knowledge absorption is critical if a going concern (or individual) is to succeed (and stay independent).32

Seen in another way, firm organization and market contracting each create a particular relationship between people that stands in for the expanded individual abilities that otherwise would be needed to accomplish a similar outcome.33 An essential function of market contracting, on the one hand, is to enable individuals to specialize in different aspects of business activity. Each person need not possess the full range of understanding or skills necessary to complete all aspects of the work by itself. The provisions of the market contract coordinate the individuals’ efforts, so that a unified product (and hence specialization itself) can emerge.

On the other hand, firm organization also enables specialization, since it too provides a means for coordinating individual efforts. However, unlike market contracting, the firm entails a second means for minimizing the impact of limited cognitive abilities. Again looking at polar cases, because the employment con-
tract creates the authority necessary for knowledge-substitution, but a market contract does not, an employee need not internalize all the insights required to choose and carry out an action, while an independent contractor must. Thus firm organization economizes on cognitive limitations through two methods: specialization and knowledge-substitution. In contrast, market contracting economizes through specialization alone.34

The gains from implementing knowledge-substitution will vary with the circumstances at hand. In highly creative or artistic endeavors, for example, knowledge-substitution may be counterproductive or impossible. Some discernment (such as how to create a new product idea, strategy, or marketing campaign) may be so tacit that important aspects of it cannot be communicated. Even if it wants to, the “manager” may not be able to give sufficient directions to permit an employee to come up with the next breakthrough.35 Productivity instead may be enhanced by allowing each party to exercise its judgment independently. In contrast, if a manager’s knowledge is not too tacit, then knowledge-substitution can leverage it (Hamel and Prahalad 1993). More people can use the know-how, because less individual internalization is required. In addition, a previously isolated individual’s understanding purposefully can be leveraged by giving managerial responsibilities (and adopting a firm). Both firm organization and market contracting thus can be used proactively as elements of organizational design, in order to foster, respectively, knowledge leveraging or intellectual independence.

Accordingly, firm organization will not always be the organizational mode of choice on knowledge-substitution grounds. Y will consider which mode provides the greater net value. Either can be the better option as to knowledge-substitution, depending on the relative knowledge-related advantages, balanced against the associated costs of operating each mode.

The cost of operating the firm includes negotiating the Y-Z employment contract, and the time and cost of Y receiving and digesting directions from Z. As discussed by Simon (1951), Y also bears the risk that Z’s postcontractual directions will be, from Y’s perspective, less as opposed to more desirable (notwithstanding that both are truthful and all Z’s directions fall within the contractually established acceptable range). In addition, Y’s evaluation of the quality of Z’s knowledge will affect (interact with) some of the costs of obtaining it. For example, the greater the faith Y has in Z’s judgment, the less may be the risk that Y believes it faces as to the palatability of Z’s directions, the cheaper may be the costs of negotiating the bound-

ary on Z’s discretionary authority in the Y-Z employment contract, and the greater may be the likelihood that the boundary will not be excessively constrained.

Alternatively, the cost of carrying on market contracting includes the expense of negotiating the Y-Z market contract, and the cost of Y internalizing the knowledge of Z. Contract negotiation costs can be expected to be lower, the more quickly and cheaply the parties can transfer enough knowledge, settle differences and arrive at agreement, and the lesser are legal costs. The degree of difference in perspectives between Y and Z will be important in this regard. If, for example, Y and Z come to diametrically opposite conclusions as to what needs to be done, market contract negotiation will be difficult and costly. This factor implies that firm organization is more likely to be preferred, all else equal, the greater is the initial difference in the knowledge, culture, etc. of Y and Z (provided that Y evaluates elements of Z’s knowledge as superior to Y’s).

Thus, depending on the value of Z’s knowledge as compared to Y’s, balanced against the cost of operating the firm as opposed to a market contract, either organizational mode can end up, as to knowledge-substitution, being preferred. It is possible for either organizational mode to yield (1) the better blending and use of the parties’ knowledge, and/or (2) the lower costs. Since Y will have some relationship with Z under either organizational mode (Figure 4), Y may prefer either a stronger knowledge input by Z (under firm organization) or a weaker one (under market contracting), depending on the benefits and costs associated with each. We note that analysis of the knowledge-substitution effect has not depended on the potential for opportunistic behavior.

38. The Flexibility Effect

The flexibility effect accounts for the relative cost, under the two organizational modes, of altering the parties’ duties and responsibilities on an ongoing basis, in order to incorporate learning or unexpected opportunities arising during the course of the work. The flexibility effect thus concerns the dynamics of future knowledge acquisition and application, and response to new, internal or external developments. The modes carry different implications for the costs of adjusting duties and responsibilities to respond to events that were uncertain or simply unanticipated originally, as market contracting requires contract renegotiation in many situations in which firm organization does not. Y will take this difference into account in picking an organizational mode, since it affects the new knowl-
edge that will end up being incorporated into the business activity and/or profitability.

Firm organization and market contracting each incorporate a mechanism for altering duties and responsibilities, in response to developments that were unforeseen or uncertain originally, or knowledge that was obtained after the initial contract was entered into. Under firm organization, the mechanism is an inherent part of the employment contract itself. Without engaging in contract renegotiation, Z, as manager, can reframe previous directions or issue new ones. Under market contracting, in contrast, changes in duties and responsibilities require renegotiating the market contract itself, aside from those changes pre-determined in the original market contract to become operative if certain, specified events occur. As to the latter, it will be difficult, if not impossible, for market contracts to provide for unanticipated acquisition of new knowledge, since it is axiomatic that the content of knowledge that may be obtained in the future is unknown at the present. Thus market contracting flexibility as to unforeseen developments or new knowledge is achieved through engaging in a series of contracts.

Just as contractual mechanisms differ for achieving flexibility under the two organizational modes, so do the associated costs. The cost under firm organization is simply that of operating the firm over the relevant time frame. From Y's point of view, it includes, as before, the time Y spends receiving and digesting new directions, as well as the risk that (a truthful) Z will err by making changes unfavorable to the enterprise in general or to Y's productivity in particular.

Under market contracting, on the other hand, the cost of implementing flexibility is the expense of the market contract renegotiations required to attain it. The renegotiation cost includes two aspects. The first and most obvious is the cost of conducting the renegotiation and redrafting the contract. As a factor contributing to these costs, it must be remembered that Y and Z have different knowledge sets (and potentially different associated linguistic conventions). This difference makes costly the transmission of ideas by Y to Z (and vice versa), which is itself an unavoidable part of the parties explaining to each other the reasons behind their respective new negotiating positions. Moreover, in a market contract, the onus is on the party seeking to convince the other to agree to disturb the status quo. The motivation is the former's perception that new, significant knowledge has been obtained, or that new circumstances or opportunities now beckon. Unless the stand-pat party can be convinced, the pre-existing contract controls (if its term has not ended) or otherwise may exert a powerful force of inertia against change.

A second aspect concerns loss of coherent vision or cohesion as to the aim or conduct of the endeavor itself. This loss may be brought on through engaging in successive rounds of incremental compromises by Y and Z. Because each round of market contracting requires the parties to come to a priori agreement on the acts to be performed by each, the compromises made in one round can become the starting point for the next round. Over a succession of rounds, it is possible for the cohesion or coherence of the visions, plans for, and activity of the endeavor as a whole to be diminished.37

Of course, the less important that flexibility is to the profitability of the endeavor, the less important is any cost difference between the organizational modes in obtaining it. The actual amount and value of flexibility that will be needed for efficient operations may become known only after entering into the contract—whether market or employment. We may think of Y as having an a priori probability distribution in its mind regarding this amount and value. Significant parts of the anticipated value that Y places on firm organization may turn on the amount (mean and variance) of flexibility that it anticipates to be necessary. The cost of flexibility in the market contract (e.g., the cost of renegotiation and redrafting) is an increasing function of the amount of flexibility that turns out to be necessary. Less needed flexibility translates into fewer negotiations and/or less important—and thus less highly contested—negotiations.

Market contract renegotiation costs increase with the continuing degree of difference in perspectives between Y and Z. This difference, in turn, is likely to increase with the frequency, importance, and complexity of changes perceived to be necessary. If Y and Z continue to come to opposite (or very different) conclusions as to what needs to be done, renegotiation will be difficult and costly. The coherence/"short-sightedness" problem discussed earlier also may be intensified. In addition, in a turbulent, hypercompetitive (D'Aveni 1994) environment, the nature of tasks or required coordination may be concomitantly complex, and the sustainability of competitive advantage may be especially uncertain. As a result, the frequency, importance, and difficulty of market contract renegotiation may increase. Similarly, the greater is the anticipated likelihood of frequent, important and involved changes in the market environment, or in the knowledge that will be obtained in the future, the higher is the likelihood of either a very short term original market contract, or
very costly negotiation of the original market contract. A very short term original contract may exacerbate the coherence problem, and costly original negotiation may reflect attempts to anticipate and set the rules for dealing with a potentially chaotic future. This would reflect a parallel increase in the complexity of the work to be accomplished.

Thus, contrary to some views, firm organization is more likely to be preferred on knowledge-based flexibility grounds, the more dynamic and uncertain is the competitive environment. Put another way, firm organization is likely to be preferred when the cognitive limitations imposed by bounded rationality, which are chief contributors to the unforecastability of events, and to the parties' differences in perspectives, may have a large impact on productivity and profitability. In such environments, commitment (Ghemawat 1991) on the part of Y, in the form of agreeing to abide by an employment contract and consequently Z's series of postcontractual directions, may add the advantage of stability of vision or strategic intent (Hamel and Prahalad 1989) to the endeavor. At the same time, the firm may preserve the lower-cost flexibility as to Y's duties, by eliminating the need for successive contract renegotiations.

Accordingly, all else equal, if Y believes a priori that the mean and variance of the distribution on needed flexibility are small, Y is more likely to select market contracting. If the mean and/or variance is high, Y is more likely to select an employment contract with Z. We note, therefore, that either firm organization or market contracting can provide the lower-cost flexibility as to amending duties and responsibilities, to reflect the dynamics of ongoing learning and new opportunities. As before, the effects discussed in this subsection do not depend on opportunistic potential.38

4. Resource-based Predictions of Organizational Mode

Resource-based predictions of Y's choice of organizational mode are displayed in Figure 5. In situation (i), substitution of Z's knowledge for some of Y's provides the higher-value insights for the business activity. In addition, an employment contract produces the cheaper flexibility. Y therefore picks a firm, as this mode dominates on both counts. In situation (iii), these conditions are reversed, leading to Y's unambiguous choice of market contracting.

In situations (ii) and (iv), the direction of advantage is mixed, so Y picks either firm organization or market contracting, depending on the balance between opposing forces. In situation (ii), if the value of substituting Z's understanding for some corresponding elements of Y's is sufficiently large, and if market contract renegotiation provides the cheaper source of flexibility, but not by a lot, then the knowledge-substitution factor dominates and Y opts for a firm. Alternatively in situation (ii), if the value of knowledge-substitution is small enough, and the flexibility-cost savings from market contracting are large enough (e.g., as might occur in a stable, mature industry), then Y picks market contracting. A parallel analysis applies to situation (iv), by reversing these conditions.

As illustrated by Figure 5, the resource-based approach thus can be extended to a theory of the existence of the firm that incorporates predictive elements. Associated with the knowledge-substitution and flexibility effects are concepts such as strategy formulation, organizational culture, motivation, learning, organizational routines, barriers to transfer of knowledge, and strategic leadership, among others. For example, in assessing who has the superior judgment, Y is likely to consider how well Y and Z each are able to evaluate market opportunities and translate them into ideas for action leading to sustained competitive advantage. More broadly, Y's assessment of the value of firm organization will depend on the leadership and vision of Z—the value of Z's insights as compared to Y's. In addition, the organizational culture and shared values that can be built up under firm organization as opposed to market contracting, are likely to impact both the success of knowledge-substitution and flexibility costs. Research on these and other concepts can be applied to the problem of predicting organizational mode, especially through their bearing on knowledge-substitution and adaptive flexibility.39

5. Comparison of Resource- and Opportunism-based Predictions

In Figure 6, we compare resource-based predictions of organizational mode with corresponding opportunism-
related ones. It is straightforward to see how the identical forecasts for situations (i) and (iii) in Figure 6 arise. In situation (i), (a) the probability of opportunistic behavior is high, thus favoring a firm, and (b) a firm provides the more valuable, opportunism-independent knowledge.\(^{40}\) Both approaches therefore anticipate firm organization. Situation (iii), in which neither does so, derives from reversing (a) and (b). We now turn to situations (ii) and (iv), which contain opposing forecasts.

Looking first at situation (ii), firm organization is not warranted from an opportunism-based view. Because the probability of opportunistic behavior is low, \(Y\) opts for market contracting. In contrast, the resource-based approach predicts the reverse: \(Y\) chooses a firm.

The key to the difference is that, in the choice of organizational mode, opportunism-independent knowledge considerations can outweigh opportunism-based ones. Even when opportunistic potential is too low to justify a firm, as occurs in situation (ii), firm organization nonetheless is selected from the resource-based perspective if it results in the more valuable knowledge applied to the activity. The resource-based approach thus implies that opportunistic potential is not all that \(Y\) considers in choosing an organizational mode.

Turning to situation (iv) of Figure 6, (a) the probability of opportunistic behavior is high, and (b) market contracting provides the more valuable, opportunism-independent knowledge. While the opportunism-based approach anticipates a firm, resource-based theory forecasts \(e\)ither a firm or market contracting. The specific resource-based prediction depends on the balance of opposing forces.

As to this balance, it is not inconsistent with the resource-based approach that high probability of opportunistic behavior will lead to a firm, even if, absent the opportunism factor, the firm provides less valuable knowledge than market contracting. The benefit of minimizing opportunism, if sufficiently great, outweighs the loss imposed by this mode essentially “costing too much” for the opportunism-independent knowledge that it provides. Alternatively, the gain from firm organization’s reduction of opportunistic potential may be too weak to compensate for the loss this mode causes by providing the less valuable, opportunism-independent knowledge. In this case, the resource-based approach anticipates market contracting.

To establish that the resource- and opportunism-based theories have distinct predictive content, we must show that situations (ii) and (iv) are empirical possibilities. For example, will we ever observe a situation in which the probability of opportunistic behavior is low, but the firm nonetheless provides, on net, the more valuable opportunism-independent knowledge (situation (ii))? We must examine whether factors that can produce a low probability of opportunistic behavior are compatible with those that can cause the firm to provide the more valuable opportunism-independent knowledge, and vice-versa.

To show that these two sets of factors can be compatible, we supply a specific example. If \(Y\) and \(Z\) each stand to gain the same amount from cooperation and each must make the same size idiosyncratic investment in order to make the cooperation work, then opportunistic potential is balanced between the parties. If opportunistic potential is balanced, the probability of opportunistic behavior is low, since the parties are in a “stand-off” situation (Klein and Leffler 1981, Williamson 1983). Yet, equal specific investment by \(Y\) and \(Z\) and equal gain from cooperation does not imply that the factors giving rise to the knowledge-substitution and flexibility effects of firm organization must be absent. Equal specific investment and equal gain are compatible, for example, with a dynamic, highly uncertain competitive environment in which the flexibility and strategic coherence that a firm provides may be of great value. Similarly, equal specific investment and equal gain are compatible with the barriers to knowledge absorption that underlie the knowledge-substitution effect. Expressed in another way, equal specific investments and equal gains need not preclude obstacles to knowledge absorption or flexibility gains to firm organization. A parallel example can be developed to demonstrate the empirical possibility of situation (iv) in Figure 6.\(^{41}\)

Analysis of Figure 6 also may help shed light on previously anomalous empirical findings. In an influential study of the costs of organization in naval shipbuilding, Masten et al. (1991) found that

the second-stage estimates indicate that the correlation between human asset specificity and the likelihood of integra-
tion found in the first stage is a consequence of a decrease in internal organization costs rather than the increase in the costs of market exchange that the [opportunism-based, transaction-cost] theory predicts (p. 19).

Masten et al. [1991, p. 21] further comment that "[f]ully satisfactory explanations for these findings are elusive." Masten et al.’s results in this regard appear consistent with the analysis of situation (ii). From the resource-based perspective, a firm is chosen in situation (ii) because the human asset specificity—dependence of $Y$ on direction by $Z$—thereby produced is more productive than the lesser level of dependence that would be generated by market contracting.

Speaking more broadly, the relationship of the resource- and opportunism-based theories is diagrammed in Figure 7. As level (a) indicates, the basic purpose of firm organization under both approaches is to increase the productivity of cooperating individuals. The theories, therefore, are consistent as to this fundamental reason for a firm.42

As indicated in level (b) of Figure 7, the divergence between the approaches is caused by different—opposite—ceteris paribus assumptions. We hold the probability of opportunistic behavior constant (at zero), in order to examine the opportunism-independent, knowledge-based influences of organizational mode more closely. Opportunism-based theory does the reverse. It implicitly holds opportunism-independent effects of organizational mode constant, in order to ascertain how organizational mode affects opportunistic potential.

Expressed in a different way, both theories are parsimonious in that each holds empirically relevant factors fixed in order to analyze others. Recognition of opportunism-based theory’s implicit ceteris paribus assumption is the basis for Conner’s (1991) earlier observation that:

Transaction cost theory assumes that the same productive activity can be carried on either within a 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 (p. 142).

It also may relate to Demsetz’ (1988) statement that:

The emphasis that has been given to transaction cost...dims our view of the full picture by implicitly assuming that all firms can produce goods or services equally well (p. 147).43

Foss (1996) also maintains that:

One problem in the contractual approach is that it is often implicitly assumed that what one firm can do on the level of production, another firm can do equally well, so that differences in economic organization are not allowed to turn on differences in production costs... (p. 474).

---

**Figure 7** Comparison of Resource- and Opportunism-based Approaches

(a)

| Objective: Increase Earnings Through Greater Productivity |
| Basic Condition: Bounded Rationality |
| Incentive for an Individual’s Investment in Productivity-enhancing Actions |
| Convergence |

(b)

| Focus on Opportunism |
| Focus on Opportunism-independent Knowledge |
| Divergence |

(c)

| Choice of Organizational Mode |

(a) Opportunism-based Approach

(b) Resource-based Approach
The opposing predictions in situations (ii) and (iv) of Figure 6, and the contrary ceteris paribus assumptions (Figure 7) that drive them, indicate the need for future theory development that incorporates both resource- and opportunism-based considerations into an explanation of the firm's existence. The now-conflicting predictions in Figure 6 need to be replaced, through creation of new theory, with single, unified ones. The new forecasts would take into account, and the new approach would explain, how nonopportunism-based and opportunism-based factors interact and are balanced in the choice of organizational mode.

6. Limits to Size and Scope of the Firm

It is not possible within the present paper to address in detail, a theory-of-the-firm's second question—limits to firm size and scope. Instead, our objective is to suggest a general approach through which this question subsequently might be treated from a resource-based perspective. Referring to limits on firm size as a "chronic puzzle," Williamson (1995a) holds that

The upshot is that limitations on firm size of a comparative institutional kind have yet to be described (p. 135).

The reason for this conclusion centers on the concept of "selective intervention." Its motivating question is, "Why can't a large firm do everything a collection of small firms can do and more?" (Williamson 1985a, p. 131) As Williamson puts it,

Selective intervention is really the key. The idea is to replicate the market mode within the firm in all respects save those where the intervention is the source of expected net gains. If hierarchical control is reserved for the latter and if the firm replicates the market in all other respects, then the firm never does worse than the market (by replication) and sometimes does better through selective intervention). Accordingly, the firm will everywhere do as well as and will sometimes do better than the market (1992, p. 339).

Williamson holds that the central issue in a theory of the limits of firm size is why selective intervention breaks down. The question is why a perfect blend of authority and market relationships cannot be achieved within the same firm, so that firm organization always outperforms the alternative of a collection of market contracts. Expressed differently, the issue is why a firm cannot accomplish a perfect mix of the right to direct employees' actions with the absence of rights to do so.44

A central challenge for resource-based theory is to evaluate whether selective intervention is an appropriate approach for assessing opportunism-independent, knowledge-based limits to the size and scope of the firm, or whether we should return to the marginal analysis originally suggested by Coase (1937).45 For example, a question as to selective intervention might be: If there are places within a firm where firm management should not possess the right to direct employees, then what is the gain from having those places inside the firm? Resource-based theory as to existence of the firm predicts that if the manager adds, on net, insufficiently valuable knowledge to an employee's activity, then it will not be carried on inside the firm. For it to have independent force, the concept of selective intervention implies that factors additional to or simply different from those affecting whether an activity is included in the firm in the first place may determine the scope of an already established firm. The selective intervention concept thus may be reaching for a source of indivisibility between the activities or assets of an established firm, which simultaneously is not a factor (or is less of one) in deciding whether the activity or asset will be part of the firm in the first place. Such a factor could account for why parts of an already existing firm might be retained even though a theory of the existence of the firm would indicate that these parts would not be added de novo to the firm, or vice-versa.

We believe that the selective intervention concept profitably can be interpreted as calling for a deeper understanding of the fundamental sources of incommensurability between the polar organizational modes. Indivisibilities pertaining to a firm's size but not to its existence, should they exist, are likely to be outgrowths of incommensurabilities between market contracting and firm organization themselves. Moreover, hybrid organizational modes, such as joint ventures or alliances, may be practical attempts to minimize inherent incommensurabilities between the polar modes, and thus may constitute practical attempts to achieve "selective intervention."

7. Conclusion

Operationalizing the resource-based view through development of predictive theory is a significant challenge. We attempt to contribute by advancing a resource-based theory of the firm. Appreciating the substantial intellectual contribution of the opportunism-based, transaction-cost view, we seek to complement and build on it, through articulation of a knowledge-based perspective that is independent of opportunistic considerations. A resource-based theory of the firm is an important step, we believe, toward
creation of more general, resource-based theory of performance differences between firms. It also provides knowledge-related reasons for vertical integration itself.

There have been at least three seminal contributions to the theory of the firm. Coase (1937) delineates the area as a research topic and establishes the comparative organizational reasoning crucial to a theory of the firm. Coase (1937) also introduces the fundamental concept of transaction costs, which he defines broadly. Simon (1957) advances the motivating behavioral assumption of bounded rationality. Simon (1951) also establishes the employment—authority—relationship as the incisive distinction between a firm and market contracting. Williamson (1975, 1985a) explores a powerful aspect of bounded rationality in the context of choosing an organizational mode: opportunistic potential. The predictive theory developed by Williamson operationalizes an important aspect of the transaction-cost approach.

While opportunistic potential may be an important ramification of self-interested, cognitively limited individuals choosing an organizational mode, it is not the only one. Holding that transaction costs arise only because of opportunism or the potential for it is, in our view, an incomplete interpretation of Coase’s, Simon’s, and Williamson’s work. Transaction costs, or frictions within independent-contractor relationships, also may arise for knowledge-based reasons that are independent of opportunistic potential. Expressed in another way, bounded rationality has other, important implications besides opportunism for how economic actors coordinate their productive activity. The implication upon which we focus is the difference in knowledge applied to business activity under the polar organizational modes of market contracting and firm organization.

The results obtained in this paper complement those of Barney (1991) and others concerned with predictive power as to identifying “resources.” A theory of the firm indicates a subset of the substantive areas of endeavor likely to be associated with a company’s advantage vis-a-vis other firms. An implication of this study is that important resources, i.e., factors contributing to above-normal earnings, may include those materially affecting (1) the quality of managerial as opposed to employee judgment (including, for example, organizational culture and human resources policies), and (2) the cost of implementing flexibility as to what employees should do. Barney’s (1991) tests of value, rarity, imperfect imitability, and nonsubstitutability can be applied to identify the specific assets within these areas that have the capability to generate competitive advantage.

An important future step is to explore the interaction between nonopportunism-based knowledge considerations and opportunism-related factors in the choice of organizational mode. The present knowledge-based approach also may be extended more specifically to organizational entities and to more complicated or hybrid modes, such as multi-level corporations, alliances, joint ventures, and legal partnerships. Examining situations involving physical, as well as human, capital also would be useful. In addition, opportunism-independent, knowledge-based limitations to the size and scope of an existing firm need to be explored.

Acknowledgment
We are grateful for the helpful comments and suggestions provided by Raphael Amit, Jay Barney, Jeffrey Conner, Ansel Karnani, Scott Masten, Will Mitchell, Elaine Mosakowski, Joanne Oxley, Richard Rumelt, José Santos, Ted Snyder, Neal Stoughton, Gordon Walker, and three anonymous referees. Kathleen Conner received financial support for this paper from the University of Michigan Business School Summer Fellowship program.

Appendix

Defining Firm Organization
An essential element of the debate within the economics literature is whether an authority relationship can be said to separate firm organization from market contracting.46 The use of authority to distinguish between firms and markets begins with Coase’s (1937) original treatment. His approach depends on authority of employer over employee under firm organization, in contrast to market contracting. Coase defines a firm in terms of the managed direction of resources, instead of coordination being achieved through market processes.

Outside the firm, price movements direct production, which is co-ordinated through a series of exchange transactions on the market. Within a firm, these market transactions are eliminated and in place of the complicated market structure with exchange transactions is substituted the entrepreneur coordinator, who directs production . . . A firm . . . consists of the system of relationships which comes into existence when the direction of resources is dependent on an entrepreneur (Coase 1937, pp. 333, 339).

While Coase’s (1937) and Simon’s (1951) managerially based distinction between firms and markets may be intuitively appealing, Masten (1988) points out a weakness. He argues that neither Coase (1937) nor Simon (1951) establishes the firm as a discrete organizational mode as compared to market contracting. That is, Coase and Simon differentiate firms and markets as a matter of degree rather than kind. Indeed, as noted by Masten, Coase foreshadows this
problem in his statement that:

Of course, it is not possible to draw a hard and fast line which determines whether there is a firm or not. There may be more or less direction (Coase 1937, p. 337).

And as to Simon (1951), Masten holds that

Under . . .[Simon's] definition . . . the employment relationship is analytically indistinguishable from any contract in which one party is empowered to alter some aspect of performance unilaterally. An example would be a fixed price, variable-quantity contract in which the buyer has the “authority” to determine the volume of trade under the agreement and can thus “direct” the production level of the seller. Such arrangements are not at all uncommon in long-term contracts. But although they conform in a technical sense to Simon's definition of an employment transaction, the relationship between the buyer and seller in such contracts would not generally be considered that of employer and employee. At best, the distinction between employer and supplier is again a matter of degree. Only the details and not the type of contract entered separate an employment from a commercial transaction, and neither the label firm nor employer has any force beyond the provisions explicitly adopted in the contract itself (Masten 1988, p. 182).

The logical problem generated is substantial. If firms and markets differ only as to the degree of authority exercised, then it is difficult to identify in a general way the separation between market contracting and firm organization. For example, how much authority is needed for the relationship to constitute firm organization? Without a generalizable answer, a managerially based theory of the firm remains remote. In order to predict choice of organizational mode, it is necessary to be able to distinguish between modes.

An implicit answer to the question of how much authority is needed—as well as a significant challenge to the authority-based view as a whole—is presented by Alchian and Demsetz (1972). They argue that the authority exercised within a firm is no stronger or any different than that which occurs under market contracting. The Alchian-Demsetz approach therefore solves the problem of how much authority is enough to produce firm organization, by eliminating authority altogether as a variable of interest in a theory of the firm. According to Alchian and Demsetz (1972):

It is common to see the firm characterized by the power to settle issues by fiat, by authority, or by disciplinary action superior to that available in the conventional market. This is delusion. The firm does not own all its inputs. It has no power of fiat, no authority, no disciplinary action any different in the slightest degree from ordinary market contracting between any two people. I can "punish" you only by withholding future business or by seeking redress in the courts for any failure to honor our exchange agreement. That is exactly all that any employer can do. He can fire or sue, just as I can fire my grocer by stopping purchases from him or sue him for delivering faulty products. What then is the content of the presumed power to manage and assign workers to various tasks? Exactly the same as one little consumer's power to manage and assign his grocer to various tasks . . . . To speak of managing, directing, or assigning workers to various tasks is a deceptive way of noting that the employer continually is involved in renegotiation of contracts on terms that must be acceptable to both parties (p. 777).

The Alchian-Demsetz view gives rise to defining the firm as a "nexus of contracts" (Jensen and Meckling 1976). The nexus-of-contracts approach underlies a significant portion of the agency-cost literature as well as of the "monitoring" school of transaction-cost economics (see, e.g., Williamson 1985a). Firms and markets are not seen to be separable entities on analytically substantive grounds. Firms are taken to be "legal fictions," i.e., little more than nodes of intersection between (market) contracts. An implication is that the theory of the firm (with the firm taken as a separable institution or substantive entity) would be replaced with a theory of contract. The latter would investigate why certain types of contracts emerge under certain circumstances. Presumably, the same type of contract could exist either under the legal fiction of "firm organization" or as a bona-fide market relationship.14

In contrast, for firm organization to function as a "governance mode," firms and markets must be able to be defined as "discrete structural alternatives" (Williamson 1991b). Contrary to the nexus-of-contracts view, a substantive difference must exist between firms and markets for opportunism-based, transaction-cost theory (and ours) to apply. The difference between the polar organizational modes needs to be one of kind rather than simply degree. As Williamson puts it,

[The] object [of transaction-cost analysis] is to match governance structures to the attributes of transactions in a discriminating way. Microanalytic attention to differences among governance structures and microanalytic definition of transactions are both needed in order for this to be accomplished [Williamson 1981, p. 1544].

As to the nexus-of-contracts view, he maintains that

[The] argument that the firm "has no power of fiat, no authority, no disciplinary action any different in the slightest degree from ordinary market contracting" (Armen Alchian and Howard Demsetz, 1972, p. 777) is exactly wrong: firms can and do exercise fiat that markets cannot (Williamson 1994, p. 325).

Masten (1988) and Williamson (1991a, 1991b, 1994) argue that the authority exercised under firm organization is of a different kind, as compared to the ability of parties to "direct" one another under market contracting. The authors analyze the law governing firm relations as opposed to market ones. Masten's (1988) purpose, for example, is to

explore the status of the employment relationship in the legal system and a comparison of corresponding doctrines in commercial contract law to see whether legal rules establish an institutional basis for the advantages and limitations most commonly associated with internal organization . . . . [T]he issue of whether the firm is a distinct institution is ultimately a question of fact: Are there mechanisms or sanctions available in employment transactions that are not similarly available to independent contractors? (pp. 182, 185).
Both Masten and Williamson find that the law indicates the traditional, authority-based notion of the firm. Masten observes, for example, that

[\text{\textquoteright}even a cursory examination of the case law governing the relationship between employers and employees\text{\textquoteright} reveals a set of obligations and responsibilities that are indeed unique to employment transactions and which often coincide precisely with the traditional emphasis in economics on the information and authority advantages of internal organization (p. 185).]

Masten (1988, p. 195) indicates that because of “differences in legal defaults, sanctions, and procedures governing commercial and employment transactions,” there is “a basis for special managerial authority or access to information” traditionally attributed to firm organization. Williamson (1994, p. 325, emphasis in original) maintains that “the (implicit) contract law of hierarchy is that of forbearance,” under which “courts will refuse to hear disputes between one internal division and another” over technical issues that otherwise would be heard if the parties instead were market contractors.

While legal analysis is valuable in establishing differences of fact with respect to firm organization and market contracting, it leaves open the issue of the cause of the differences. In particular, why is authority—managerial direction—exercised by one party over another in some circumstances, but not in others? As Williamson (1994) comments,

\begin{quote}
I conjecture that forbearance law is the evolutionary product of business and the law grogping for a contractual logic that “worked,” which is to say that forbearance law was not consciously invented but mainly evolved (p. 325).
\end{quote}

A theory of the firm emphasizing authority differences between the polar organizational modes can be seen as a way to explain why these differences have come to exist.

Endnotes
1 Examples of the knowledge-related emphasis in the resource-based and related literature include the following. Winter (1988) holds that

Fundamentally, business firms are organizations that know how to do things… Firms perform their function[s] as repositories of knowledge… (pp. 175, 177).

Prahalad and Hamel (1990), in developing the concept of core competence, state that:

Core competencies are the collective learning in the organization, especially how to coordinate diverse production skills and integrate multiple streams of technologies (p. 82). According to Teece et al. (1990):

[\text{\textquoteright}It is not only the bundle of resources that matter, but the mechanisms by which firms learn and accumulate new skills and capabilities, and the forces that limit the rate and direction of this process (p. 11).]

And Kogut and Zander (1992) comment as follows:

[The theoretical challenge is to understand the knowledge of a firm as leading to a set of capabilities that enhance the chances for growth and survival (p. 384).]

Lippman and Rumelt (1982) formally link the existence of privately held knowledge—in the form of causal ambiguity—to a firm’s ability to earn above-normal returns through productive activity. As expressed separately by Barney (1986):

When firms seeking to acquire resources to implement a strategy…and firms who currently own or control these resources…have exactly the same, and perfectly accurate, expectations about the future value of product market strategies before they are actually implemented, then the price of the resources needed to implement these strategies will approximately equal their value once they are actually implemented… In such markets, all pure profits that could have been had when the strategy in question was implemented will be anticipated and competed away (p. 1233).

Cyert et al. (1993) also conclude:

It is the existence of knowledge of internal production techniques or external opportunities in the hands of a small number of firms that creates the market imperfections necessary to generate rents for the firm. Put another way, it is proprietary knowledge that creates a comparative advantage for the firm (p. 57).

2 Opportunity-based theory suggests that firms exist because of the potential for cheating that accompanies investments, when value “depends on the continuation of a particular relationship” (Milgrom and Roberts 1990, p. 62), creating asset specificity. Firm organization is seen as a method for reducing the net incentive to act opportunistically in situations in which asset specificity is present (see, e.g., Williamson 1975, 1985a).

3 Our interest coincides with Alchian and Demsetz’ (1972) definition of teamwork:

Team production . . . is production in which several types of resources are used and the product is not a sum of separable outputs of each cooperating resource. An additional factor creates a team organization problem—not all resources used in team production belong to one person (p. 779).

4 Hereafter, by “organizational mode,” we mean whether actors join in a firm, or cooperate by means of a market contract (the polar cases).

5 We do not suppose that all of an employee’s actions are so influenced, but rather that some of them necessarily are, if firm organization (management) is to have any effect.

6 Throughout the paper, we use “it” as a personal pronoun, to avoid gender references and because the theory applies more broadly to entities, not just natural persons.

7 We are grateful to an anonymous referee for suggesting that we address this issue.

This is a central issue in the strategic management literature (see, e.g., Barney 1986, 1991; Rumelt et al. 1994).

As Schumpeter (1950, p. 85) put it in another context, this type of competition (from market contracting) "acts not only when in being but also when it is merely an ever-present threat. It disciplines before it attacks."

Components 1 and 2 in Figure 2 include both asset stocks and flows, in Dierickx and Cool’s (1989) sense.

From the early strengths-weaknesses-opportunities-threats (SWOT) analysis and accompanying concept of distinctive competence (e.g., Christensen et al. 1978), to the generic strategies and value chain approaches developed by Porter (1980, 1985), the emphasis in the strategic management field has been on a firm’s competitive advantage as realized through superior productive activity, including, among other aspects, procurement of materials, manufacturing, distribution, marketing, and human resource policies. Given its intellectual roots within the strategic management field, it is not surprising that the resource-based perspective carries through the field’s traditional emphasis on the firm “search[ing] for rent” (Bowman 1974, p. 47) through how it gains advantage from aspects of productive activity. This emphasis can be seen in Wernerfelt’s (1984) initial treatment and virtually all of the resource-based and related literature.

Conner states in an earlier paper that

[b]ecause resource-based theory is centrally motivated by understanding performance differentials between firms . . . the firm’s 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.

(Conner 1991, p. 139, italics in original)

Foss (1996, p. 473) holds that such “separation of the existence issue into two parts is a non-issue.” He maintains that a single, general explanation exists under opportunism-based, transaction-cost theory that “explains both the existence of the firm relative to the market and why a particular firm owns a particular combination of assets/resources.” Our view is that existence of a firm, as opposed to other firms absorbing its assets, depends on the firm’s performance in comparison to other firms, i.e., on creating sustained competitive advantage. While a firm’s survival in a firm-versus-firm context is affected by factors related to the theory of the firm (Component 2 in Figure 3), these are not the only influencing elements. Differences in the “base” assets controlled by competing firms also are significant (Component 1 in Figure 3). Thus factors additional to those analyzed in a theory of the firm are required for a theory of performance differences between firms.

An additional reason for interest in a theory of the firm concerns its second aspect, viz. factors limiting the size and scope of firms. These factors relate to questions of diversification, vertical integration, and strategic alliances, among others, which are central in the strategic management literature.

As Alchian (1950, p. 367) puts it, “[k]nowing why ‘firms’ exist without clearly denoting what is meant by a ‘firm’ creates opportunities for dispute or nonrefutable generalities.”

As described by Coase (1937), “[o]utside the firm . . . production . . . is co-ordinated through a series of exchange transactions on the market” (p. 333).

This view is consistent with that of Coase (1937), who earlier wrote that it is important to note the character of the contract into which a factor enters that is employed within a firm. The contract is one whereby the factor, for a certain remuneration (which may be fixed or fluctuating), agrees to obey the directions of an entrepreneur within certain limits. The essence of the contract is that it should only state the limits to the powers of the entrepreneur. Within these limits, he can therefore direct the other factors of production . . . All that is stated in the contract is the limits to what the persons supplying the commodity or service is expected to do. The details of what the supplier is expected to do is not stated in the contract but is decided later by the purchaser. When the direction of resources (within the limits of the contract) becomes dependent on the buyer in this way, that relationship which I term a “firm” may be obtained (p. 336-337, italics in original).

By “postcontractually,” we mean after the formation of the contract, not after its term has ended.

Indeed, Williamson holds that “I expressly identify first as one of the distinguishing features of hierarchies, as compared with markets, in my very first article on transaction costs . . . .” (1992, p. 339). This might seem contrary to his treatment of “peer groups,” which he takes to be “the simplest non-market alternative” (1975, p. 41). Williamson holds that peer groups “do not entail subordination” (1975, p. 42). However, he later indicates that “the hypothetical Peer Group is a romantic vision” and that “the Pure Peer Group is unattainable” (1985b, p. 243, 244).

Receiving its start from Grossman and Hart (1986), the incomplete contracting perspective incorporates opportunistic behavior, but includes conceptual differences as compared to Williamson’s framework and also involves mathematical formalization through game theory. A chief difference is that asset specificity is not seen necessary to generate firm organization. Instead, “the key to evaluating the efficacy of market transactions is the costs of negotiating suitably detailed short-term contracts” where “short-term refers to a period short enough so that all the information that is relevant for current decisions is already available” (Milgram and Roberts 1990, p. 65). The basic concept is that if short-term intervals are allowed to be brief enough, then it is possible to define one in which, even given
bounded rationality, the parties do not disagree regarding what should be done. The problem over a longer period thus is the cost of negotiating a sequence of such short-term contracts. It is held that a firm will be preferred when bargaining costs associated with negotiating this sequence (including opportunism-based ones) are sufficiently great. As to firm organization, Milgrom and Roberts (1990) comment that:

the crucial distinguishing characteristic of a firm is not the pattern of asset ownership but the substitution of centralized authority for the relatively unfettered negotiations that characterize market transactions.... [W]hat most distinguishes any centralized organization is the authority and autonomy of its top decision makers or management—that is, their broad rights to intervene in lower-level decisions and the relative immunity of their decisions from intervention by others (pp. 72, 79).

In focusing on reputation effects, Kreps (1990) adopts a similar perspective:

I develop... a dichotomy in transactions that correlates well with the distinction between firms and markets.... The dichotomy is between hierarchical transactions and, for lack of a better name, specified transactions. Roughly, in a specified transaction all terms are spelled out in advance. In a hierarchical transaction, certain terms are left unspecified; what is specified is that one of the two parties has, within broad limits, the contractual right to specify how the contract will be fulfilled.... Here I will concentrate on arrangements where one party has the authority to determine ex post fulfillment, comparing this with cases where there is no need for such authority. (This notion is far from original to me; see, for example, Simon 1951) (p. 99).

21 Differences in the knowledge possessed by different individuals is implicit in the concept of asset specificity, which figures prominently in opportunism-based, transaction-cost theory (Williamson 1985a). It also is involved in the condition of information impactedness (Williamson 1975). More broadly, these differences motivate individuals to specialize in various aspects of business activity, including management.

22 Williamson (1985a) states that

By opportunism I mean self-interest seeking with guile. This includes but is scarcely limited to more blatant forms, such as lying, stealing, and cheating. Opportunism more often involves subtle forms of deceit.... More generally, opportunism refers to the incomplete or distorted disclosure of information, especially to calculated efforts to mislead, distort, disguise, obfuscate, or otherwise confuse (p. 47).

In our analysis, Y and Z engage in neither blatant nor subtle forms of opportunism. Thus, using Williamson's terminology, we assume "open or simple self-interest seeking" in which "individuals...are...fully open and honest in their efforts to realize individual advantage...." Williamson refers to this as a "semistrong form of self-interest seeking" (with opportunism being the strong form) (p. 49). As a related point, implications of trustworthy behavior for strategy and competitive advantage are addressed by, among others, Barney 1990; Barney and Hansen 1994; Donaldson 1990a, 1990b; Ghoshal and Moran 1996; Hill 1990; Hosmer 1994, 1995; Mahoney et al. 1994; and Maitland et al. 1985.

23 The result for our analysis is the same, whether we formulate the issue as one of irreducible differences in knowledge because of bounded rationality, or that, faced with the same information, individuals' hunches will differ about such future events as rivals' reactions or election outcomes, or about how those future events will impact the business.

24 An interesting question arises as to whether one party could simply pay the other to go along with the first when disagreements arise, therefore eliminating the negotiations and frictions that otherwise would be associated with market contracting. We can imagine, especially in the case of a one-time decision (e.g., where to place an oil well), that such a payment could be made, potentially resulting in market contracting being the preferred alternative. If, however, the work involves a series of future decisions, then one party would have to pay the other to agree to let the first have decision-making authority in the future. This amounts to an employment contract, and thus to firm organization as we have discussed it earlier.

25 Williamson (1985a) comments that

Plainly, if it were not for opportunism, all behavior could be rule governed. This need not, moreover, require comprehensive preplanning. Unanticipated events could be dealt with by general rules, whereby the parties agree to be bound by actions of a joint profit-maximizing kind. Thus problems during contract execution could be avoided by ex ante insistence upon a general clause of the following kind: I agree candidly to disclose all relevant information and thereafter to propose and cooperate in joint profit-maximizing courses of action during the contract execution interval, the benefits of which gains will be divided without dispute according to the sharing ratio herein provided (p. 48).

26 Interestingly, Williamson (1991a) recognizes the problem as it pertains to the incomplete-contracting perspective. He states that "common knowledge is a very expansive assumption.... Is common knowledge reasonable even between the immediate parties?" (p. 174).

27 This statement is not the same as assuming that Y's or Z's compensation in a firm or market contract is an express function of the profit of the enterprise. Our point simply is that, the more profitable is the enterprise, the more Y or Z is likely to receive, whether as salary or profit-sharing. And even a fixed-compensation individual cares that the enterprise takes in enough to be able to pay the fixed compensation.

28 Y's knowledge endowment is presumed to be the same, whether it picks firm organization or market contracting. A parallel assumption pertains to Z.

29 We intend "direction" by Z to be interpreted broadly. For example, such directions can include giving detailed instructions on how to complete a task, assigning workgroup membership, deciding or agreeing to certain delegations of decision-making authority, setting the premises for decision by others, and taking actions that affect the norms or culture of the organization. In addition, by "contractual partner," we merely mean a party to a market contract.
There is no presumption that all of an employee’s behavior under firm organization is direct. As McNulty (1984, p. 244) states, “the employment contract is usually one in which the employee has a certain discretion and freedom to choose how he will use his time on the job.”

In a related vein, Fredrickson (1986, p. 294) suggests that “[o]rganizations that differ in their dominant structure are likely to make strategic decisions using a very different process.”

Believing Z’s knowledge to be valuable is required for Y to want to engage in joint production with Z, under either organizational mode. A useful area for future research concerns how Y evaluates the overall quality of Z’s “package” of knowledge without knowing and understanding, in advance, all of the steps or elements that this knowledge incorporates. One possibility is that Y uses Z’s reputation—past history—to help assess the likely quality of Z’s knowledge (see, e.g., Weigelt and Camerer 1988).

Teece et al. (1994, p. 15) make the related point that “[w]hile individual skills are of relevance, their value depends upon their employment in particular organizational settings.”

We are indebted to Gordon Walker for pointing out that Cheung (1983) contains perhaps the first indication that a knowledge-substitution-like effect may exist, although Cheung expressly rejects it as a reason for the firm. Cheung asks:

Could it be that letting someone else make the decision [as to what to do] is often more productive? The answer is no. . . . Errors are bound to be less frequent when price information guides every activity performed (p. 5).

Cheung holds that a manager’s judgment cannot be superior to an employee’s, since the best guide to the employee’s behavior would be for the employee to observe price signals itself and adjust its actions accordingly. Cheung’s argument, however, presumes that the web of market contracts is so dense, complete and reproducible that (1) price signals exist for all relevant alternative courses of action, (2) price signals are easy to decipher and separate into their components, and (3) an actor easily can find a contracting partner that will offer to replicate with the actor some other parties’ observed market contract (and thus price). We believe that in many situations, the web is not very dense, complete or reproducible, so price signals will not be an adequate guide. Indeed, sustainable competitive advantage may depend on the absence of such price signals (see, e.g., Lippman and Rumelt 1982, Barney 1986). Because the manager and the employee have knowledge differences that are irreducible, the manager can be better (or worse) at interpreting the implications of market conditions than is the employee. Thus letting someone else decide what to do indeed can be the more productive alternative.

The manager itself may be uncertain how it creates the innovations. This is an example of the causal ambiguity explored by Lippman and Rumelt (1982).

The costs discussed in the previous subsection concern implementation of the original market contract.

Of course, there is a point at which, for both parties, the cost of backing down on desired points exceeds the value of the arrangement; if this point is reached, the relationship will be terminated once the contract expires.

The flexibility effect is related to one of Coase’s (1937) originally-identified transaction costs: “For this series of [market contracts] is substituted one” (p. 336). Coase’s argument, like ours, does not depend on opportunism.

In taking issue with a knowledge-based approach, Foss (1996, p. 474) holds that “[w]e cannot do without concepts such as opportunism if we wish to explain the existence of the firm.” We believe that a contrary conclusion is indicated by the foregoing analysis, which relies on opportunism-independent, knowledge-based factors to explain why firms exist. Foss (1996, p. 473) also maintains that a knowledge-based approach, such as in Conner (1991), “commit[s] the fallacy of technological determinism.” To avoid the mistake of technological determinism, Williamson provides the following test.

As illustrated by Figure 4 above, our model holds “technology” constant across organizational modes. Knowledge endowments are identical under both firm organization and market contracting, since the same actors participate in each. The only asymmetry in the model is that under firm organization, Z exercises authority as to Y, whereas under market contracting, Z does not. Since, using Williamson’s test, the initial knowledge of the actors—and hence the technology open to them—is constant across modes, it is difficult to see how our approach is an example of “technological determinism” in his sense. Firm organization may lead to the better use of knowledge, or more valuable learning through greater flexibility and coherence of strategic vision over time. In other words, firms may have advantages over markets. But this is no more “technological determinism” than is opportunism-based theory’s view that more valuable asset-specific investments will be made in firms because of this mode’s reduction of opportunistic potential.

We use “more valuable, opportunism-independent knowledge” in this section in a broad sense, encompassing both the knowledge-substitution and flexibility effects discussed earlier.

A more formal treatment would examine the degree of positive or negative correlation, or lack thereof, between four sets of events: (a) high and (b) low probability of opportunistic behavior, and (c) affirmative and (d) negative answers to whether the firm provides, on net, the more valuable knowledge.

The linkage of opportunism-based theory to productivity is established through the concept of asset specificity. If firm organization lessens opportunistic potential, then cooperating individuals optimally invest in more productive molding to one another. This generates a gain in their joint productivity as compared to market contracting. (For discussion of the optimal degree of asset specificity under an opportunism-based approach, see, e.g., Riordan and Williamson 1985.)
Demsetz (1988) includes in “firms,” both multiperson and single-person enterprises—the two organizational modes considered in this paper.

Based on the concept of selective intervention, Williamson eschews his earlier approach.

I...[previously addressed] the firm size dilemma by invoking bounded rationality and noting that limited spans of control are thereby implied. If any one manager can deal directly with only a limited number of subordinates, then increasing firm size necessarily entails adding hierarchical levels. Transmitting information across these levels experiences...losses..., which are cumulative and arguably exponential in form. As firm size increases and successive levels of organization are added, therefore, the effects of control loss eventually exceed the gains. A limit upon radical expansion is thus reached in this way... Plausible as the argument seemed at the time, it does not permit selective intervention of the kind described above. Rather, [in my previous conception] the entire firm is managed from the top (Williamson 1985*, pp. 134–5).

Milgrom and Roberts (1990) also comment that

[m]any of the arguments purporting to explain the limits of organization fail when confronted with the policy of replacing previously autonomous units with semiautonomous ones in whose operations and decisions central managers intervene only when uncoordinated or competitively oriented decisions are inefficient. Any adequate explanation of why all economic activity is not brought under central management must confront this possibility (p. 70).

Coase (1937, pp. 340–341) originally wrote that

[f]irst, as a firm gets larger, there may be decreasing returns to the entrepreneur function, that is, the costs of organizing additional transactions within the firm may rise... Naturally, a point must be reached where the costs of organizing an extra transaction within the firm are equal to the costs involved in carrying out the transaction in the open market, or, to the costs of organizing by another entrepreneur. Secondly, it may be that as the transactions which are organized increase, the entrepreneur fails to place the factors of production in the uses where their value is greatest, that is, fails to make the best use of the factors of production. Again, a point must be reached where the loss through the waste of resources is equal to the marketing costs of the exchange transaction in the open market or to the loss if the transaction was organized by another entrepreneur. Finally, the supply price of one or more of the factors of production may rise, because the “other advantages” of a small firm are greater than those of a large firm... Of course, the actual point where the expansion of the firm ceases might be determined by a combination of the factors mentioned above. The first two reasons given most probably correspond to the economists’ phase of “diminishing returns to management.”

As Masten (1988, p. 181) comments,

[fifty years after the publication of Ronald Coase’s seminal deliberations on the subject, economists have yet to reach a consensus on the nature of the firm. While many continue to regard the firm as a distinct institution, usually ascribing to it some superior control, information, or adaptive properties, others reject the notion that any unique governance advantages accrue to integration, noting that neither human nature nor technology or information are altered by the purely nominal act of “internalization.” For the latter, the word firm is merely descriptive, a collective noun denoting a particular cluster of otherwise ordinary contractual relationships.

For example, according to Jensen and Meckling (1976),

[contractual relations are the essence of the firm, not only with employees but with suppliers, customers, creditors, etc... It is important to recognize that most organizations are simply legal fictions which serve as a nexus for a set of contracting relationships among individuals... Viewed this way, it makes little or no sense to try to distinguish those things which are “inside” the firm (or any other organization) from those things that are “outside” of it. There is in a very real sense only a multitude of complex relationships (i.e., contracts) between the legal fiction (the firm) and the owners of labor, material and capital inputs and the consumers of output (pp. 310–311, emphasis in original).

As Masten (1988) explains,

[s]uch criticisms have led many economists to deny the existence of administrative solutions to contractual failures, asserting that the same transactional frictions confront employers as independent contractors. The firm, at least in governance respects, thus becomes no more than a coalition or “nexus” of contractual relationships, and the choice faced by transactors is only among the details to include in the contract (p. 183).

Dow also criticizes the nexus-of-contracts approach, as follows.

A venerable tradition in the economics of organization holds authority relations to be the distinguishing features of the firm (Coase 1937) and indeed of organization quite generally... Authority relations form the bedrock of internal organization in a way which ownership or the degree of performance monitoring do not. For example, employment relationships are widely agreed to fall within the ambit of internal organization, although ownership of human and physical capital remains separated. Intensive monitoring often appears in conjunction with internal organization (Alchian and Demsetz 1972), but does not imply it, since such monitoring can equally well be used to enforce promises arrived at through bargaining in a market setting. Williamson is therefore correct to use “degree of autonomy” as the fundamental dimension along which governance structures are arrayed; another way of expressing this is to say that as one moves
toward the ‘hierarchy’ end of the contractual spectrum, authorship relations grow in scope and complexity (1987, p. 16).

References


KATHLEEN R. CONNER AND C. K. PRAHALAD  A Resource-based Theory of the Firm


(1957), Administrative Behavior, New York: Macmillan Co.


Accepted by Jay B. Barney.


