Empirical Research in Transaction Cost Economics:
A Review and Assessment

Howard A. Shelanski
Kellogg, Huber, Hansen, Todd, and Evans, Washington, D.C.

Peter G. Klein
University of Georgia

This article summarizes and assesses the growing body of empirical research in transaction cost economics (TCE). Originally an explanation for the scale and scope of the firm, TCE is now used to study a variety of economic relationships, ranging from vertical and lateral integration to transfer pricing, corporate finance, marketing, the organization of work, long-term commercial contracting, franchising, regulation, the multinational corporation, company towns, and many other contractual relationships. The main insights and predictions of TCE—in particular, the importance of governing transactions—are becoming increasingly accepted. The empirical support for these claims, however, is much less known. We believe the empirical literature, on the whole, is remarkably consistent with the predictions of TCE—more so than is typically the case in economics. After presenting an overview of the theory and a discussion of some theoretical and methodological preliminaries, we summarize the major findings and discuss their implications, particularly the potential applications to public policy. In an appendix we provide a more comprehensive list of articles, arranged by type of study, as a reference aid for researchers.

1. Introduction

[The new institutional economics] suggests a whole agenda of microeconomic empirical work that must be performed. . . . Until that work has been carried out . . . the new institutional economics and related approaches are acts of faith, or perhaps of piety.

—Herbert Simon (1991: 27)

Transaction cost theories of exchange, part of what has been termed the "New Institutional Economics," have been the subject of growing interest in recent years. Originally an explanation for the scale and scope of the firm, transaction
cost economics (TCE) is now used to study a variety of economic phenomena, ranging from vertical and lateral integration to transfer pricing, corporate finance, marketing, the organization of work, long-term commercial contracting, franchising, regulation, the multinational corporation, company towns, and other contractual relationships, both formal and informal. The basic insight of transaction cost economics—that transactions must be governed as well as designed and carried out, and that certain institutional arrangements effect this governance better than others—is now increasingly accepted. The purpose of this article is to outline this empirical literature and to offer a preliminary assessment of its impact on the development of TCE theory and related fields. We make no effort to examine all the existing empirical work in the TCE tradition; such a project would require several essays. Instead, we focus on several key empirical issues or phenomena on which we think TCE has enabled researchers to make substantial progress. We find that, on balance, a remarkable amount of the empirical work we have examined is consistent with TCE predictions—much more so, perhaps, than is the case with most of industrial organization.¹

The article is organized as follows. Section 2 presents a brief overview of TCE, and Section 3 discusses some theoretical and methodological issues related to the testing of TCE hypotheses. In Section 4 we summarize the relevant empirical research, organized by topic; Section 5 covers the public policy implications of this evidence. We summarize and conclude in Section 6. The appendix provides a comprehensive list of articles, arranged by type of study, as a reference aid for researchers.

2. Transaction Cost Economics: An Overview

Transaction cost economics studies how trading partners protect themselves from the hazards associated with exchange relationships. As developed by Williamson (1975, 1985, 1995), Klein, Crawford, and Alchian (1978), and more formally by Grossman and Hart (1986) and Hart and Moore (1990), TCE maintains that in a complex world, contracts are typically incomplete.² Because of this incompleteness, parties who invest in relationship-specific assets expose themselves to a hazard: If circumstances change, their trading partners may try to expropriate the rents accruing to the specific assets. One way to safeguard those rents is through integration, where the parties merge and eliminate adversarial interests. Less extreme options include reciprocal buying arrangements, in which each party exposes itself to form a mutual exchange of "hostages," and partial ownership agreements. In general, a variety of such "governance structures" may be employed; the appropriate one depends on the particular characteristics of the relationship. In this way, TCE may be considered the


². This may be because agents are boundedly rational, or because certain quantities or outcomes are unobservable (or not verifiable to third parties, such as the courts), in which case contracts cannot be made contingent on these variables or outcomes.
study of alternative institutions of governance. Its working hypothesis, as expressed by Williamson (1991: 79), is that economic organization is really an effort to "align transactions, which differ in their attributes, with governance structures, which differ in their costs and competencies, in a discriminating (mainly, transaction cost economizing) way." Simply put, TCE tries to explain how trading partners choose, from the set of feasible institutional alternatives, the arrangement that offers protection for their relationship-specific investments at the lowest total cost.

Transactions differ in a variety of ways: the degree to which relationship-specific assets are involved, the amount of uncertainty about the future and about other parties' actions, the complexity of the trading arrangement, and the frequency with which the transaction occurs. Each matters in determining the preferred institution of governance, although the first—asset specificity—is held to be particularly important.\(^3\) Williamson (1985: 55) defines asset specificity as "durable investments that are undertaken in support of particular transactions, the opportunity cost of which investments is much lower in best alternative uses or by alternative users should the original transaction be prematurely terminated." This could describe a variety of relationship-specific investments, including both specialized physical and human capital, along with intangibles such as R&D and firm-specific knowledge or "capabilities."

Governance structures can be described along a spectrum. At one end lies the pure, anonymous spot market, which suffices for simple transactions such as basic commodity sales. Market prices provide powerful incentives for the exploitation of profit opportunities, and market participants are quick to adapt to changing circumstances as information is revealed through prices. When specialized assets are at stake, however, and when product or input markets are thin, bilateral coordination of investment decisions may be desirable, and combined ownership may be efficient. At the other end of the spectrum from the simple, anonymous spot market thus lies the fully integrated firm, where trading parties are under unified ownership and control. TCE posits that such hierarchies offer greater protection for specific investments and provide relatively efficient mechanisms for responding to change where coordinated adaptation is necessary. Compared to decentralized structures, however, hierarchies provide managers weaker incentives to maximize profits and normally incur additional bureaucratic costs.\(^4\) Between the two poles of market and hierarchy are a variety of "hybrid" modes, such as complex contracts and partial ownership arrangements. The movement from market to hierarchy thus entails a trade-off between the high-powered incentives and adaptive properties of the market, and the safeguards and central coordinating properties of the firm.\(^5\)

\(^3\)Indeed, TCE (associated mainly with Williamson) is sometimes described as the "governance" branch of the New Institutional Economics, as opposed to the "measurement cost" or agency-theoretic branch (associated with Alchian and Demsetz, 1972).

\(^4\)An example of these bureaucratic costs would be the "influence costs" studied by Milgrom and Roberts (1990).

\(^5\)The general theoretical framework of TCE is now sufficiently accepted to have been incorpo-
Implicit in TCE is a notion that market forces work to bring about an "efficient sort" between transactions and governance structures, so that exchange relationships observed in practice can be explained in terms of transaction cost economizing. The existence of this selection mechanism is usually assumed rather than explained, though, and thus TCE is subject to some of the same criticisms that evolutionary economists (Nelson and Winter, 1982, for example) have made of standard microeconomic theory. Some students of business organization have also charged TCE with having a too narrowly "economic" or efficiency-oriented view of individual and firm behavior. To be sure, TCE usually abstracts away from issues of market power, resource dependence, social embeddedness, and the like; the bulk of the empirical literature inspired by TCE takes as given an economizing framework. The basic framework is applicable to a wide range of phenomena. While vertical and lateral integration are perhaps the best known examples, there are many others.

3. Some Theoretical and Methodological Preliminaries
Much of the empirical work in TCE can be considered a variation of the following basic model. The efficient form of organization for a given economic relationship—and, therefore, the likelihood of observing a particular organizational form or governance structure—is a function of certain properties of the underlying transaction or transactions: asset specificity, uncertainty, complexity, and frequency. Organizational form is the dependent variable, while asset specificity, uncertainty, complexity, and frequency are independent variables. Specifically, the probability of observing a more integrated governance structure depends positively on the amount or value of the relationship-specific assets involved and, for significant levels of asset specificity, on the degree of uncertainty about the future of the relationship, on the complexity of the transaction, and on the frequency of trade.

Organizational form is often modeled as a binary variable—"make" or "buy," for example—though it can sometimes be parameterized by a continuous variable. Of the independent variables, asset specificity is the most difficult to measure. Among the common proxies are component "complexity," qualitatively coded from survey data, as a proxy for physical asset specificity (Masten, 1984); worker-specific knowledge, again coded from survey data, as a proxy for human asset specificity (Monteverde and Teece, 1982b); physical proximity of contracting firms, as a proxy for site specificity (Joskow, 1985, 1987, 1988b, 1990; Spiller, 1985); and R&D expenditure, as a proxy for physical asset specificity. Other proxies, such as fixed costs or "capital intensity," have more obvious limitations and are rarely used.

The empirical work in TCE uses a variety of econometric and historical methods. In general, these studies fall into one of three categories: qualitative case studies, quantitative case studies, and cross-sectional econometric analy-

---

rated in a number of textbook treatments. See, for example, Kreps (1990: 744–90), Baye and Bell (1994), Milgrom and Roberts (1992), and Rubin (1990).
ses. Williamson’s (1976) study of cable TV franchising in Oakland, California, is an example of the first category, while Masten’s (1984) investigation of contracting practices in a large aerospace corporation is an example of the second, and Levy’s (1985) study of vertical integration across industries is an example of the third. The bulk of the empirical literature in TCE consists of case analyses of various kinds. This is primarily because the main variables of interest to transaction cost economists—asset specificity, uncertainty, frequency—are difficult to measure consistently across firms and industries. Typically, these characteristics are estimated based on surveys or interviews: for example, a manager might be asked to rate on a Likert-type scale of 1 to 7 the degree to which an investment has value in outside uses. Such data are of course subject to the general limits of survey data; namely, that they are based on the respondents’ stated beliefs, rather than on their beliefs or valuations as revealed through choice. More important, since these measurements are based on ordinal rankings, it is hard to compare them from industry to industry. What is ranked as a relatively specialized asset in one firm may be rated differently in another firm or industry. Similarly, what one firm considers a comparatively uncertain production process may be the standard operating environment in another. Multi-industry studies therefore may contain variables that are labeled the same thing but are really incommensurable or, conversely, may contain variables that are identical but labeled differently.

Besides these measurement difficulties, empirical research in TCE is often hampered by confusion about the definitions, and therefore the empirical parameterizations, of key variables. The primary conceptual problem that we have found lies in the treatment of uncertainty as a factor that raises transaction costs and increases the probability of integration. This confusion may explain some seemingly contradictory results on the effects of sales volume uncertainty on the vertical integration decision.

Uncertainty about future events is of course a common feature of many trading relationships; sales volume uncertainty due to volatile market conditions is an obvious example. Empirical studies sometimes treat this kind of uncertainty as an independent variable, regressing the choice of organizational form on the variance of sales or another variable, but without including any measure of asset specificity in the model. Absent fixed investments, however, TCE does not predict that uncertainty would itself lead to hierarchical governance. Changes in circumstances allow for expropriation only when there are quasi-rents at risk; that is, when one side’s investment is exposed. When there are no relationship-specific investments at stake, it may be less costly for a firm to contract on the market for goods and services in an uncertain environment than to assume the risk of producing them internally. In this way, the effect of uncertainty depends on competitive conditions. If there is no asset specificity and thus there are many potential suppliers of a component for which future demand is uncertain, it may be cheaper to buy the component than to make it internally.

The effect of uncertainty on governance structure thus hinges on asset specificity and the consequent bilateral dependency. The failure of some studies
to take this into account may explain a few conflicting results on the effects of uncertainty. Hence, Harrigan's (1986) finding that uncertainty reduced the probability of integration in a large, cross-sectional sample may be reconciled with opposite results by Levy (1985) and Macmillan, Hambrick, and Pennings (1986), as Harrigan abstracts from asset specificity in her study. In Walker and Weber's (1987) study of automobile parts procurement, they test the interactive effects of uncertainty and competition by dividing the sample according to the level of supplier-market competition for that component, and then testing the role of uncertainty on each part of the sample separately. They find that sales volume uncertainty, as expected, increases the probability of a "make" rather than "buy" decision, for those components produced in thin markets.6

Asset specificity has been more successfully treated in the empirical literature than has uncertainty. Relationship-specific physical, site, and human capital investments have all been studied, both independently and comparatively. Further refinement and analysis need to be done here, however, particularly in the area of measurement. Proxies such as capital intensity or fixed costs are very imperfect, and may not capture whether an investment has alternative value outside the transaction for which it was initially made. Another concern is that asset-specificity effects may be confused with market power. While specific investment may lead to bilateral monopoly, the existence of a small-numbers bargaining situation is not by itself evidence of relationship-specific investment.

Besides the difficulties of measurement and definition that are unique to TCE, empirical TCE is also subject to the problems found in empirical work more generally: namely, alternate hypotheses that could also fit the data are rarely stated and compared. Usually, the data are only found consistent or inconsistent with the hypothesis at hand. We believe there is a need and opportunity for studies that explicitly compare competing, observationally distinct, hypotheses about contractual relationships, because rival theories commonly posit mutually exclusive outcomes. One example is Spiller's (1983) comparison of asset-specificity and market-power explanations of vertical merger, explanations that have rival predictions about the size of the gains from merger under various competitive conditions. Another prototype for such a project might be MacDonald's (1985) cross-sectional study of vertical integration, which incorporated elements of both TCE and Stigler's theory of the vertical "life-cycle" of the firm (though it did not attempt to distinguish between them). Further studies

6 There may also be situations where uncertainty is so great that efficient governance structures cannot be crafted at all, in which case trade may fail to materialize. While there is a considerable stream of theoretical literature, following Akerlof (1970), on the possibility that markets might break down due to private information, there is relatively little theoretical or empirical work on nonmarket exchange under these conditions. An exception is Wiggins and Libecap's (1985) study of utilization agreements in oil production. Under such an agreement, producers designate a single firm to develop a given field, with the net returns shared among all producers. This reduces recovery costs and improves oil yields by eliminating the negative externalities associated with concurrent independent development of a single field. Yet very few oil fields are unitized. Wiggins and Libecap argue that asymmetric information encourages opportunistic holdout strategies that usually have prevented the agreements from being signed.
of this kind are essential to a better understanding of the predictive power of
the transaction cost model.

4. Comparative Contracting: A Sampling of the Evidence
The goal of transaction cost economics is to explain contracting arrangements
observed in practice. Where possible, TCE tries to explain these phenomena
on efficiency grounds. We consider below five major categories of empiri-
cal phenomena explained by TCE: vertical integration, "hybrid" contracting
modes, long-term commercial contracts, informal agreements, and franchise
contracting. We believe that TCE has added to our understanding of each of
these practices.

Because asset specificity is usually the main variable of interest in these
studies, we should keep in mind Williamson's (1983) distinction between four
distinct types of asset specificity. The first is site specificity, in which parties
are in a "cheek-by-jowl" relationship to minimize transportation and inventory
costs, and assets are highly immobile once in place. The second is physical
asset specificity, referring to relationship-specific equipment and machinery.
The third is human asset specificity, describing transaction-specific knowledge
or human capital, achieved through specialized training or learning-by-doing.
The fourth is "dedicated assets," referring to substantial, general-purpose in-
vestments that would not have been made outside a particular transaction, the
commitment of which is necessary to serve a large customer.

We now consider some main findings of the empirical literature.

4.1 Vertical Integration
Vertical integration, or the "make-or-buy" decision, has been described as the
"paradigm problem" of TCE. Much of the earliest empirical work addresses this
topic.° Monteverde and Teece (1982b) made one of the first systematic efforts to
test a contractual interpretation of vertical integration. They examine the effects
of asset specificity, defined here as worker-specific knowledge or "applications
engineering effort," on the decision to produce components in-house or to obtain
them from outside suppliers. Starting with a list of 133 automobile components,
each coded as either made or bought, the authors enlisted an automobile design
engineer to develop an index measuring the degree of applications engineering
effort involved in the production of each component. Their thesis is that "[t]he
greater is the applications engineering effort associated with the development
of any given automobile component, the higher are the expected appropriable
quasi-rents and, therefore, the greater is the likelihood of vertical integration of
production for that component" (207).

Monteverde and Teece test a probit model of the component sample, re-
gressing "make" or "buy" on the degree of applications engineering effort.
They also include proxies for whether the component is specific to the man-

7. For more on these early efforts, see the discussions in Williamson (1985: 103–30), and Joskow
ufacturer or generic, and a proxy for the specific company involved (Ford or G.M., to pick up any idiosyncratic firm-specific effects). They find applications engineering effort to be a statistically significant determinant of backward integration. The results are consistent with an earlier finding by Globerman (1980) on firm-specific technical knowledge and integration in the Canadian telecommunications industry. Globerman studied evidence from public hearings and found a tendency toward common ownership of telephone lines and equipment as the research and development demands of a carrier on its equipment suppliers become more complex and uncertain and require more relationship-specific investments.  

Other studies of component procurement in the auto industry find similar support for transactional explanations of vertical relationships. Two studies by Walker and Weber (1984, 1987) focus on uncertainty as a determinant of vertical integration. Like Monteverde and Teece, they work with a list of automobile components, coded as made or bought, as the dependent variable. They find that greater uncertainty about production volume raises the probability that a component is made in-house, but that "technological uncertainty," measured as the frequency of changes in product specification and the probability of technological improvements, has little effect. Their second study (1987) includes measures of market competition, testing the interactive effects of both uncertainty in production and competition among suppliers, and adds the qualification that volume uncertainty matters only when supply markets are thin.

In a further refinement, Masten, Meehan, and Snyder (1989) attempt to distinguish among types of specific assets, comparing the relative importance of relationship-specific human and physical capital. They also study automobile component production, finding that engineering effort, as a proxy for human asset specificity, appears to affect the integration decision more than physical or site specificity. Klein (1988), in a discussion of the G.M.–Fisher Body case, also suggests that specific human capital in the form of technical knowledge was a major determinant of G.M.'s decision to buy out Fisher. Other studies have documented a similar link between integration and research and development (R&D), which usually involves specific human capital (Armour and Teece, 1980; Globerman, 1980; Joskow, 1985; Pisano, 1990).

Asset specificity, then, appears to be an important determinant of vertical

---

8. In a related study, Monteverde and Teece (1982a) look at "quasi-integration" in the auto industry, where quasi-integration refers to assembler ownership of tooling equipment. They find a significant, positive relationship between appropriable rents and quasi-integration, although the proposed explanation accounts for only a small proportion of the total variation in integration patterns.

integration, particularly when examined together with uncertainty and product complexity. Site specificity, dedicated assets, and the need for specifically tailored products or production facilities have been shown to increase vertical integration in a variety of industries, including electricity generation (Joskow, 1985), aerospace (Masten, 1984), aluminum (Stuckey, 1983; Hennart, 1988), forestry (Globerman and Schwindt, 1986), chemicals (Lieberman, 1991), and offshore oil gathering (Hallwood, 1991).

Spiller (1985) examines site specificity in an attempt to distinguish between TCE and market-power explanations for vertical mergers. While TCE predicts that the gains from merger should be increasing in the degree of asset specificity, market-power considerations suggest that the gains will be increasing in the degree of supplier-market concentration. Using site specificity, defined as the proximity of the merging firms, to represent asset specificity, Spiller studies the gains from merger according to unexpected changes in the firms’ stock prices at the announcement of the merger. He finds the total gain from merger to be smaller the greater the distance between the merging firms (the lower the site specificity), whereas the industry concentration ratio has no significant effect. This appears to support the asset-specificity explanation over the market-power explanation.

Except for Spiller (1985), all the papers cited above are case studies of particular industries or production processes. As such, they avoid the problems discussed in Section 2 of inconsistent measurement across industries. Indeed, the limits of interindustry studies in industrial organization more generally have been recognized for some time (Joskow, 1988a: 111). These case studies have measurement difficulties of their own, however. The classification of dichotomous variables like “make-or-buy,” for example, typically is based on survey data and may require the researcher to exercise a certain amount of discretion or intuitive judgment. Nonetheless, most of the empirical work in TCE on vertical integration has been of this type. While it is of course difficult to generalize the results, the cumulative evidence from different studies and industries is quite consistent with the basic theory.

Also, there do exist some cross-sectional studies on transactional determinants of vertical integration using multi-industry data. An early effort by Levy (1985) uses the ratio of value-added to sales as a cross-industry measure of vertical integration, the number of firms and amount of R&D spending as measures of asset specificity, and the variance of sales as a measure of uncertainty. Using data from 69 firms representing 37 industries for the years 1958, 1963, 1967, and 1972, he finds each of the independent variables to have a statistically significant effect on the likelihood of vertical integration. Macmillan, Hambrick, and Pennings (1986) obtain very similar results with a larger sample. Harrigan (1985), by contrast, finds sales variability to result in a lower chance of vertical integration, although she does not include a measure for asset specificity.

Accounting constructs like value-added-to-sales ratios, such as those used by Levy, are highly problematic as measures of vertical integration. Caves and Bradburd (1988) construct a more complicated cross-industry measure of integration based on an input–output matrix of distribution shipments across
several industries. They use this metric to compare asset specificity, small-numbers bargaining conditions, and risk as determinants of vertical integration. They find asset specificity and small-numbers situations, but not risk, to be significant. TCE-based hypotheses thus do well in their study as compared to competing approaches. The approach of Caves and Bradburd is promising and warrants further exploration where possible. Unfortunately, their procedures are exceptionally data-intensive and may not be feasible in many cases. Other potentially fruitful approaches use financial data on merging firms’ pre- and post-merger performance, either to study the gains from merger as a function of asset specificity (Spiller, 1985) or to examine the likelihood of merger as a function of pre-merger bilateral relationships (Weiss, 1992).

While economists typically think of vertical integration as backward integration into components, materials, or R&D, forward integration into marketing and distribution may be just as important. Several studies of the integration of marketing channels have used TCE as an explanatory framework. Anderson and Schmittelein (1984) consider two marketing alternatives for an electronics component producer: the use of employees as a direct sales force (a form of vertical integration) vs. reliance on independent manufacturers’ representatives. This choice is regressed on managers’ perceptions of the importance of specific human capital, sales volume uncertainty, and measurement uncertainty (all based on survey data), each of which is predicted to increase the likelihood of a direct sales force. Both specific human capital and measurement uncertainty are statistically significant, though sales uncertainty is not. A second study by Anderson (1985), also on the electronics industry, finds the same basic results, as does work by John and Weitz (1988) using data from a variety of industrial-product industries. Marketing and distribution depend on other factors as well, of course. A recent study of the carbonated beverage industry by Muris, Scheffman, and Spiller (1992) finds that the shift from independent bottlers to captive subsidiaries over the last 20 to 30 years can be explained without reference to changes in asset specificity. Instead, they account for the shift in terms of the emergence of national cola markets, which required greater coordination of advertising and promotional activities. Along with changing technologies in cola production and distribution (namely, falling transportation and communication costs), it was this need for more centralized decision-making—for given levels of asset specificity—that explains the change toward a more vertically integrated industry.

To sum up, the evidence on the transactional determinants of vertical integration seems quite striking. Asset specificity and uncertainty appear to have significant effects on the vertical structure of production. This is especially remarkable when compared with the relative dearth of evidence on market-power explanations for integration, and with the results of rare studies that explicitly compare TCE-based theories with market-power theories (Spiller, 1985).

4.2 Complex Contracting and "Hybrid" Modes
Integration, as discussed previously, is an extreme form of internal governance. Intermediate forms also exist: long-term contracts, complex contracts with
reciprocity agreements, agreements to provide offsetting specific investments ("hostages"), equity linkages, and so on. These may be adopted because the benefits of full integration are not worth the costs, given the properties of the transaction, or because integration is prohibited by regulation.

A simple example of a hybrid or intermediate form of governance is an exclusive dealing contract. Gallick (1984) examines such contracts in the U.S. tuna industry. Exclusive dealing, he argues, is an efficient means of discouraging opportunism by fishing boat captains. Because most tuna sold in the U.S. is canned, it is cheaper for tuna processors to buy a boat's output at a price reflecting average quality than to pay for the inspection, sorting, and grading usually found in fresh fish markets. Exclusive dealing arrangements prevent the boat captains from selling the higher-quality tuna, ex post, to rival processors at higher prices. Until the mid-1960s, reputation effects were sufficient to enforce the contracts; the practice of frequent, small deliveries prevented the processors from trying to renegotiate the terms of trade after the catches had been made. Interestingly, Gallick reports, when fishing technology changed such that catches could be much larger—increasing the short-term gains to the processor from reneging on an exclusive dealing agreement after the catch—reputation was no longer an efficient enforcement mechanism, and a new institutional practice emerged: co-ownership of fishing vessels by the boat captains and the processors.10

In another context, Heide and John (1988) study marketing relationships between manufacturers and sales agencies. To service a particular manufacturer, sales agencies typically make investments specific to that manufacturer—most often, a human-capital investment in developing a sales territory for the manufacturer's product. Because agencies are small relative to manufacturers, they cannot safeguard their investments by backward integration into manufacturing. Similarly, they lack the bargaining power to demand long-term contracts with manufacturers. Instead, they protect their relationship-specific assets by making other specific investments, namely in routines or procedures that tie or "bond" them with a manufacturer's customers. These might be the establishment of personal relationships with the customers, the development of an identity separate from the manufacturer's particular product, or the creation of specialized procedures for ordering, shipping, and servicing the product. In this way they "balance their dependence" on the manufacturer with the customers' dependence on them.

Heide and John use data from 199 manufacturers' agents to test this dependence-balancing hypothesis. The evidence is supportive: agencies with specific assets invested in the agency—manufacturer relationship tended to make more offsetting investments to lessen their dependence on the manufacturer. This

10. On the antitrust implications of such exclusionary practices, see Masten and Snyder's (1993) analysis of United States v. United Shoe Machinery Corporation. They interpret United's equipment-leasing provisions as a means of governing the operation and maintenance of complex shoe machines.
suggests quite clearly the importance of asset specificity in determining the governance of marketing channel exchange.

Pisano (1990) asks why firms may rely on equity linkages instead of contracts to support certain transactions. He argues that partial ownership will dominate contractual governance when a relationship involves uncertainty, transaction-specific capital, and other variables. He tests the hypotheses that (i) equity linkages are more likely when R&D is to be performed during collaboration; (ii) equity arrangements are more likely when collaboration encompasses multiple projects; and (iii) equity arrangements are less likely in environments in which there are more potential collaborators. Each of the three hypotheses is supported by tests using data from 195 collaborative arrangements in the biotechnology industry. The study corroborates TCE explanations for the role of equity linkages in governing exchange. Pisano, Russo, and Teece (1988) apply a similar analysis to the telecommunications equipment business and find that the same basic framework can explain the choice between equity linkages and other forms of cooperative ventures (joint ventures, consortiums, or non-equity linkages).

In another hybrid mode case study, Eccles (1981) uses a transaction cost framework to explain the existence of the prime-contractor/subcontractor organizational unit (“quasi-firm”) in the home construction business. He argues that the quasi-firm can be explained as an efficient governance structure for the construction industry. Using data from interviews with 38 home-building firms, Eccles finds strong support for the role of governance costs in explaining the quasi-firm structure. Also, he finds that the quasi-firm structure allows market suppliers to be governed like integrated organizational units. If such contractual relations are efficient at high levels of asset specificity, then TCE may assign too much attention to asset specificity for the integration decision. Alternatively, it could be that the observed level of asset specificity is close to a threshold or switchover value, in which case either organizational alternative is efficient. Clearly, more work needs to be done on hybrid modes of organization before the implications of Eccles's results can be fully understood.

4.3 Long-term Commercial Contracts
A series of studies by Joskow (1985, 1987, 1988b, 1990) investigates the effects of asset specificity on contract duration and price adjustment in agreements between coal suppliers and coal-burning electrical plants. He examines a large sample of coal contracts and finds strong support for the hypothesis that the greater are relationship-specific investments (in this case, site specificity and dedicated assets), the longer are the periods covered by the contract. Furthermore, he finds that long-term contracts performed well despite large fluctuations in the nominal price of coal. This suggests that long-term contracting can be a feasible alternative to integration when asset specificity is moderate. Crocker and Masten (1988) found similarly that contracts in the natural gas industry tended to cover longer terms when specific assets were involved. More generally, they argue that efficient contract duration depends on the costs of contracting—contract terms become shorter, for example, as uncertainty in-
creases. Goldberg and Erickson (1987) analyze 90 contracts for petroleum coke written between 1946 and 1973 and conclude that many provisions of the contracts can best be interpreted as efforts by the parties to protect themselves against expropriation of specialized investments.\(^{11}\)

DeCanio and Frech (1993) tried to measure more precisely the efficiency gains from long-term contracts in natural gas. Relationship-specific investments are critical for transactions between wellhead owners and pipelines. For that reason, "take-or-pay" contracts, in which the buyer is required to pay for some minimum quantity even if delivery is not taken, are often used to safeguard against buyer (pipeline) opportunism.\(^{12}\) In 1987, the Federal Energy Regulatory Commission (FERC) eliminated such long-term agreements. The authors use data from before and after the FERC order to test its effect on spot gas prices and prices at the wellhead. They find that FERC’s interference with parties’ ability to craft long-term governance mechanisms raised natural gas prices between 21 and 31 percent in the year following FERC’s order. The results support TCE explanations for the relative efficiency of long-term contracts where asset specificity is required, while representing an effort to quantify that efficiency gain.

Pirrong (1993) has recently argued that long-term contracts (and sometimes vertical integration) can be efficient in the presence of smaller contracting hazards—even when obvious physical, human, and site asset specificities are absent. In a study of bulk shipping, he finds that more integrated governance structures can dominate spot trading in the presence of what Masten, Meehan, and Snyder (1991) call “temporal specificities.” When a processing or refinery plant contracts with a particular bulk carrier, for example, both plant and carrier capacities suddenly become specific assets. Small delays in delivery can then result in large losses of quasi-rents for the plant, just as the plant’s refusal to take full delivery can impose substantial losses on the carrier. Hence, Pirrong concludes, “spatial/temporal proximity is a form of relationship-specific capital” (1993: 943), at least when markets are thin. To avoid costly strategic bargaining, then, these parties will choose a complex, long-term agreement.

A key feature of long-term contracts is their incompleteness. Indeed, TCE holds that all complex contracts are necessarily incomplete; otherwise, why would specialized governance arrangements be necessary? Yet the degree of incompleteness of a contract need not be exogenous. If there are degrees of incompleteness—the extent to which renegotiation procedures are specified, for example—it then becomes important to study how complete a contract should be. In a recent study of Air Force engine procurement, Crocker and Reynolds (1993) tested the relationship between contractual incompleteness and the likelihood of opportunistic behavior. Using a sample of procurement agreements from the 1970s and 1980s, they assigned each contract a measure of incompleteness and regressed this on variables representing the contractor’s history of

---


12. Mullerin (1986) and Masten and Crocker (1985) also examine “take-or-pay” contracts.
litigiousness and a dummy for dual sourcing (representing the expected degree of ex post opportunism); the time between contract agreement and delivery, and historical failure rates for the engine type (proxies for environmental uncertainty); and other structural variables.

The results are significant and in the expected directions for a variety of specifications (including ordinary least squares and ordered probit). Contracts are observed to be more complete when the contractor has a history of disputes with purchasers and less complete when there are increases in associated intertemporal or technological uncertainty (increasing the cost of writing more complete contracts). One implication is that federal rules governing military procurement should allow for flexibility in contract design, because the optimal contract will vary from case to case, depending on the attributes of the transaction. Most important, this study shows that the degree of contractual completeness may reasonably be treated as an endogenous variable.

4.4 Informal Agreements

TCE pays special attention to the importance of "private ordering" for dispute resolution, in contrast to the older tradition of "legal centralism" (Williamson, 1985: 20–21). Several studies have investigated whether informal trade arrangements, which are not legally enforceable, may also be motivated by the desire to make exchange more efficient. Important work in this area has been done by Palay (1984, 1985). In two closely related papers, Palay studies the role of informal, legally unenforceable agreements between rail-freight carriers and shippers. He argues that Interstate Commerce Commission regulation of the industry, which prohibits vertical integration of carriers and shippers, was geared to "classical contracting" (Macneil, 1978) but is inappropriate for transactions requiring more complex agreements. Shipment of items like automobile parts and chemicals, for example, requires specially designed rail cars and equipment that cannot be easily redeployed for other uses. Palay's hypothesis is that informal agreements, substituting for combined ownership, would emerge both to encourage and to protect these relationship-specific investments. Furthermore, he argues that the underlying characteristics of a transaction predict whether it will be supported by an informal agreement. Evidence from 51 case studies of shipper–carrier transactions reveals a pattern of informal agreements highly consistent with TCE. Equipment tailored for particular users—custom carrier racks for automobile parts, tank and covered hopper cars for specific volatile chemicals, and so on—was owned by individual shippers. Equipment for more standardized shipments would be owned by rail carriers. The informal agreements also provided handling procedures for unusual circumstances related to shipment. The transactions that did not use informal contracting all involved nonspecialized capital such as standard box cars. All of this suggests the importance of asset specificity for complex contracting.

13. The recent work on private law and its evolution by Ellickson (1991) and Benson (1990), for example, is in this same spirit.
Two studies of New England fishing industries also examined the role of transaction costs in determining trade agreements and market structure. Wilson (1980) conducted an intensive study of the New England fresh fish market. He found that underlying the smooth functioning of the market was a system of mutual dependence created by the particular trade arrangements there; reputation effects provided an enforcement mechanism. Acheson's (1985) study of the Maine lobster market reached similar conclusions, finding the lobster market to be characterized by long-term, informal relationships between fishermen and lobster-pound operators. Fisherman and pound operators typically crafted agreements designed to reduce the costs of information and the possibility of opportunistic use of informational asymmetries. The agreements were reinforced by reputation considerations and interdependencies arising from the sharing of scarce resources, such as market information, boat fuel, and bait. Informal agreements and norms in eighteenth- and nineteenth-century whaling have been studied similarly by Ellickson (1989) and Gifford (1993).

Finally, in an interesting application of TCE to the context of personal relationships, Brinig (1990) employs transaction cost reasoning to explain the sudden increase in the demand for diamond engagement rings in the mid-1930s. The increase, she argues, can be traced to the abolition in several states of the "breach of promise to marry action" around the same time. Before this action was abolished, a broken engagement could trigger a lawsuit, because a woman in this situation faced considerable loss of reputation. Once the cause of action was eliminated, however, another arrangement was needed to ensure the credibility of the marriage commitment. Diamond engagement rings filled that role. In this way, rings may be seen as a governance structure: they safeguard the future bride's relationship-specific investment—her good reputation.

In general, although none of the agreements in the above studies were legally enforceable, they were not easily broken. The reputation effects and reciprocity provisions embodied in these arrangements evidently work well and provide strong safeguards for the parties involved; the short-term gains from opportunism are largely offset by long-term losses from a damaged reputation in the particular industry community. These empirical studies support transaction cost reasoning not only because they find that observed arrangements can be explained in terms of asset specificity, uncertainty, and the like, but also because they reflect an emphasis on private ordering over resort to the courts.

4.5 Franchise Contracting

Williamson's (1976) case study of the Oakland, California, cable TV (CATV) franchise was an early empirical study using transactional reasoning. Responding to the Posner-Demsetz argument that competitive bidding for monopoly franchises would result in competitive prices, Williamson claimed that once idiosyncratic investments are in place, what was a large-numbers bargaining situation during the bidding process is transformed into a bilateral monopoly. Because of this, the terms of the original contract may no longer be applicable. Williamson outlined the difficulties faced by the city of Oakland in the early 1970s over its CATV franchise. The franchise was awarded to the lowest bidder
in 1970. After the franchise was awarded, however, the construction process went more slowly than expected, fewer households signed up than predicted, and costs escalated. Consequently, the franchisee requested a renegotiation of the contract. A complex dual-source agreement was eventually reached, and the outcome in no way reflected the intent of the initial agreement.

Two later studies of CATV have looked for similar problems, with mixed results. Zupan (1989a) examined a series of public cable franchise agreements, comparing the terms of trade struck during the original franchise agreement with those prevailing at the time of renewal, after relationship-specific investments had been made; he found no significant differences in those terms. Frager (1990), however, found that opportunistic behavior by the franchisee, as perceived by cable customers, was higher for franchises awarded through competitive bidding.14

Of course, it is not always the franchisee who is engaged in opportunism; the franchisor may behave in opportunistic fashion as well. Grandy’s (1989) examination of 19th-century railroad regulation in New Jersey finds that the railroads in that state were willing to make large specialized investments only when they were protected by “special corporation charters” limiting state action against them. Levy and Spiller’s (1994) comparative study of telecommunications regulation in Argentina, Chile, Jamaica, the Philippines, and the U.K. shows that private investment is forthcoming only when regulators can commit not to pursue arbitrary administrative actions. Furthermore, many private franchise contracts can also be explained in terms of TCE (Norton, 1989; Dnes, 1992).

Besides the contractual phenomena described above, TCE has been brought to bear on such diverse topics as labor market contracts and regulation (Barker and Chapman, 1989), tie-ins and “block booking” (Kenney and Klein, 1983), international trade and the multinational corporation (Hennart, 1989; Yarbrough and Yarbrough, 1987b; Gatignon and Anderson, 1988; Klein, Frazer, and Roth, 1990), company towns and company stores (Fishback, 1986, 1992), land tenure agreements (Roumasset and Uy, 1980; Alston and Higgs, 1982; Alston, Datta, and Nugent, 1984; Datta, O’Hara, and Nugent, 1986) and even indentured prostitution (Ramseyer, 1991). These and other “nonstandard” contracting practices, when viewed through a transaction cost lens, often turn out to have efficiency properties, particularly in offering safeguards for specific investments.

5. Public Policy Implications and Influence

Theoretical and empirical TCE research has strong implications for antitrust, regulation, and other aspects of public policy. The full title of Williamson’s 1975 book, after all, is Markets and Hierarchies: Analysis and Antitrust Implications. A basic conclusion of TCE is that vertical mergers, even when

14. Even when successful at curbing opportunism, though, the agreements do not always induce efficient pricing (Zupan, 1989b).
there are no obvious gains in technological possibilities, may enhance efficiency by reducing governance costs. Hence Williamson (1985: 19) takes issue with what he refers to as the "inhospitality tradition" in antitrust; namely, that firms engaged in nonstandard business practices like vertical integration, customer and territorial restrictions, tie-ins, franchising, and so on, must be seeking monopoly gains. In the 10 years between the celebrated Schwinn (1967) and GTE-Sylvania (1977) cases, Williamson argues, economists began to incorporate transaction cost considerations into their understanding of vertical restrictions. This change in the intellectual climate was reflected in the Supreme Court's reversal in GTE-Sylvania of its earlier position that vertical restraints are necessarily anticompetitive.15

However, as Joskow (1991: 79–80) points out, this change may reflect sensitivity to claims that vertical integration and restraints need not reduce competition, rather than to claims that such arrangements provide contractual safeguards. While TCE proponents argued that nonstandard business practices may reduce transaction costs, Chicago school writers like Posner, Peltzman, and Bork were maintaining that such practices do not necessarily result in reduced competition. Therefore it is not certain to what extent TCE, as compared to complementary though distinct developments in industrial organization, has contributed to the observed changes in antitrust enforcement.

Joskow argues, more generally, that much of TCE is problematic for policy purposes:

The hard problems in antitrust and regulatory economics often involve potential tradeoffs between apparent increases in market power and potential reductions in costs or between regulatory imperfections and organizational or contractual imperfections. To perform such tradeoffs, we need more than an ordinal ranking of the efficiency of different organizational arrangements. We would like to know how much we lose by going from the best to the next best... Unless we can find good ways to quantify the magnitude of the differences in costs of alternative institutional arrangements it will be very difficult to do the necessary tradeoffs even when we convince antitrust authorities or regulators that tradeoffs are appropriate. Unfortunately, it is unlikely that the data econometricians typically rely on, drawn from actual organizational choices, will reflect the ideal natural experiment to perform such computations. (1991:81–82)

Joskow commends the recent attempt by Masten, Meehan, and Snyder (1991) to measure directly the costs of internal organization (though not the benefits). Their study represents an early effort to estimate actual costs and benefits of alternate institutional arrangements, rather than rank them ordinarily by reference to reduced-form estimation. This type of analysis, were it to become more

15. For a discussion of Schwinn and GTE-Sylvania, see the recent exchange between Williamson (1993) and Posner (1993).
common, could be used in making the comparative judgments to which Joskow refers.

TCE also has direct implications for many other contracting practices and regulations, though it does not yet appear to have had any influence in those areas. Barker and Chapman (1989) argue, for example, that closed-shop agreements in labor markets may serve to protect worker's job-specific training rather than to exploit a monopoly position. They attack New Zealand's "blanket coverage clause," which effectively prohibits the closed shop, supporting their claims with arguments based on TCE. Studies of optimal contract design such as Crocker and Reynolds's (1993) examination of Air Force procurement contracts are also relevant as a guide to public policy toward government purchases of goods and services. Other contracts between government agencies and private firms, such as franchise contracts for the provision of public utilities (like cable TV), can be evaluated using TCE reasoning. TCE also points out how the potential for opportunism by the state affects private incentives to make specific investments (Levy and Spiller, 1994). This is particularly important for economic and political reform in the former communist countries, where the need to provide incentives for private investment is paramount.

6. Summary and Conclusions

Empirical studies of transaction cost economizing cover a broad range of phenomena. Topics range from problems traditionally within the domain of industrial organization to those more frequently addressed by sociologists, business strategists, or organization theorists. The articles discussed above generally support the TCE predictions. Studies that examine the make-or-buy decision and the structure of long-term contracts, in particular, overwhelmingly confirm transaction cost economic predictions.

In each area, however, there are also results that contradict fundamental and important TCE arguments, and others that provide only weak or tangential support for the framework. The purpose of this article is not to address each of these contradictions or to assess rigorously the validity of individual studies. It suffices to say that researchers in TCE must address empirical challenges to the theory. Taken as a whole, the body of empirical research in TCE shows that a good deal of economic activity aligns with transactions in the manner predicted by the theory. As Joskow concludes, the growing body of empirical work in transaction cost economics is in many ways in "much better shape than much of the empirical work in industrial organization generally" (1991: 47). Nonetheless, much remains to be done, both in applying those approaches already developed to additional data, and in further refining and developing the methods used to test transaction cost hypotheses.

Appendix

This appendix lists a large sample of empirical studies that either directly test transaction cost economic hypotheses or have important implications for TCE. While the list is not exhaustive, we believe it is reasonably comprehensive. The arrangement is by topic and, within that, mainly by type of study. Clearly,
some articles could easily fit within more than one category, while others fit only uneasily into any category at all. Our purpose is simply to provide a picture of the scale and scope of empirical TCE research, and to provide readers with a general road map through the literature. Full citations for each article are in the references.

A.1 Comparative Contracting

This section, which contains the bulk of the empirical work on TCE, focuses on the choice between internal and external procurement of components and supplies and between internal and external distribution and marketing of final products.

A.1.1 Vertical Integration. Tests of the effects of transaction costs on vertical integration cover a broad spectrum of industries and methods. They also focus on different sources of transaction costs. Some studies focus on asset specificity, some on uncertainty or small-numbers exchange conditions, and some on a combination of these variables. Much of the work listed below consists of focused single-industry studies, though several studies test TCE hypotheses using multi-industry data.


*Studies using multi-industry data:* Butler and Carney (1983); Caves and Bradburd (1988); Harrigan (1986); Levy (1985); MacDonald (1985); Macmillan, Hambrick, and Pennings (1986); Mahoney (1992); Weiss (1992).

*Forward integration into marketing and distribution:* Anderson (1985); Anderson and Schmittein (1984); John and Weitz (1988); Lilien (1979); Muris, Scheffman, and Spiller (1992); Noordewier, John, and Nevin (1990).


A.1.2 Complex Contracting and "Hybrid Modes." For a large class of transactions, simple market exchange is not feasible, yet the transactions are not vertically integrated. The decision not to integrate may be due to regulatory restrictions, or to the fact that relatively efficient arrangements short of unified ownership can be set up to govern the transaction. Examples include complex contracts with reciprocity agreements, offsetting specific investments, or other safeguards. Another type of arrangement, closer to integration along the market–hierarchy continuum, is equity linkage between firms. This section lists research into how transaction costs determine the structure of exchange relations that lie between market and hierarchy ("hybrid modes"). Included are studies of both formal contracts and informal agreements.

*Long-term commercial contracts:* Crocker and Masten (1988); Crocker and Reynolds (1993); DeCanio and Frech (1993); Eccles (1981); Heide and John (1988); Joskow (1985, 1987); Leffler and Rucker (1991); Masten and Crocker...
(1985); Mulherin (1986); Pirrong (1993); Pisano (1990); Pittman (1991); Walker and Poppo (1991).

Franchising and franchise bidding: Dnes (1992); Kaufmann and Lafontaine (1994); Lafontaine (1993); Norton (1989); Prager (1990); Williamson (1976); Zupan (1989a, 1989b).


Land tenure agreements: Alston, Datta, and Nugent (1984); Alston and Higgs (1982); Datta, O'Hara, and Nugent (1986); Roumasset and Uy (1980).

Informal exchange relations: Acheson (1985); Brinig (1990); Ellickson (1989); Gifford (1993); Jones and Pustay (1988); Palay (1984, 1985); Wilson (1980).


A.1.3 Price Adjustment in Long-Term Contracts. TCE predicts that long-term contracts should be designed to protect fixed investments and to limit the extent to which either side can benefit from market changes not anticipated at the time of bargaining. These articles examine more carefully the adjustment mechanisms for price and/or quantity in long-term contracts: Crocker and Masten (1991); Crocker and Reynolds (1993); Goldberg and Erickson (1987); Hubbard and Weiner (1986, 1991); Joskow (1988b, 1990).

A.1.4 Multinational Corporations and the Structure of Foreign Trade. The complexities of transacting across national boundaries include ownership restrictions, government participation, and a variety of political factors. Exchange agreements often take apparently peculiar forms. These studies use transaction cost analysis to explain the structure of multinational corporations and foreign trade agreements: Anderson and Coughlan (1987); Davidson and McFetridge (1984, 1985); Gatignon and Anderson (1988); Hallwood (1990); Hennart (1989); Klein, Frazer, and Roth (1990); Murtha (1993); Teece (1977); Yarbrough and Yarbrough (1987a).

A.2 Effects of Organizational Form

The internal structure of firms, and the effects of internal organization on firm performance, have been subject to relatively few empirical TCE studies. One possible reason is that data for internal transactions are difficult to obtain. Another may be that the internal workings of a firm comprise a complex system, in which social, managerial, economic, and technological forces all operate. Nonetheless, several important studies have been carried out.

A.2.1 Effects of Vertical Integration. This section lists research into the effects of vertical integration on several types of firm performance. The range of work includes transaction cost analyses of governance cost savings, of changes in activities that the firm can undertake, and of changes in capital costs due to vertical integration.


A.2.2 Comparative Studies of Organizational Form. The comparative performance of firms adopting different methods of internal organization has received considerable attention. Most often, this consists of attempts to test the "M-form hypothesis" associated with Chandler and Williamson. This hypothesis states that firms adopting a particular internal governance structure—namely, the multidivisional or M-form structure—will outperform firms organized either as traditional unitary (U-form) structures or as holding companies (H-form). The evidence on relative M-form performance is decidedly mixed, especially when comparing results from the U.S. and the U.K. These comparative studies include: Armour and Teece (1978); Blackwell, Brickley, and Weisbach (1994); Burton and Obel (1988); Butler (1983); Cable and Dirrheimer (1983); Cable and Yasuki (1985); Dwyer and Oh (1988); Harris (1983); Hill (1988); Hill and Pickering (1986); Jones (1987); Shelanski (1993); Steer and Cable (1978); Teece (1981); Thompson (1981); Williamson (1981).

A.2.3 Firm Ownership and Governance. The application of transaction cost principles to corporate governance and ownership has led to several theoretical predictions (Williamson, 1985: chap. 12). These theories so far have received limited, but increasing, empirical attention. Such studies include: Balakrishnan and Fox (1993); Baysinger and Butler (1985); Baysinger and Zardkoohi (1986); Brickley and James (1987); Romano (1985).

References


Journal 75–94.


