

"NOTICE: This **material** may be protected
by **copyright** law (**Title 17 U.S. Code**)."

The ECONOMIC INSTITUTIONS of CAPITALISM

Firms, Markets, Relational Contracting

OLIVER E. WILLIAMSON
Yale University



THE FREE PRESS
A Division of Macmillan, Inc.
NEW YORK

Collier Macmillan Publishers
LONDON

UNIVERSITY OF CONNECTICUT
AT STAMFORD-LIBRARY
SCOFIELDTOWN ROAD
STAMFORD, CONN. 06903

Transaction Cost Economics

Firms, markets, and relational contracting are important economic institutions. They are also the evolutionary product of a fascinating series of **organizational innovations**. The study of the economic institutions of capitalism has not, however, occupied a position of importance on the social science research agenda.

Partly this neglect is explained by the inherent complexity of those institutions. But **complexity** can and often does serve as an inducement rather than a deterrent. The primitive state of our knowledge is at least equally explained by a reluctance to admit that the details of organization matter. The widespread conception of the modern corporation as a “black box” is the epitome of the noninstitutional (or pre-microanalytic) research tradition.

Merely to acknowledge that the microanalytic details of organization matter does not, however, suffice. The salient structural features of market, hierarchic& and quasi-market forms of organization need to be identified and **linked** to economic consequences in a systematic way. Lack of agreement on (or misconceptions regarding) the main purposes served by economic organization has also been an impediment to research progress.

A chapter in some yet unwritten history of economic thought will be needed to sort those matters out. Whatever the eventual explanation, the fact is that the study of economic institutions has witnessed a renaissance. Thus,

whereas the study of institutional economics reached a nadir in the immediate postwar period, a renewal of interest in institutions and a reaffirmation of their economic importance can, with the benefit of hindsight, be traced to the early **1960s**.¹ Operational content began to appear in the early **1970s**.² A common characteristic of the new line of research is that the concept of firm as production function is supplanted (or augmented) by the concept of firm as governance **structure**. Research of the New Institutional Economics kind had reached a critical mass by **1975**.³ The ensuing decade has witnessed exponential growth.

Transaction cost economics is part of the New Institutional Economics research tradition. Although transaction cost economics (and, more generally, the New Institutional Economics) applies to the study of economic organization of all kinds, this book focuses primarily on the economic institutions of capitalism, with special reference to firms, markets, and relational contracting. That focus runs the gamut from discrete market exchange at the one extreme to centralized hierarchical organization at the other, with myriad mixed or intermediate modes filling the range in between. The changing character of economic organization over time--within and between markets and hierarchies--is of particular interest.

Although the remarkable properties of neoclassical markets, where prices serve as sufficient statistics, are widely conceded--as Friedrich Hayek put it, the market is a "marvel" (1945, p. 525)—**opinions** differ in assessing transactions that are organized within quasi-market and nonmarket modes of organization. At best the administrative apparatus and private ordering sup-

¹The early contributions include Ronald Coase's **reconceptualization** of social costs (1960), **Armen** Alchian's pioneering **treatment** of property rights (1961). Kenneth **Arrow's** work on the troublesome economic properties of information (1962, 1963), and Alfred Chandler, Jr.'s contribution to business history (1962).

²**These** include **my** first efforts to recast the vertical integration problem in transaction cost terms (Williamson, 1971) and efforts to generalize that approach in the context of markets and hierarchies (Williamson, 1973); the treatments by **Armen** Alchian and Harold Demsetz of the "classical capitalist firm" in terms of team organization (1972) and their related work on property rights (1973); the proposed reformulation of economic history by Lance Davis and Douglass North (1971); the important work by Peter Doeringer and Michael **Piore** (1971) on labor markets; and Janos Komai's provocative treatment of disequilibrium economics (1971).

³**Some** of this is described in the first chapter of *Markets and Hierarchies* (1975), which is titled "Toward a New Institutional Economics." **The** conference on "The Economics of Internal Organization" held at the University of Pennsylvania in 1974 (the papers from which were published in 1975 and 1976 in the *Bell* journal of Economics) helped to redefine the research agenda. Many of the articles in the *Journal of Economic Behavior and Organization*, which first began publication in 1980, are in the New **Institutionalist** spirit. For recent commentary and contributions to this literature, see the March 1984 issue of the *Journal of Institutional and Theoretical Economics* and the forthcoming book of readings edited by Louis **Putterman** and Victor Goldberg.

ports that attend these transactions are messy. Some scholars decline even to deal with them. Others regard the deviations as evidence of a pervasive condition of “market failure.” Until very recently the primary economic explanation for nonstandard or unfamiliar business practices was monopoly:⁴ “[I]f an economist finds something—a business practice of one sort or another—that he does not understand, he looks for a monopoly explanation” (Coase, 1972, p. 67). That other social scientists should regard these same institutions as antisocial is unsurprising. The enforcement of antitrust from 1945 through 1970 reflected that orientation.

To be sure, a net negative social assessment is sometimes warranted. A more subtle and discriminating understanding of the economic institutions of capitalism has nevertheless been evolving. Many puzzling or anomalous practices have been cast into different relief in the process. This book advances the proposition that the economic institutions of capitalism have the main purpose and effect of economizing on transaction costs.

Main purpose is not, however, to be confused with sole purpose. Complex institutions commonly serve a variety of objectives. This is no less true here. The inordinate weight that I assign to transaction cost economizing is a device by which to redress a condition of previous neglect and undervaluation. An accurate assessment of the economic institutions of capitalism cannot, in my judgment, be reached if the central importance of transaction cost economizing is denied.⁵ Greater respect for organizational (as against technological) features and for efficiency (as against monopoly) purposes is needed. This theme is repeated, with variation, throughout this book.

I submit that the full range of organizational innovations that mark the development of the economic institutions of capitalism over the past 150 years warrant reassessment in transaction cost terms. The proposed approach adopts a contracting orientation and maintains that any issue that can be formulated as a contracting problem can be investigated to advantage in transaction cost economizing terms. Every exchange relation qualifies. Many other issues which at the outset appear to lack a contracting aspect turn out, upon scrutiny, to have an implicit contracting quality. (The cartel problem is an example.) The upshot is that the actual and potential scope of transaction cost economics is very broad.

As compared with other approaches to the study of economic organiza-

⁴Important exceptions to this tradition—which, however, were widely ignored—are Lester Telser’s (1965) and Lee Preston’s (1965) treatments of restrictive trade practices.

⁵A balanced view of the economic institutions of capitalism will await more concerted attention to the sociology of economic organization, which, happily, is in progress. For recent work of this kind, see Harrison White (1981), Martha Feldman and James March (1981), Arthur Stinchcombe (1983), Mark Granovetter (forthcoming), and James Coleman (1982).

tion, transaction cost economics (1) is more microanalytic, (2) is more self-conscious about its behavioral assumptions, (3) introduces and develops the economic importance of asset specificity, (4) relies more on comparative institutional analysis, (5) regards the business firm as a governance structure rather than a production function, and (6) place greater weight on the *ex post* institutions of contract, with special emphasis on private ordering (as compared with court ordering). A large number of additional implications arise upon addressing problems of economic organization in this way. The study of the economic institutions of capitalism, as herein proposed, maintains that the transaction is the basic unit of analysis and insists that organization form matters. The underlying viewpoint that informs the comparative study of issues of **economic** organization is this: Transaction costs are economized by assigning transactions (which differ in their attributes) to governance structures (the adaptive capacities and associated costs of which differ) in a discriminating way.⁶

Given the complexity of the phenomena under review, transaction cost economics should often be used in addition to, rather than to the exclusion of, alternative approaches. Not every approach is equally instructive, however, and they are sometimes **rival** rather than **complementary**.

The nature of transaction costs is developed in section 1. A cognitive map of contract, in which alternative approaches to economic organization are described and with respect to which transaction cost economics is located, is set out in section 2. The relation between behavioral assumptions and **alter-**native conceptions of contract is presented in section 3. A rudimentary contracting schema on which the argument in the book repeatedly relies is developed in section 4. Contractual issues that arise in organizing the company town are examined in section 5. Other applications are sketched in section 6. Concluding remarks follow.

1. Transaction Costs

1.1 *Frictionlessness*

Kenneth Arrow has defined transaction costs as the “costs of running the economic system” (1969, p. 48). Such costs are to be distinguished from production costs, which is the cost category with which neoclassical analysis

⁶Indeed, transaction cost economizing is central to the study of economic organization quite generally-in capitalist and noncapitalist economies alike.

has been preoccupied. Transaction costs are the economic equivalent of friction in physical systems. The manifold successes of physics in ascertaining the attributes of complex systems by assuming the absence of friction scarcely require recounting here. Such a strategy has had obvious appeal to the social sciences. **Unsurprisingly**, the absence of friction in physical systems *is* cited to illustrate the analytic power associated with “unrealistic” assumptions (Friedman, 1953, pp. 16-19).

But whereas physicists were quickly reminded by their laboratory instruments and the world around them that friction was pervasive and often needed to be taken expressly into account, economists did not have a corresponding appreciation for the costs of running the economic system. There is, for example, no reference whatsoever to transaction costs, much less to transaction costs as the economic counterpart of friction, in Milton Friedman’s famous methodological essay (1953) or *in* other postwar treatments of positive economics.⁷ Thus although positive economics admitted that **frictions** were important in principle, it had no language to describe frictions in **fact**.⁸

The neglect of transaction costs had numerous ramifications, not the least of which was the way in which nonstandard modes of economic organization were interpreted. Until express provision for transaction costs was made, the possibility that nonstandard modes of organization—customer and territorial restrictions, tie-ins, block booking, franchising, vertical integration, **and the like**—operate in the service of transaction *cost economizing* was little **appreciated**. Instead, most economists invoked monopoly explanations—be it of the leverage, price discrimination, or entry barriers **kinds**—when confronted with nonstandard contracting practices (Coase, 1972, p. 67). Donald Turner’s views are representative: “I approach customer and territorial restrictions not hospitably in the common law tradition, but inhospitably in the tradition of **antitrust**.”⁹ As discussed below, the research agenda and public policy toward business were massively influenced by that monopoly predisposition. The prevailing view of the firm as production function was centrally implicated in that situation.

⁷**Herbert** Simon’s treatments of decision-making in economics focus mainly on individual rather than **institutional** features of economic organization (1959; **1962**).

⁸**To** be sure, the market failure literature was concerned with many of the relevant issues. But it rarely posed the issues in transaction cost terms. Arrow’s remarks are thus prescient: “I contend that market failure is a more general category than externality. . . . [Moreover], market failure is not absolute; it is better to consider a broader category, **that** of transaction costs, which **in** genera! impede and in particular cases completely block the formation of markets” (1969, p. 48).

⁹**The** quotation is attributed to Turner by Stanley Robinson, 1968, **N.Y.** State Bar Association, Antitrust Symposium, p. 29.

1.2 *Explication*

Transaction cost economics poses the problem of economic organization as a problem of contracting. A particular task is to be accomplished. It can be organized in any of several alternative ways. Explicit or implicit contract and support apparatus are associated with each. What are the costs?

Transaction costs of *ex ante* and *ex post* types are usefully distinguished. The first are the costs of drafting, negotiating, and safeguarding an agreement. This can be done with a great deal of care, in which case a complex document is drafted in which numerous contingencies are recognized, and appropriate adaptations by the parties are stipulated and agreed to in advance. Or the document can be very incomplete, the gaps to be **filled** in by the parties as the contingencies arise. Rather, therefore, than contemplate all conceivable **bridge** crossings in advance, which is a very ambitious undertaking, only **actual** bridge-crossing choices are addressed as events unfold.

Safeguards can take several forms, the most obvious of which is common ownership. Faced with the prospect that autonomous traders will experience contracting difficulties, the parties may substitute internal organization for the market. This is not, to be sure, without problems of its own (see Chapter 6). Moreover, *ex ante* interfirm safeguards can sometimes be fashioned to signal credible commitments and restore integrity to transactions. The study of “nonstandard” contracting is centrally concerned with such matters.

Most studies of exchange **assume** that efficacious rules of law regarding contract disputes are in place and are applied by the courts in an informed, sophisticated, and low-cost way. Those assumptions are convenient, in that lawyers **and** economists are relieved of the need to examine the variety of ways by which individual parties to an exchange “contract **out** of or away from” the governance structures of the state by devising private orderings. Thus arises a division of effort whereby economists are preoccupied with the economic benefits that accrue to specialization **and** exchange, while legal specialists focus on the technicalities of contract law.

The “legal centralism” tradition reflects the latter orientation. It maintains that “disputes require ‘access’ to a forum external to the original social setting of the dispute [and that] remedies will be provided as prescribed in some body of authoritative learning and dispensed by experts who operate under the auspices of the state” (Galanter, 1981, p. 1). The facts, however, disclose otherwise. Most disputes, including many that under current rules could be brought to a court, are resolved by avoidance, self-help, and the like (Galanter, 1981, p. 2).

The **unreality** of the assumptions of legal centralism can be defended by **reference** to the fruitfulness of the pure exchange model. That is not **disputed** here. My concern is that the law and economics of private ordering have been pushed into the background as a consequence. That is unfortunate, since in “many instances the participants can devise more satisfactory solutions to their disputes than can professionals constrained to apply general rules on the basis of **limited** knowledge of the dispute” (Galanter, 1981, p. 4).¹⁰

The issues here are akin to those that were of **concern** to Karl Llewellyn in his discussion of contract in 1931 but have been systematically evaded since.¹¹ But for the limitations of legal **centralism**, the *ex post* side of contract **can** be disregarded. Given the very real limitations, however, with which court ordering is beset, the *ex post* costs of contract unavoidably intrude. Transaction cost economics insists that contracting **costs** of all kinds be accorded **parity**.

Ex post costs of contracting take several forms. These include (1) the maladaptation costs incurred when transactions drift out of alignment in relation to what Masahiko Aoki refers to as the “shifting contract curve” (1983),¹² (2) the **haggling** costs incurred if bilateral efforts are made to correct *ex post* misalignments, (3) the setup and running costs associated with the governance structures (often not the courts) to which disputes are referred, and (4) the bonding costs of effecting secure commitments.

Thus suppose that the contract stipulates x but, **with the** benefit of **hindsight** (or **in the fullness of knowledge**), the parties discern that they should have done y. Getting from x to y, however, may not be easy. The manner in which the associated benefits are divided is apt to give rise to intensive, self-interested bargaining. Complex, strategic behavior may be elicited. Referring the dispute to another forum may help, but that will vary with the circumstances. An incomplete adaptation will be realized if, as a consequence of efforts of both kinds, **the parties move not to y but to y'**.

A complicating factor in all of this is that the *ex ante* and *ex post* costs of contract are interdependent. **Put** differently, they must be addressed simultaneously rather than sequentially. Also, costs of both types are often difficult

¹⁰Marc Galanter elaborates as follows: “The variability of preferences and of situations, compared to the small number of things that can be taken **into** account by formal rules . and the loss of meaning in **transforming** the dispute into professional categories suggest limits on the desirability of conforming outcomes to authoritative **rules**” (1981, p. 4).

¹¹See “Prologue,” Section 1.2.

¹²The *ex post* transaction costs are related to, but plainly differ from, what Michael Jensen and William **Meckling** refer to as agency costs, which they define as the sum of “(1) the monitoring expenditures of the principal, (2) the bonding expenditures by the agent, and (3) the **residual** loss” (1976, p. 308)—**this** last being a very expansive category.

to quantify. The difficulty, however, is mitigated by the fact that transaction costs are always assessed in a comparative institutional way, in which one mode of contracting is compared with another. Accordingly, it is the difference between rather than the absolute magnitude of transaction costs that matters. As Herbert Simon has observed, the comparison of discrete structural alternatives can employ rather primitive apparatus- “such analyses can often be carried out without elaborate mathematical apparatus or marginal calculation. In general, much cruder and simpler arguments will suffice to demonstrate an inequality between two quantities than are required to show the conditions under which these quantities are equated at the margin” (1978, p. 6). Empirical research on transaction cost matters almost **never** attempts to measure such costs directly. Instead, the question is whether organizational relations (contracting practices; governance structures) line up with the attributes of transactions as predicted by transaction cost reasoning or not.

1.3 *The Larger Context*

--

This book concentrates on transaction cost economizing, but the costs need to be located in the larger context of which they are a part. Among the relevant factors-to which I sometimes (but not continuously) refer-are the following:

1. Holding the nature of the good or service to be delivered constant, economizing takes place with reference to the sum of production and transaction costs, whence tradeoffs in this respect must be recognized .
2. More generally, the design of the good or service to be delivered is a decision variable that influences demand as well as costs of both kinds, whence design is appropriately made a part of the calculus.
3. The social context in which transactions are embedded-the customs, mores, habits, and so on-have a bearing, and therefore need to be taken into account, when moving from one culture to another. ^{1 3}
4. The argument relies in a general, background way on the efficacy of competition to perform a sort between more and less efficient modes and’ to shift resources. in favor of the former. This seems plausible, especially if the relevant outcomes are those which appear over **inter-**

¹³See Mark Ciraoovetter (I 983) for a discussion of the importance of embeddedness. Also see Douglass North (1981).

vals of five and ten years rather than in the very near **term**.¹⁴ This intuition would nevertheless benefit from a more fully developed theory of the selection process. Transaction cost arguments are thus open to some of the same objections **that** evolutionary economists have made of orthodoxy (Nelson and Winter, 1982, pp. 356–70), though in other respects there are strong complementarities (pp. 34–38).

5. Whenever private and social benefits and costs differ, the social cost calculus should govern if prescriptive treatments are attempted.

2. A Cognitive Map of Contract

The field of specialization with which transaction c&t economics is most closely associated is industrial organization. A number of the leading approaches to the study of industrial organization and the relation that transaction cost economics bears to them are examined here.

Industrial organization examines contract in terms of the purposes served. What are the parties trying to accomplish? Here as elsewhere in industrial organization, monopoly and efficiency purposes are usefully distinguished. The cognitive map shown in Figure 1–1 begins with this distinction.

2.1 *The Monopoly Branch*

All of the approaches to contract shown in Figure 1-1, monopoly and efficiency alike, are concerned with the same puzzle: What purposes are served by supplanting classical market exchange—whereby product is sold at a uniform price to all comers without restriction—by more complex forms of contracting (including nonmarket modes of economic organization)? The monopoly approaches ascribe departures from the classical norm to monopoly purpose. The efficiency approaches hold that the departures serve economizing purposes instead.

“This intuition is akin to **that** expressed by Michael Spence in his conjecture that entry barrier arguments give way to contestable markets in the long run (1983, p. 988). Although the long run for Spence probably exceeds five or ten years, some of the evolutionary phenomena of **interest** to me also span half a century. One way of putting it is that I subscribe to weak-form rather than strong-form selection, the distinction being that “in a relative sense, **the fitter** survive, but there is no reason to suppose that they **are fittest in** any absolute sense” (Simon, 1983, p. 69; emphasis in original).

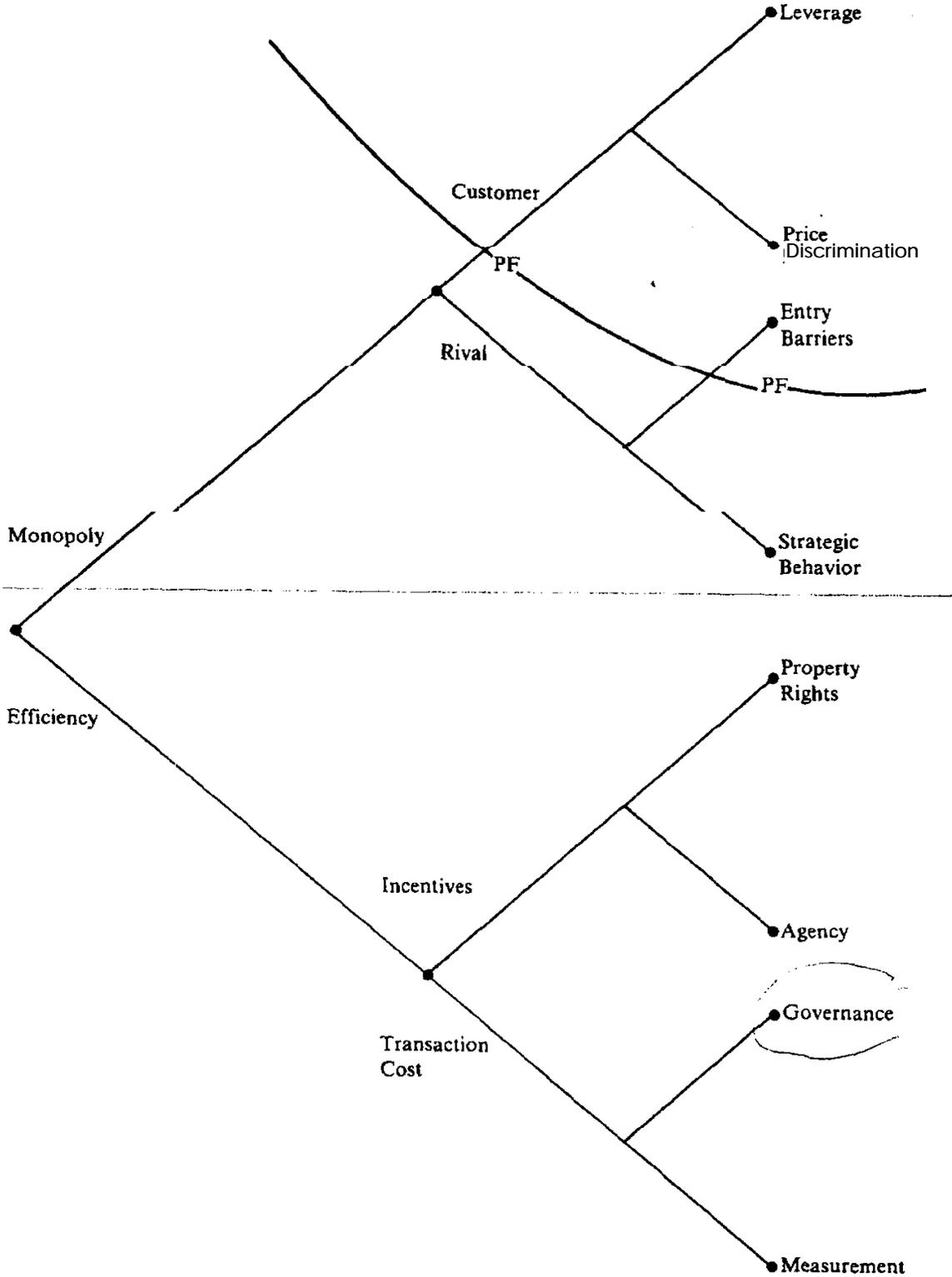


FIGURE 1- 1. Cognitive Map of Contract

The four monopoly approaches to contract are grouped under two headings. The first examines the uses of customer and territorial restrictions, resale price maintenance, exclusive dealing, vertical integration, and the like in relation to buyers. The second is concerned with the impact of such practices on rivals.

The “**leverage**”¹⁵ theory of contract and the price discrimination interpretation of nonstandard contracting both focus on buyers. Richard Posner (1979) associates leverage theory with the (earlier) “Harvard School” approach and price discrimination with the “Chicago School” approach to antitrust economics. Leverage theory maintains that original monopoly power can be extended and that nonstandard contracting practices accomplish this. Although leverage theory is largely discredited among economists,¹⁵ it maintains an appeal to many lawyers and continues to find its way into legal briefs¹⁶ and court opinions.¹⁷

The price discrimination approach to nonstandard contracting maintains that original monopoly power is unchanged. Price discrimination is merely a means by which latent monopoly power is actualized. This interpretation of nonstandard contracting has **been** advanced by Aaron Director and Edward Levi (1956) in conjunction with tie-in sales and by George Stigler (1963) in relation to block booking. Tie-in sales and block booking are purportedly devices by which sellers are able to discover underlying product valuation differences among consumers and to monetize consumers’ surplus.

The other two **monopoly** approaches examine nonstandard contracting **practices** in relation to rivals. They are expressly concerned with the enlargement of monopoly power by large established firms in relation to smaller actual or potential rivals. The barriers to entry literature, which is prominently associated with the work of Joe Bain (1956), is in that tradition. The early work in the area has *come* under considerable criticism, much of it originating

¹⁵The “main” leverage argument applies to the sale of complementary goods in the downstream market. Chicago maintains that tying in **these** circumstances cannot extend monopoly power but merely represents an effort to effect price discrimination: “absent price discrimination, a monopolist **will** obtain **no** additional monopoly profits from monopolizing a complementary product” (Posner, 1976, p. 173). This assessment is widely conceded. Timothy Brennan and Sheldon **Kimmel** have since examined the special but interesting case where the tie occurs not in the downstream but in the upstream market. They show that a tie here can effect a monopoly in the second market if “economies of scope or demand conditions . . . [are] such that the producers cannot sell the second good profitably unless they **sell** the first good to the monopolist as well” (1983, p. 21).

¹⁶See, for example, the *amicus* brief prepared by Lawrence A. Sullivan in support of the **respondent** in *Monsanto Company v. Spray-Rite Service Corporation*.

¹⁷Although the majority opinion in *Jefferson Parish Hosp. Dist. No. 2 v. Hyde* (44CCH S. Ct. Bull., P.) reaches the correct result, it also muddies the opinion by passing reference to **leverage** theory.

with the Chicago School. The main problems with the early work are that it was static and did not carefully identify the essential preconditions for entry barrier arguments to go through. The more recent literature on strategic behavior relieves many of the objections.¹⁸ Investment and information asymmetries are expressly introduced. Intertemporal attributes are recognized; and reputation effect features are developed. The use of nonstandard contracting as a means of “raising rivals’ costs” (Salop and Scheffman, 1983) is an especially intriguing possibility.

The recent strategic behavior literature excepted, all the monopoly approaches to contract work within the neoclassical framework, where the firm is regarded as a production function. Inasmuch as *the natural boundaries of the firm are therein defined by technology, any effort by the firm to extend its reach* by recourse to nonstandard contracting was presumed to have monopoly purpose and effect.¹⁹ This “applied price theory” approach to industrial organization was the prevailing postwar orientation. As Coase observed (1972, p. 61), it informed both of the leading industrial organization texts—the one by Joe Bain (1958); the other by George Stigler (1968). The inhospitality approach to antitrust law enforcement, to which I referred in 1.1, was similarly oriented. Much of the strategic behavior literature, by contrast, is more closely associated with the governance structure conception of the enterprise (see Chapter 14). So as to highlight this important monopoly distinction, the dashed curve (denoted PF) in Figure 1-1 separates the earlier production function approaches from the more recent strategic conception of contract.

2.2 *The Efficiency Branch*

Most of what I refer to as the New Institutional Economics is located on the efficiency branch of contract. The efficiency branch of contract distinguishes between those approaches in which incentive alignments are emphasized and those which feature economies of transaction costs. The incentive alignment literature focuses on the *ex ante* side of contract. New forms of property rights and complex contracting are thus interpreted as efforts to overcome the incentive deficiencies of simpler property rights and contracting traditions. Ronald Coase (1960), Armen Alchian (1961; 1965), and Harold Demsetz (1967;

*See Chapter 14.

¹⁹To be sure, it can be argued that price discrimination is efficient, which it ordinarily is if it can be effected at zero transaction cost and if income distribution effects wash out. The zero transaction cost assumption is rarely warranted, however. Private and social valuations of price discrimination can yield contradictory results for this reason (Williamson, 1975, pp. 11-13).

1969) are prominently associated with the property rights literature.²⁰ Leonid Hurwicz (1972; 1973), Michael Spence and Richard Zeckhauser (1971), Stephen Ross (1973), Michael Jensen and William Meckling (1976), and James Mirrlees (1976) opened up the agency approach.*

The property rights literature emphasizes that ownership *matters*, where the rights of ownership of an asset take three parts: the right to use the asset, the right to appropriate returns from the asset, and the right to change the form and/or substance of an asset (Furubotn and Pejovich, 1974, p. 4). Upon getting the property rights straight, it is commonly assumed (often implicitly; sometimes explicitly) that asset utilization will thereafter track the purposes of its owners. This will obtain if (1) the legally sanctioned structure of property rights is respected and (2) human agents discharge their jobs in accordance with instructions.²²

Thus, whereas the monopoly branch of **contract** interprets nonstandard forms of exchange as having monopoly purpose and effect, the property rights literature would inquire whether mistaken property rights assignments were responsible for resource **misallocations**. Redescribing property rights, possibly in complex (nonstandard) ways, is what explains contractual irregularities. Put differently, discrete market contracting is supplanted by more complex forms of contracting, because that is the way residual rights to control can be placed in the hands of those who can use those rights most productively.

The agency literature, **particularly** the early agency literature, **emphasizes that** ‘principals **contract** in full awareness of the hazards that contract execution by agents poses. Thus although the separation of ownership from control attenuates profit incentives, that is anticipated at the time separation occurs and is fully reflected in the price of new shares (Jensen and Meckling, 1976). The future therefore holds no surprises; **all** of the relevant contracting action is packed into *ex ante* incentive alignments.

Actually, as Michael Jensen’s influential survey points out (1983), the agency literature has developed in two parts. He refers to the one branch as the positive theory of agency. Here, “capital intensity, degree of **specializa-**

²⁰For recent survey, see Louis De Alessi (1983). For an earlier survey, see Eirik Furubotn and Steve Pejovich (1974).

²¹For a recent survey, see Stanley Baiman (1982).

²²The recent treatment of vertical integration by Sanford Grossman and Oliver Hart illustrates both of these propositions. Thus they view asset ownership as control over residual rights: “Each **asset** will have a single owner and that owner has the right to control the asset in the case of a missing [contractual] provision” (1984, p. 7). They further contend that the owner of physical assets “**can** order plant employees” to utilize these assets in accordance with his directions (1984, p. 17). Differences between market organization and vertical integration are thus entirely attributed to the asset ownership differences that distinguish them.

tion of assets, information costs, capital markets, and internal and external labor markets are examples of factors in the contracting environment that interact with the costs of various monitoring and bonding practices to determine the contractual forms” (Jensen, 1983, pp. 334-35). The positive branch repeatedly asserts that natural selection processes are **reliably** efficacious (Fama, 1980; Jensen, 1983, p. 331; Fama and Jensen, 1983, pp. 301, 327)—**Armen** Alchian’s classic but highly **nuanced** and very cautious statement of the evolutionary approach to economics (1950) being cited as the main authority.

Jensen refers to the second type of agency literature as that of “principal-agent” (1983, p. 334). This relatively mathematical literature features *ex ante* incentive alignments in superlative degree. It has come to be known more recently as the mechanism design approach. This line of research is akin to the earlier contingent claims contracting **literature**²³ but moves beyond it by admitting contracting complications in the form of private information. Complex problems of incentive alignment are posed (which the contingent claims contracting literature had ignored) if full and candid disclosure of private information cannot be assumed; In--other respects; **however,** the mechanism--design and contingent claims contracting literatures are very similar: Both resolve all the relevant contracting issues in a comprehensive *ex ante* **bargain**;²⁴ and both assume that court ordering is **efficacious**.²⁵ Again, efficiency rather than monopoly purposes drive the argument.

The transaction cost literature also maintains the rebuttable presumption that nonstandard forms of contracting have efficiency purposes. Greater attention is shifted, however, to the contract execution stage. As shown in Figure

²³**Mervyn** King characterizes the Arrow-Debreu model as **follows**:

. . . commodities are distinguished not only by physical and spatial characteristics, and by the date at which the commodity is made available, but also by the “state of the world” in which it is delivered. A “state of the world” is defined by assigning values to all the uncertain variables which are relevant to the economy and comprises a complete list of **all** these variables. These states of the world are mutually exclusive, and together form an exhaustive **set**. . . Commodities are now defined as contingent on the occurrence of **certain** events, and the market system comprises markets in all these contingent commodities. [1977, p. 128]

²⁴**The** mechanism design literature assumes that the parties to a contract have the cognitive competence to craft contracts of **unrestricted** complexity. In the language of Chapter 2, the parties **to** a contract have unbounded rationality; see Bengt Holmstrom (1984). By contrast with the **property** rights literature, the mechanism design approach holds that “since each party’s obligation to the other is completely specified for every state of nature, there are no residual rights of control over assets to be allocated” (Grossman and Hart, 1984, p. 7). Complex contracts are therefore not **concerned** with residual rights but with getting the obligations defined at the outset--due provision for private information having been acknowledged.

²⁵**See** Baiman (1982, p. 168).

1-1, the transaction cost approach is split into a governance branch and a measurement branch. Of the two, this book places greater emphasis on the former. Both, however, are important and in fact are interdependent.

In common with the property rights literature, transaction cost economics agrees **that** ownership matters. It furthermore acknowledges that *ex ante* incentive alignments matter. But whereas the property rights and mechanism design approaches work within the tradition of legal centralism, transaction cost economics disputes that court ordering is efficacious. Attention is shifted instead to private ordering. What institutions are created with what adaptive, sequential decision-making and dispute settlement properties? To ownership and incentive alignment, therefore, transaction cost economics adds the proposition that the *ex post* support *institutions* of contract *matter*.

James Buchanan has argued that “economics ‘comes closer to being a ‘science of contract’ than a ‘science of choice’ [on which account] the maximizer must be replaced by the arbitrator, the outsider who tries to work out compromises among conflicting claims” (1975, p. 229). The governance approach adopts the science of contract orientation but joins the arbitrator with an institutional design specialist. The object is not merely to resolve conflict in progress but **also** to recognize potential conflict in advance and devise governance structures that forestall or attenuate it.

Transaction cost economics maintains that it is impossible to concentrate all of the relevant bargaining action at the *ex ante* contracting stage. **Instead, bargaining is pervasive**—on which account **the institutions** of private ordering and the study of contracting in its entirety take on critical economic significance. The behavioral attributes of human agents, whereby conditions of bounded rationality and opportunism are joined, and the complex attributes of transactions (with special reference to the condition of asset specificity) are responsible for that condition.

The measurement branch of transaction cost economics is concerned with performance or attribute ambiguities that are associated with the **supply** of a good or service. The Alchian-Demsetz (1972) treatment of technological nonseparabilities (team organization) is an example. The issues have since been addressed by William **Ouchi (1980b)** in the context of work organization and Yoram **Barzel (1982)** with respect to the organization of markets. A recent interesting application is the study by Roy Kenney and Benjamin Klein **(1983)** of what they refer to as “oversearching.” They take exception with **Stigler’s** view that block booking has monopoly (price discrimination) **pur-**poses and argue instead that it serves to economize on measurement costs.

AS indicated, this book deals **mainly** with the governance branch of transaction cost economics. Measurement aspects are also treated, **however**—as indeed they must be, as governance and measurement are interdependent.

3. The World of Contract

The world of contract is variously described as *one* of (1) planning, (2) promise, (3) **competition**, and (4) governance (or private ordering). Which of these descriptions is most applicable depends on the behavioral assumptions that pertain to an exchange and on the economic attributes of the good or service in question.

As developed more fully in Chapter 2, the study of economic organization turns critically on two behavioral assumptions. What cognitive **competencies** and what self-interest seeking propensities are ascribed to the human agents engaged in exchange? Transaction cost economics assumes that human agents are subject to bounded rationality, whence behavior *is* “**intendedly** rational, but only **limitedly** so” (Simon, 1961, p. xxiv), and are given to opportunism, which is a condition of self-interest seeking with guile. Transaction cost economics further maintains that the most critical dimension for describing transactions is the condition of asset specificity. Parties engaged in a trade that is supported by nontrivial investments in transaction-specific assets are effectively operating in a **bilateral** trading relation with one another. Harmonizing the contractual interface that joins the parties, thereby to effect adaptability and promote continuity, becomes the source of real economic value.

But for uncertainty, problems of economic organization are relatively uninteresting. Assume, therefore, that uncertainty is present in nontrivial degree and consider the ramifications for contract of differences in bounded rationality, opportunism, and asset specificity. Assume, in particular, that each of these conditions can take on either of two values: Either it is present in significant degree (denoted +) or it is presumed to be absent (denoted 0). Consider the three cases in which only one of these factors is presumed to be absent and then that in which all three are joined. Table 1-1 shows the four conditions to be compared and the contracting model that is associated with each.

The case where parties are opportunistic and assets are specific but economic agents have unrestricted cognitive competence essentially describes the mechanism design literature (Hurwicz, 1972; 1973; Meyerson, 1979; Harris and Townsend, 1981). Although the condition of opportunism requires that contracts be written in such a way as to respect private information, whence complex incentive alignment issues are posed, **all** the relevant issues of contract are settled at the **ex ante** bargaining stage. Given unbounded rationality, a comprehensive bargain is struck at the outset, according to which appropriate adaptations to subsequent (publicly observable) contingent

Table 1-1. Attributes of the Contracting Process

<i>Behavioral Assumption</i>			<i>Implied Contracting Process</i>
<i>Bounded Rationality</i>	<i>Opportunism</i>	<i>Asset Specificity</i>	
0	+	+	Planning
+	0	+	Promise
+	+	0	Competition
+	+	+	Governance

events are fully described. Contract execution problems thus never arise (or defection from such agreements is deterred **because court** adjudication of all disputes is assumed to be efficacious (Baiman, 1982, p. 168)). Contract, in the context of unbounded rationality, is therefore described as a world of planning.

Consider alternatively the situation where agents are subject to bounded rationality and transactions are supported by specific assets, but the condition of opportunism is assumed to be absent, which implies that the word of an agent is as good as his bond. Although gaps will appear in these contracts, because of bounded rationality, they do not pose execution hazards if the parties take recourse to a self-enforcing general clause. Each party to the contract **simply pledges at** the outset to execute the contract efficiently (in a joint profit maximizing manner) and to seek only fair returns at contract renewal intervals. Strategic behavior is thereby denied. Parties to a contract thus extract all such advantages as their endowments entitle them to when the **initial** bargain is struck. Thereafter contract execution goes efficiently **to** completion because promises of the above-described kind are, in the absence of opportunism, self-enforcing. Contract, in this context, reduces to a world of promise.

Consider, then, the situation where agents are subject to bounded rationality and are given to opportunism, but asset specificity is presumed to be absent. Parties to such contracts have no continuing interests in the identity of one another. This describes the world where discrete market contracting is efficacious, where markets are fully **contestable**,²⁶ and where franchise bid-

²⁶**Differences** between transaction cost economics and “contestability theory” (Baumol, Panzer, and Willig, 1982) in asset-specificity respects are noteworthy. Both approaches to the study of economic organization acknowledge the importance of asset specificity, but they view it from opposite ends of the telescope. Thus contestability theory reduces asset specificity to insignificance, so that hit-and-run entry is easy. Transaction cost economics, by contrast, magnifies the condition of asset specificity. The existence of durable, firm specific assets is held to be widespread, and accordingly hit-and-run entry is often infeasible.

ding for natural monopoly goes through. Inasmuch as fraud and egregious contract deceits are deterred by court **ordering**,²⁷ contract, in this context, is described by a world of competition.

Each of the **three** devices fails when bounded rationality, opportunism, and asset specificity are joined. Planning is necessarily incomplete (because of bounded rationality), promise predictably breaks down (because of opportunism), and the **pairwise** identity of the parties now matters (because of asset specificity). This is the world of governance. Since the efficacy of court ordering is problematic, contract execution falls heavily on the institutions of private ordering. This is the world with which transaction cost economics is concerned. The organizational imperative that emerges in such circumstances is this: *Organize transactions so as to economize on bounded rationality while simultaneously safeguarding them against the hazards of opportunism.* Such a statement supports a different and **larger** conception of the economic problem than does the imperative “Maximize profits!”

4 . A Simple Contracting S c h e m a -

Assume that a good or service can be supplied by either of two alternative technologies. One is a general purpose technology, the other a special purpose technology. The special purpose technology requires greater investment in transaction-specific durable assets and is more efficient for servicing **steady-state** demands.

Using k as a measure of transaction-specific assets, transactions that use the general purpose technology are ones for which $k = 0$. When transactions use the special purpose technology, by contrast, a $k > 0$ condition exists. Assets here are specialized to the particular needs of the parties. Productive values **would** therefore be sacrificed if transactions of this kind were to be prematurely terminated. The bilateral monopoly condition described above and elaborated in Chapter 2 applies to such transactions.

Whereas classical market contracting--“sharp in by clear agreement; sharp out by clear performance” (Macneil, 1974, p. 738)—suffices for transactions of the $k = 0$ kind, unassisted market governance poses hazards whenever nontrivial transaction-specific assets are placed at risk. Parties have an incentive to devise safeguards to protect investments in transactions of the latter kind. Let s denote the magnitude of any such safeguards. An $s = 0$ condition is one in which no safeguards are provided; a decision to provide safeguards is reflected by an $s > 0$ result.

²⁷The assumption that court ordering is efficacious in a regime of **bounded rationality and opportunism** is plainly gratuitous, but it is the maintained assumption nonetheless.

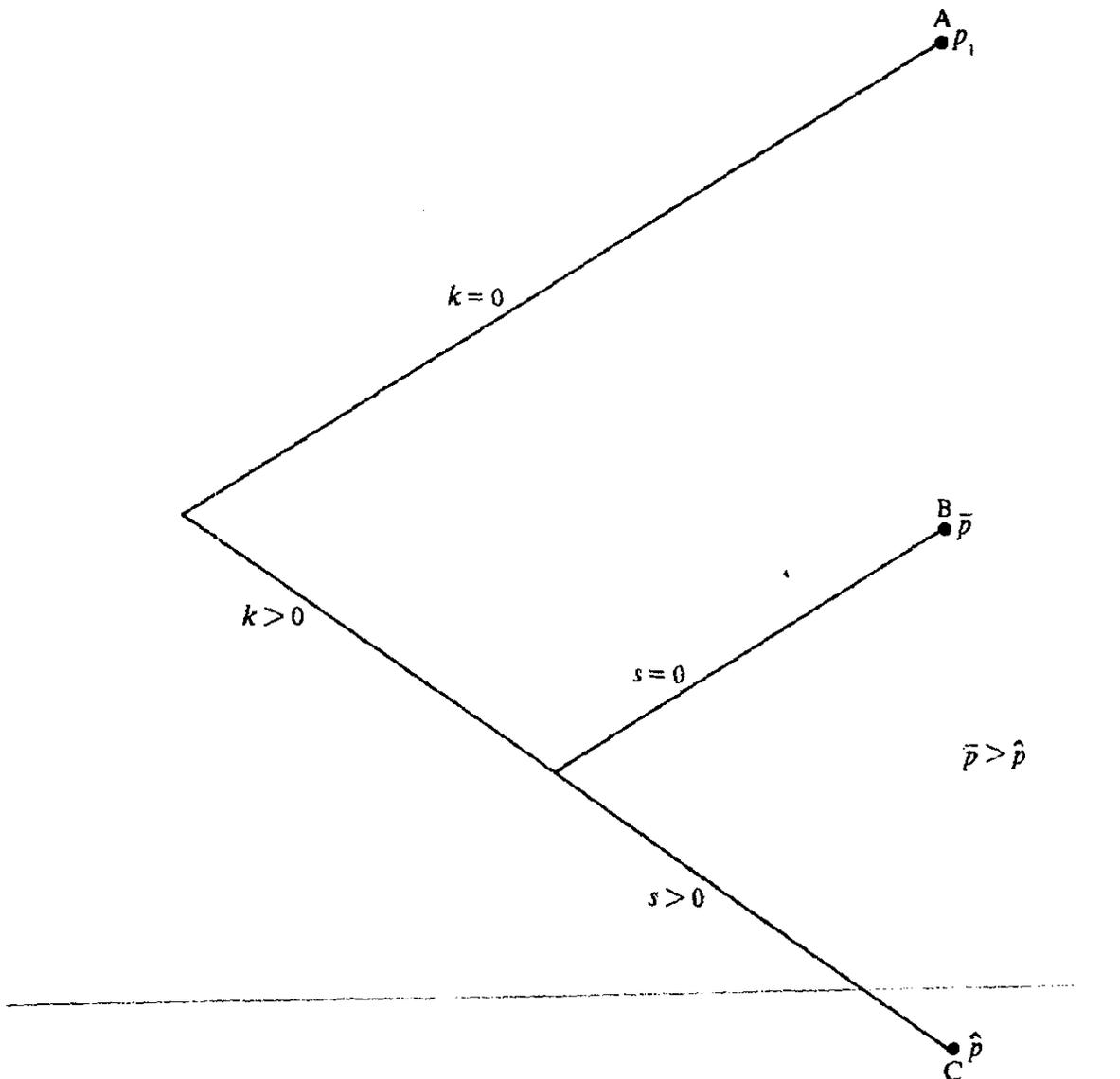


FIGURE 1-2. A Simple Contracting Schema

Figure 1-2 displays the three contracting outcomes corresponding to **such** a description. Associated with each node is a price. So as to facilitate comparison between nodes, assume that suppliers (1) are risk neutral, (2) are prepared to supply under either technology, and (3) will accept any safeguard condition whatsoever so long as an expected breakeven result can be projected. Thus node A is the general purpose technology ($k = 0$) supply relation for which a breakeven price of p_1 is projected. The node B contract is supported by transaction-specific assets ($k > 0$) for which no safeguard is offered ($s = 0$). The expected breakeven price here is \bar{p} . The node C contract also employs the special purpose technology. But since the buyer at this node provides the supplier with a safeguard, ($s > 0$), the breakeven price, \hat{p} , at node C is less than \bar{p} .

The protective safeguards to which I refer normally take on one or more of three forms. The **first** is to realign incentives, which commonly involves

some type of severance payment or penalty for premature termination. A second is to create and employ a specialized governance structure to which to refer and resolve disputes. The use of arbitration, rather than litigation in the courts, is thus characteristic of node C governance. A third is to introduce trading regularities that support and signal continuity intentions. Expanding a trading relation from unilateral to bilateral exchange-through the concerted use, for example, of reciprocity-thereby to effect an equilibration of trading hazards is an example of that last.

This simple contracting schema, which will subsequently be elaborated, applies to a wide variety of contracting issues. It facilitates comparative institutional analysis by emphasizing that technology (k), contractual governance/safeguards (s) and price (p) are fully interactive and are determined simultaneously. Repeated reference to the schema will be made throughout the book. Indeed, it is gratifying that so many applications turn out to be variations on a theme. As Hayek observed, “whenever the capacity of recognizing an abstract rule which the arrangement of these attributes follows has been acquired in one field, the same master mould will apply when the signs for those abstract attributes are evoked by altogether different elements” (1967, p. 50).²⁸

By way of summary, the nodes A, B, and C in the contractual schema set out in Figure t-2 have the following properties:

1. Transactions that are efficiently supported by general purpose assets ($k = 0$) are located at node A and do not need protective governance structures. Discrete market contracting suffices. The world of competition obtains.
2. Transactions that involve significant investments of a transaction-specific kind ($k > 0$) are ones for which the parties are effectively engaged in bilateral trade.
3. Transactions located at node B enjoy no safeguards ($s = 0$), on which account the projected breakeven supply price is great ($\bar{p} > \hat{p}$). Such transactions are apt to be unstable contractually. They may revert to node A (in which event the special purpose technology would be replaced by the general purpose ($k = 0$) technology) or be relocated to node C (by introducing contractual safeguards that would encourage the continued use of the $k > 0$ technology).

²⁸Although I was aware that the contractual approach to vertical integration on which I was working in 1971 would have other applications (including labor market organization and regulation), little did I imagine that nonstandard modes of contracting and corporate governance would also yield to the same type of analysis upon restating the issues in contracting terms.

4. Transactions located at node C incorporate safeguards ($s > 0$) and thus are protected against expropriation hazards.
5. Inasmuch as price and governance are linked, parties to a contract should not expect to have their cake (**low price**) and eat it too (no safeguard). More generally, it is important to study *contracting in its entirety*. Both the *ex ante* terms and the manner in which contracts are thereafter executed vary with the investment **characteristics** and the associated governance structures within which transactions are embedded.

5. Economic Organization of the Company Town

The company town is mainly regarded as a painful reminder of labor abuses associated with an earlier era. Surely there is nothing favorable, much less redeeming, that can be said about such a condition.

Still, company towns were the exception rather than the rule. The question, moreover, needs to be asked, why would anyone accept employment under patently unfavorable terms? More generally, what are the relevant contractual alternatives for which a comparative assessment is needed? Inasmuch as the study of extreme instances often helps to illuminate the essentials of a situation (Behavioral Sciences Subpanel, 1962, p. 5), an **examination of the problems of organization faced by the** company town may be instructive.

The issues are addressed in two stages. The first illustrates the advantages and the second the limitations of studying economic organization from the standpoint of “contracting in its entirety.”

5.1 *Contract Analysis*

Assume the following: (1) A remote mineral source has been located, the mining of which is deemed to be economical; (2) the mineral can be mined only upon making significant investments in durable physical assets that are thereafter nonredeployable; (3) requisite labor skills are not firm-specific to any significant degree, but there are set-up costs associated with labor relocation; (4) the weather in the region is severe, which necessitates the provision of durable housing for protection from the elements; (5) the community of miners is too small to support more than one general store; and (6) the nearest city is forty miles away.

I wish to focus on two issues: Should the workers or the mining **firm** own

the homes in the community? And how should the general store be owned and operated? So as to display the relevant features more clearly, two different mobility scenarios will be considered.

a. THE IMMOBILE SOCIETY

This is **the** pre-automobile era. The firm advertises for workers, and describes the terms of employment. Given the remote location, workers will be concerned not merely with wages but also with housing and with the economic infrastructure.

Were the firm to decide to construct housing itself, it could then (1) sell the homes to the workers, (2) rent the homes to workers on short-term leases, (3) write long-term leases with severe penalties for early termination by the lessee, or (4) write long-term leases that bind the **firm** but permit easy termination by the lessee. Alternatively, the firm could (5) require workers to construct their own housing.

Given the thin market, workers who constructed their own homes would, in effect, be making firm-specific investments. Lacking contractual safeguards—buy-back clauses (whereby the company guarantees a market in the event of layoff or termination), long-term employment guarantees, lump sum severance awards, **death** benefits, and the like—workers will agree to make such investments only if offered a sign-on bonus **and/or** a wage premium. Expressed in terms of the contractual schema in Figure 1–1, that last corresponds to a node B rather than a node C result (which is to say, a $\bar{w} > \hat{w}$ outcome).

Node B outcomes, however, are notoriously inefficient. The marginal costs of the firm will be driven up by a \bar{w} wage bargain, whence the firm will make layoffs according to an inefficient criterion. Home designs chosen by the workers will likewise be compromised in consideration of the hazards. The advantages of concentrating all the specific investments on the mining firm are thus apt to be apparent to both parties at the outset (or **will** become obvious during negotiations). Accordingly, home ownership by the mining firm coupled with efficient lease terms ought to be observed. Option **4**—long-term leases that bind the lessor but provide easy release for the lessee—have obvious **attractions**.²⁹

Consider **the** general store. The leading possibilities here are: (1) The

²⁹Whether the defects of the **first** three options described in the text exceed those of option (4) is **left** as an exercise for the reader. Rental arrangements are favored **by** assuming that occupants will exercise due care, which requires that deterrents for abuses that can be implemented to the satisfaction of the parties be included in the lease agreement. For a related discussion, see Alchian (1984, p. 40).

store is owned by the mining firm and (a) operated as a monopoly, (b) placed under a fair rate of return constraint, or (c) placed under a market basket (index number) constraint; (2) a multiyear franchise is awarded to the highest bidder, the receipts from which bidding competition are (a) paid to the company treasury, (b) divided among the initial group of workers, or (c) placed in a money market fund and paid out to customers over the life of the franchise in proportion to purchases; and (3) the store is owned and operated by the workers as a cooperative. Although none of these is unproblematic, options 2c and 3 have much to recommend them.³⁰ Whatever the determination, the more general point is this: The wage bargain to which the workers agree will be conditional, rather than independent of, the way in which the general store is owned and operated if, as assumed, contract realizations reflect all of the salient features of which the ownership and governance of the general store are plainly germane.

b. THE MOBILE SOCIETY

The appearance of the automobile, mobile homes, home freezers, mail order houses, and the like greatly relieve the contracting difficulties of the premobility era. The need for site-specific investments in homes is alleviated by the invention of suitable assets on wheels, which is what the mobile home option represents. Exclusive reliance on the general store is relieved by the possibility of shopping at a distance, which cheap transportation to the nearby city and purchases from mail order houses permit. Changes in markets and technology thus have sweeping contracting ramifications. In effect, a viable node A alternative has been introduced into what had previously been a contractually complicated node B/node C choice set.

To be sure, remote mining communities may present still other issues for which careful comparative institutional assessments will be needed. Plainly, however, contractual strains of the earlier era are greatly alleviated by the mobility that assets on wheels and competition permit.

5.2 *Some Reservations*

If contracting in its entirety reliably obtains, then an efficient configuration of wages, home ownership, company store operations, and the like will appear, whatever the mobility condition of the population. What then explains the widespread discontent with the organization of company towns in the premobility era?

³⁰Again, this is left as an exercise for the reader.

There are two leading possibilities. One is that students of company towns have not performed the relevant comparative institutional tests. Rather than describe and evaluate the actual set of contractual choices from which company town organization is constrained to choose, company towns are compared instead with noncompany towns. Unsurprisingly, company towns fare poorly in the comparison. Inasmuch, however, as such a comparison is operationally irrelevant, it is wholly unhelpful to an understanding of the organizational problems with which the company town is faced.

The second possibility is that, especially in the context of labor market organization, contracting in its entirety is rarely realized. Company towns would be a good deal less objectionable if they were actually organized along efficient contracting principles. But what company town was ever organized as a cooperative? A chronic problem with labor market organization is that workers and their families are irrepressible optimists. They are taken in by vague assurances of good faith, by legally unenforceable promises, and by their own hopes for the good life. Tough-minded bargaining in its entirety never occurs or, *if it* occurs, comes too late. An objective assessment of employment hazards that should have preceded any employment agreement thus comes only after disappointment. "Demands" for redress in those circumstances are apt to be regarded as bluff-based, as they are, on weakness. Collective organization may help, but it entails a struggle. Ensuing settlements may stanch the losses rather than effect a transfusion.

I submit that both factors contribute to the low opinion with which company towns are held. As stated at the outset, however, this book does not attempt a comprehensive treatment of all the relevant factors. Instead, I consistently assume that the parties to a contract are hard-headed and that the ramifications of alternative contracts are intuited if not fully thought through. This often sheds insights, but not without cost. Omissions and distortions sometimes result. Such costs are less severe, I believe, where commercial contracting practices (including vertical integration and supporting internal governance structures) are under review than when labor market organization is being studied. In any event, my emphasis on previously neglected transaction cost features is meant to redress an earlier imbalance. I fully concur that complex contracting will be better understood if examined from several **well-**focused perspectives.

6. Applications

The applications of transaction cost economics sketched here are developed more extensively in later chapters. The object is merely to motivate the