The Economics of Organization –

with Reference to Transaction Cost Economics

and More Generally

Oliver E. Williamson

University of California, Berkeley

April 25, 2009

The Economics of Organization –

With Reference to Transaction Cost Economics

And More Generally

Oliver E. Williamson*

The study of economic organization is a huge subject and benefits from being examined from different perspectives, both within and among disciplines. For economists, if not more generally, organization matters if and as it is made susceptible to analysis. This paper describes the general framework out of which transaction cost economics works with special emphasis on little noted but consequential forks in the road where transaction cost economics takes one branch while other studies of economic organization take others.

The general framework begins with what I take to be the first move in the study of economic organization: choose the lens through which the phenomena to be studied are to be examined. For transaction cost economics, this is the lens of contract/governance. Section 2 discusses the phenomena of interest, namely the study of firm and market organization. Sections 3 and 4 deal with the choosing of the exemplar (focal) transaction and the analysis thereof. Applications to phenomena other than the focal transaction and to puzzles and challenges arise in conjunction with the study of organization are sketched in Section 5. The conclusions and an appendix on pragmatic methodology follow.

The forks in the road to which I call special attention are these: the aforementioned choices of (1) a lens and (2) a focal transaction, (3) the description of human actors and, more generally, (4) of the disciplinary framework, (5) naming the unit of analysis (to include the dimensionalization thereof (or not)), (6) naming the main purposes served by organization and (7) describing firm and market mechanisms (interfaces) as these relate thereto, (8) revisiting (or

not) nagging puzzles of organization, (9) the relation of the project to the four precepts of pragmatic methodology, and (10) confronting unmet challenges, to include full formalism.

1. The Lens

James Buchanan avers that "mutuality of advantage from voluntary exchange is … the most fundamental of all understandings in economics" (2001, p. 29). He further contends that this fundamental understanding is better realized by examining economics through the underused lens of contract rather than the overused lens of choice (Buchanan, 1975) – where, by the latter, he means the neoclassical resource allocation paradigm. Indeed, Buchanan (1975, p. 225) holds that economics as a discipline went "wrong" in its preoccupation with the science of choice and the optimization apparatus associated therewith. Wrong or not, the parallel development of a science of contract was slow to develop and is still a work-in-progress.

As perceived by Buchanan, the principal needs for a science of contract were in the field of public finance and took the form of public ordering: "Politics is a structure of complex exchange among individuals, a structure within which persons seek to secure collectively their own privately defined objectives that cannot be efficiently secured through simple market exchanges" (1987, p. 296). Thinking contractually in the public ordering domain leads to focus on the rules of the game. Constitutional economics issues are posed (Buchanan and Tullock, 1962; Brennan and Buchanan, 1985).

Whatever the rules of the game, the lens of contract is also usefully brought to bear on the play of the game. This latter is what I refer to as private ordering, which entails efforts by the immediate parties to a transaction to align incentives and to craft governance structures that are better attuned to their exchange needs.

Figure 1 sets out the main distinctions (Williamson, 2002). The initial divide is between the science of choice (orthodoxy) and the science of contract. The latter divides into public ordering (constitutional economics) and private ordering parts, where the second is split into two

related branches.[1] One branch concentrates on the study of ex ante incentive alignment. The second branch deals with ex post governance, with emphasis on “good order and workable

arrangements" (Fuller, 1954, p. 477), to include both spontaneous order in the market and purposeful order, if and as needed, in both markets and hierarchies.

The study of governance is an interdisciplinary undertaking in which law, economics, and organization theory are joined. Economics is the main analytical engine, but both law and organization theory play important roles. Indeed, but for my study of organization theory in the interdisciplinary PhD program at Carnegie, I have grave doubts that I would have undertaken the study of transaction cost economics as described herein.[2]

If and as economics moves beyond prices and output, supply and demand to deal with the modern corporation and to explain non-standard and unfamiliar contractual practices, all-purpose reliance on the resource allocation paradigm is, to say the least, strained. As Harold Demsetz observed, it is a "mistake to confuse the firm of [neoclassical] economic theory with its real world namesake. The chief mission of neoclassical economics is to understand how the price system coordinates the use of resources, not the inner workings of real firms" (1983, p. 377; emphasis added). Similar considerations apply to markets, where the mechanisms of simple market exchange and of complex market exchange differ consequentially in ways and for reasons that need to be uncovered and explicated.

David Kreps captures the spirit of what is different about the transaction cost economics enterprise by noting that the firm is akin to the agent (consumer) in textbook economics but is of the genus of the market in transaction cost economics. Thus (Kreps, 1990, p. 96):

> The [neoclassical] firm is like individual agents in textbook economics … Agents have utility functions, firms have a profit motive; agents have consumption sets, firms have production possibility sets. But in transaction-cost economics, firms are more like markets – both are arenas within which the individual can transact.

Plainly, the lens of contract/governance and the orthodox lens of choice work out of different conceptual setups, on which account it should come as no surprise that these two employ different apparatus and describe and interpret firm and market organization differently.[3]

2. The Lapse

Ronald Coase pushed the basic assumption of standard economic theory that transaction costs were zero to completion, thereby to display a troublesome implication: economic organization was indeterminate. If firm and market are "alternative methods of coordinating production" (Coase, 1937, p. 388), then the decision to use one mode rather than the other should be derived. Accordingly, economists needed (1937, p. 389):

> … to bridge what appears to be a gap in [standard] economic theory between the assumption (made for some purposes) that resources are allocated by means of the price mechanism and the assumption (made for other purposes) that this allocation is dependent on the entrepreneur-coordinator. We have to explain the basis on which, in practice, this choice between alternatives is effected.

Coase thereafter averred that "the main reason why it is profitable to establish a firm would seem to be that there is a cost of using the price mechanism," the most obvious of which is "that of discovering what the relevant prices are" (1937, p. 390). Coase also described the different mechanisms through which firm and market work: outsourcing is accomplished by an arm's length contract between buyer and supplier whereas an employment contract is used if a firm produces to its own needs. The latter entails creating an authority relation according to which the employee "agrees to obey the directions of an entrepreneur within certain limits" (p. 391; emphasis in original).

Not only was the challenge posed by Coase (1937) little noted at the time, but Coase would later describe the 1937 paper as "much cited and little used" (1972, p. 63). It was much cited because it posed an unanswered puzzle; but it was little used because, once the black

boxes of firm and market organization were opened, positive transaction costs turned out to be "everywhere." Since any outcome whatsoever could be "explained" by invoking some type of transaction cost to fit the need, transaction cost reasoning earned a "well-deserved bad name" (Fischer, 1977, p. 322).

Coase's 1960 paper on "The Problem of Social Cost" served to deepen the puzzle. The "trick" that Coase introduced was to push the logic of zero transaction costs to completion. With reference to externalities, the effect was shocking: externalities vanished – because the parties would costlessly bargain to the same efficient result whichever way property rights were initially assigned. Technological nonseparability aside, firm and market were equally efficacious modes of organization.

Kenneth Arrow's examination of "The Organization of Economic Activity" (1969) further made the case for positive transaction cost constructions with his argument that (1969, p. 48):

> … market failure is not absolute; it is better to consider a broader category, that of transaction costs, which in general impede and in particular cases completely block the formation of markets. It is usually though not always emphasized that transaction costs are costs of running the economic system. An incentive for vertical integration is replacement of the costs of buying and selling on the market by the costs of intrafirm transfers; the existence of vertical integration may suggest that the costs of operating competitive markets are not zero, as is usually assumed in our theoretical analysis.

3. Applying the Lens of Contract to Economic Organization

As indicated, economics is the mother discipline for transaction cost economics to which, as needed, aspects of organization theory and the law are selectively introduced. The main ways in which organization theory bears on the enterprise are in describing the attributes of

human actors[8] and in revealing the importance of key intertemporal regularities. With reference to the first, all complex contracts are incomplete by reason of bounded rationality[5] (as relieved, however, by making provision for "feasible foresight")[6] and self-interest manifests itself as defection from the spirit of cooperation (opportunism) when the stakes are great.[7] Moreover, albeit a tool, organization "nevertheless has a life of its own" (Selznick, 1949, p. 10), as a result of which significant intertemporal regularities arise that are properly included in the calculus as these have a bearing on comparative economic organization. Also, as discussed in conjunction with "scaling up" (see subsection 5.1, below), the appearance and purposes served by the condition of near-decomposability in hierarchies (Simon, 1962) is also pertinent. As for the law, the main way in which the law is introduced into the study of comparative economic organization is with reference to contract laws (plural), as against contract law (singular), as discussed in subsection 4.1 below.

These matters are all developed more extensively elsewhere (Williamson, 1991, 2002a, 2002b, 2005). My purpose here is merely to emphasize that but for selective appeal to the contiguous social sciences, transaction cost economics, as described herein, would be a very different enterprise and, I think, would bear a less veridical relation to the phenomena.

I successively discuss the precursory contributions of John R. Commons to the transaction cost economic project, the importance of choosing an exemplar (focal) transaction, and the rudiments of the governance of contractual relations as viewed through the lens of contract.

3.1 John R. Commons

Of the many good ideas that originated with Commons, none was more important to the economics of governance than his abiding interest in “going concerns.” As against the preoccupation of orthodoxy with simple market exchange, Commons observed that the continuity of an exchange relationship was often important, whereupon he reformulated the problem of economic organization as follows: “the ultimate unit of activity … must contain in

itself the three principles of conflict, mutuality, and order. This unit is a transaction" (Commons, 1932, p. 4). Commons thereafter recommended that "theories of economics center on transactions and working rules, on problems of organization, and on the … [ways] the organization of activity is … stabilized" (1950, p. 21).

Commons appealed to W. N. Hohfeld's analysis, terminology, and classification of legal relations in his efforts to introduce operational content into institutional economics. Alas, what Hohfeld provided was a vocabulary of "jural opposites" and "jural correlatives" that are classified as to rights, privilege, power, and immunity (Commons, 1924, pp. 91-134). These concepts provided an elaborate taxonomy, but a theory replete with refutable implications never materialized. As with the legal realism movement in the United States (Schlegel, 1979, p. 459), so with older style institutional economics: for lack of operationalization, both movements "ran themselves into the sand."

Although lack of a positive research agenda spelled the demise for older style institutional economics in the United States (Stigler, 1983, p. 170), this fate does not, contrary to Coase, imply an absence of good ideas (1983, p. 230). Indeed, transaction cost economics subscribes to both parts of Commons' formulation of the problem of economic organization. Not only is the transaction the basic unit of analysis, but governance is the means by which to infuse order, hereby to mitigate conflict and realize mutual gains.

3.2 The exemplar (focal) transaction

As Thomas Kuhn observed, concentrating attention on a "narrow area of trouble" is often instructive (1970, p. 88). What he subsequently describes as "exemplars" serve this purpose (1970, p. 187). Ideally, examining the focal (exemplar) transaction with respect to the aforementioned governance (or other) purpose will not only reveal key attributes and mechanisms that explain observed regularities but will also permit related phenomena to be interpreted as variations on a theme. Among the candidate contractual alternatives from which

to choose the focal transaction are the employment relation, externalities, the intermediate product market transaction, the finance transaction, regulation, and the list goes on.

The two transactions that have been most often treated as focal are the employment relation (Coase, 1937; Simon, 1951; Hart and Moore, 2008; Hart, 2008) and the intermediate product market transaction (Williamson, 1971, 1979; Klein, Crawford, and Alchian, 1978). As between these two, the advantages that I associate with taking the intermediate product market transaction rather than the employment relation to be focal are these: (1) it is simpler; (2) the two parties to the intermediate product market transaction are more on a parity; (3) the key attributes of transactions and governance structures come more readily to the fore (especially as the contractual condition of bilateral dependency, by reason of asset specificity (in its various forms) had hitherto been neglected); (4) the same is true of the interface differences between make and buy, in that attention is directed not to the contractual interface between <u>individuals</u> but to the contractual interface <u>between successive stages of production</u>; (5) this last has a bearing on scaling up from individual transactions to a series of technologically separable transactions; and (6) the regularities that arise in conjunction with intermediate product market transaction also appear in many other commercial transactions, which are interpreted as variations on a theme. Taking the intermediate product market to be focal was a fortuitous choice.

Having chosen a focal transaction for studying economic organization, it is further instructive to ask what main purpose is being served. Transaction cost economizing is the obvious response, but, as previously indicated, that manifests itself in many ways. Both the economist Friedrich Hayek (1945) and the organization theorist Chester Barnard (1938) declared that adaptation is the central purpose of economic organization. Such agreement notwithstanding, there were interesting differences. Thus Hayek (1945, pp. 526-527) focused on the adaptations of autonomous economic actors who adjust spontaneously to changes in the market, mainly as signaled by changes in relative prices. The marvel of the market resides in

"how little the individual participants need to know to be able to take the right action" (Hayek, 1945, p. 527). By contrast, Barnard featured coordinated adaptation among economic actors working through deep knowledge and the use of administration. In his view, the marvel of hierarchy is that coordinated adaptation is accomplished not spontaneously, but in a "conscious, deliberate, purposeful" way (1938, p. 4).

Because a high-performance economic system will display adaptive properties of both kinds, the problem of economic organization is properly posed not in the old ideological way as a choice between markets or hierarchies, but rather in a pragmatic way as the efficient deployment of markets and hierarchies.

3.3 Implementation[8]

A predictive theory of economic organization will explicate the factors that are responsible for the adaptive differences between markets and hierarchies and how and why transactions differ in their adaptive needs. Transaction cost economics is implemented by (1) identifying the critical dimensions with respect to which transactions differ – namely, asset specificity (which can take different forms to which different organizational ramifications accrue), disturbances (to which adaptations are required), and frequency (which has a bearing on both the cost-effectiveness of internal dispute resolution and reputation effects in the market), (2) identifying the critical dimensions with respect to which governance structures (especially markets and hierarchies) differ in their adaptive capacities – namely incentive intensity, administrative control, and contract law regime, (3) deriving refutable implications by invoking the discriminating alignment hypothesis – to wit, transactions, which differ in their attributes are aligned with governance structures, which differ in their cost and competence, thereby to effect a transaction cost economizing result, and (4) inviting empirical testing.

These matters are all discussed elsewhere (1979, 1985, 1991, 1996). I merely add brief comments on two crucial aspects of the project: asset specificity and empirical testing.

Asset specificity. I have grave doubts that asset specificity would have surfaced as a crucial concept had not the intermediate product market transaction been made focal. Given that the focal contract was between firms, there was much less reason to feature differential risk aversion, differential competence, different numbers of players (as with one supplier and many buyers (as in the final product market) or many suppliers and one buyer (as in the labor market), or to appeal to language from the insurance literature (on adverse selection and moral hazard). Instead, the make-or-buy transaction was of a bilateral kind and issues of complex contracting could be revisited anew, as a result of which new contractual problems were uncovered to which comparative contractual reasoning could be applied.

Considerations of bounded rationality ruled out both ex ante contractual completeness and ex post costless Coasian renegotiation. Inasmuch as simple market contracting worked well for many transactions, evidently the problems resided in "more complex" contracting. Non-generic transactions were possibility, but not in the customary sense where product differentiation could pose a problem with final product market transactions. The transactions in question being intermediate product market transactions between firms, there must be something else, possibly something more basic? Bilateral dependency by reason of asset specificity (often as this evolved during contract execution and at the contract renewal interval – wherein a large numbers competition at the outset underwent a Fundamental Transformation and became small numbers supply relation thereafter) was an intriguing possibility. That is a transaction cost economics insight and construction (Williamson 1971, 1979; Klein, Crawford, and Alchian, 1978).

Empirical testing. The aforementioned discriminating alignment hypothesis sounds all well and good in principle, but can it be tested and are the data corroborative? Involving, as it does, microanalytic features of transactions that were not routinely reported in accounting statements or in census reports or in other data bases, could this be an interesting but ultimately untestable hypothesis?

Empirical tests of transaction cost economics began to appear in the 1980s and have since grown exponentially. Thus "despite what almost 30 years ago may have appeared to be insurmountable obstacles to acquiring the relevant data [of a microanalytic kind], today transaction cost economics stands on a remarkably broad empirical foundation" (Geyskins, Steinkamp, and Kuman, 2006). Surveys of the empirical literature – beginning with Paul Joskow (1988); the most recent by Jeffrey Macher and Barak Richman (2008) – record that published empirical studies in this area now number over 1000. This empirical work has contributed significantly to the growing acceptance of transaction cost economics (Whinston, 2001, 2003) and to a growing interest in organization. Thus although some skeptics doubt that a "science of organization" replete with refutable implications for which empirical tests can be conducted is within reach(Flyvbjerg, 2001), Barnard's aspirations for such a science (1938, pp. 9, 270) are being realized.

4. Gap Filling

Three features of transaction cost economics to which little attention has been directed are contract laws (plural), the criteria by which viable and nonviable modes of governance are distinguished, and the different interface mechanisms through which markets and hierarchies work.

4.1 Contract laws (plural)

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. What has been referred to as the "legal centralism" tradition 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).

Given such an orientation, (nearly costless) dispute settlement becomes the province of the lawyers. That being an analytical convenience, why should economists hold otherwise?

Still, there were dissenters. Karl Llewellyn took exception with this legal rules contracting tradition and introduced the concept of "contract as framework" (1931, pp. 736-737):

> [T]he major importance of legal contract is to provide a framework for well-nigh every type of group organization and for well-nigh every type of passing or permanent relation between individuals and groups … − a framework highly adjustable, a framework which almost never accurately indicates real working relations, but which affords a rough indication around which such relations vary, an occasional guide in cases of doubt, and a norm of ultimate appeal when the relations cease in fact to work.

Marc Galanter likewise emphasized the importance of “private ordering." Contrary to the legal centralist approach, most disputes including many that under current rules could be brought to a court, are resolved by avoidance, self-help, and the like (1981, p. 2). This is because 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” (1981, p. 4).

In addition to the foregoing distinction of contract as legal rules and contract as framework, "forbearance law" is a third type of dispute settlement mechanism (Williamson, 1991). This last is applicable to disputes within, rather than between, firms. Thus whereas courts routinely grant standing in contracts between firms should there be disputes over prices, the damages to be ascribed to delays, failures of quality, and the like, courts will refuse to hear disputes between one internal division and another over identical technical issues. The upshot is that hierarchy becomes its own court of ultimate appeal. This is consequential for interface management (as discussed below).

4.2 Viable modes

I will take as given that alternative modes of governance are described by the three attributes referred to in Section 3.3, namely incentive intensity, administrative command and control, and contract law regime. I further assume that each attribute can take on either of two values, much (+) or nil (0), such that there are $2^3 = 8$ possible combinations. As among the eight combinations shown in Table 1, which of these are internally consistent? Which define markets and hierarchies?

Viable and nonviable commercial modes of governance are distinguished as follows: (1) because high-powered incentives are compromised by administrative command and control and vice versa, these two should take on opposite values; and (2) inasmuch as the contract law regime performs a supporting role, the operative contract law is that which is supportive of whichever operating attribute takes on the (+) value. Application of criterion (1) eliminates modes I, II, VII, and VIII. Criterion (2) eliminates modes IV and V. The internally consistent syndromes are thus III (high-powered incentives, negligible control at the interface, and a legal rules contract law regime) and VI (low-powered incentives, managerial coordination at the interface, and a forbearance law conflict resolution regime). These two correspond to market and hierarchy, respectively.

4.3 Interface differences

Upon reflection, the autonomous adaptations to which Hayek referred are "obviously" supported by high-powered incentives (and a lack of administration) and the coordinated adaptations emphasized by Barnard are "obviously" supported by administration (and much weaker incentives). A legal rules contracting regime supports the former; a forbearance law contracting regime supports the latter. Presumably the mechanisms that operate at the interface in support of make (hierarchy) and buy (market) reflect these differences. But can we be more explicit about how these mechanisms differ?

A heuristic display of the key interface differences is shown in Figure 2, where market mediated governance is shown in the top panel and hierarchically mediated exchange is shown in the bottom.

Very briefly, if the interface is market mediated, then the supplier delivers a good or service as prescribed by the contract and receives a fixed payment in return; changes in the terms or implementation of the agreement need to be renegotiated; and irreconcilable disputes are presented to the courts where legal rules are applied and settlement is reached by the award of money damages. The schematic shown in the top panel of Figure 2 applies. Given ownership autonomy, each stage appropriates its own net receipts and each adapts to price signals in the market on its own motion. Market mediated exchange obtains. Autonomous adaptations are served in the process.

Hierarchically mediated exchange is shown in the bottom panel. Inasmuch as the object here is to promote coordinated adaptations, independent ownership gives way to unified ownership of a symmetrical (rather than directional) kind. A new economic actor – the peak coordinator – is created to manage the interface, where the job of the peak coordinator is to (1) delegate operating responsibilities to each stage, (2) effect coordination at the interface as the need for coordinated adaptations arise (also, to effect convergent expectations as needed for investment purposes), (3) settle disputes, and (4) implement settling up (mainly of a cost-plus kind, as determined by the application of internal accounting routines as verified by internal auditing). Hierarchy, so described, is akin to double-feedback, as described by W. Ross Ashby (1960). Thus whereas local disturbances elicit local responses (often by the application of routines) within each operating stage, more consequential disturbances that require coordinated adaptations between stages are accomplished with the participation of the peak coordinator in the secondary (strategic) feedback loop. Management by exception is thus the peak coordinator's job. Coordinated adaptations at the exchange interface is what distinguishes hierarchy. Markets and hierarchies thus differ not merely in attribute to the differences shown in

Table 1 but also because these syndromes are implemented through the interfaces shown in Figure 2. The action is in the implementation differences.

5. Beyond the Focal Transaction

Extensions upon the foregoing include variations on a theme, re-examining longstanding puzzles, and unmet challenges.

5.1 Variations on a theme

As discussed in Appendix I, more fruitful theories will permit other-than-focal transactions to be interpreted as variations on a theme. The regulation of natural monopoly, the governance of corporate finance, the separation of ownership from control, and "rainmakers" are four examples. The condition of asset specificity is pivotal for all four.

Natural monopoly. Harold Demsetz (1968) challenged the widely held view that there is no good solution for natural monopoly but only a choice among the "three evils" named by Milton Friedman for dealing with natural monopoly: "private unregulated monopoly, private monopoly regulated by the state, and government operation" (Friedman, 1962, p. 128). Upon examining natural monopoly from a contractual perspective, Demsetz introduced a fourth alternative: franchise bidding for natural monopoly, wherein the monopoly pricing consequences of private unregulated monopoly could purportedly be avoided by using an ex ante bidding competition to award the monopoly franchise to the firm that offers to supply product on the best terms.

Demsetz described the logic of franchise bidding and offered the stamping and painting of automobile license plates as an example. Richard Posner (1970, 1972) was persuaded and subsequently argued that franchise bidding could and should be used to solve the problem of local monopoly in the cable television industry. Inasmuch as Posner described the mechanisms for implementing franchise bidding for CATV and since I had some familiarity with the issues

from my work on Mayor John Lindsay's CATV task force (1969-1970), this presented a good opportunity to put the contractual reasoning on which transaction cost rested to the test.

As I describe in Williamson (1976), franchise bidding for CATV experienced a number of unexamined complications, the most important being that nonredeployable investments in physical plant (also, possibly, in human assets) gave the initial winning bidder a significant advantage at the contract renewal interval. This was compounded by (1) the inherent limits of accounting for deciding asset valuation disputes,[9] (2) the limits of the courts for sorting these matters out, and (3) the vulnerability of the bidding process to being politicized. I concluded that (at least as of the 1970s when the technology was still in flux) franchise bidding for CATV was deeply problematic.[10] Jean-Jacques Laffont and Jean Tirole (1993, Chap. 8) would subsequently conclude similarly.

Regulation and deregulation have become lively topics in the years since. The challenge is to address them in an objective way, to which Joskow's reservations with the deregulation of electric power in California are pertinent (2002, pp. 527-528; emphasis added):

> Electricity sector reforms necessarily must be built upon an infrastructure made up of long-lived historical sunk investments made over past decades. The investments were made within an institutional environment which did not contemplate the kinds of opportunism, coordination, and market power problems that can emerge in a decentralized system with many independent firms owning and operating different pieces of an industry. Market power problems, network congestion management, and coordination problems arising from restructuring of the existing configuration of assets should be expected and their existence carefully identified *ex ante* as an integral part of the design and implementation of liberalization reforms. Accordingly, electricity restructuring programs need to consciously and carefully include transition mechanisms to mitigate these problems until investments in new generating and transmission capacity can be

made to move the system toward a new asset configuration that is less susceptible to them. These mechanisms will include contracts to deal with local market power problems, carefully structured congestion management protocols and rules for injecting and withdrawing power from the grid, and transitional contracts between generators and those entities responsible for procuring power for retail consumers that both protect consumers from exploitation and diminish incentives that generators may have to exercise market power. These transition mechanisms must be put in place at the outset of the restructuring program because they are difficult to implement *ex post*, after problems emerge, since incumbent interests are likely to have a strong stake in preserving the status quo.

To be sure, there are some cases for which deregulation is an easy call and easy to implement. Electric power is an industry where the issues need to be worked through in a "modest, slow, molecular, definitive way."

Corporate finance. Examining the corporate finance transaction in transaction cost economics terms is accomplished by treating individual investment projects as transactions and interpreting debt and equity not merely as modes of finance but also as modes of governance, where debt is rules based (akin to the market) and equity is a more flexible and discretionary mode of finance (akin to hierarchy). The same regularities that characterize the intermediate product market transaction reappear: whereas
debt is well-suited to generic investments, a shift to equity appears as the investment projects become more specific (less redeployable).

Rules based debt requires the debtor to make stipulated interest payments, meet certain liquidity tests, repay the principal at the loan expiration date, and, in the event of default, the debt-holders have a pre-emptive claim against the assets in question. By contrast, equity displays the following financial and governance properties: (1) it bears a residual-claimant status to the firm in both earnings and asset-liquidation respects; (2) it contracts for the duration

of the life of the firm; and (3) a board of directors is created and awarded to equity that (a) is elected by the pro rata votes of those who hold tradable shares, (b) has the power to replace the management, (c) decides on management compensation, (d) has access to internal performance measures on a timely basis, (e) can authorize audits in depth for special follow-up purposes, (f) is apprised of important investment and operating proposals before they are implemented, and (g) in other respects bears what Eugene Fama and Michael Jensen (1983) refer to as a decision review and monitoring relation to the firm's management. The board of directors thus arises endogenously and serves as a credible commitment for equity investors, the effect of which is to reduce the cost of capital for projects that involve limited redeployability. Not only do the added controls to which equity has access provide added assurance, but equity is also more forgiving than debt. Efforts are therefore made to work things out and realize adaptive benefits under equity finance that would be sacrificed under rules-based debt financing when disturbances push the parties into a maladapted state of affairs.

Debt and equity are more, therefore, than different modes of finance. They are also different modes of governance, the choice between which is consequential.[11]

Ownership and control. It is common to observe that ownership confers control. The relation, however, between ownership and control can be and often is attenuated. Two examples are discussed here: the use of inside contractors and corporate governance.

John Buttrick has described the "inside contracting" system as follows (1952, pp. 201-202):

> Under the system of inside contracting, the management of a firm provided floor space and machinery, supplied raw material and working capital, and arranged for the sale of the final product. The gap between raw material and finished product, however, was filled not by paid employees arranged in [a] descending hierarchy … but by [inside] contractors, to whom the production job was

delegated. They hired their own employees, supervised the work process, and received a piece rate from the company.

This is an imaginative system whereby the capitalist takes responsibility for financing and constructing the plant, providing the raw material, and selling the final product while the inside contractor is responsible for hiring the workers and supervising the work. Indeed, it can be thought of as a variant of franchise bidding, in that the right to administer successive stages of production is auctioned off to a series of inside contractors based on the piece rate that each bids. Unlike hired supervisors, inside contractors appropriate the difference between the piece rate and the labor costs incurred, thereby to experience added incentive intensity.

This effort to inject high powered incentives into the production process induced inside contractors to behave strategically – which resulted in added quality defects at successive exchange interfaces, malutilization of equipment, innovation biases in favor of labor as against capital, and strategic timing of innovative changes awaiting contract renewal. Also, by reason of the acquisition of deep knowledge through experience, bidding advantages accrued to incumbents at the contract renewal interval. (What is furthermore noteworthy is that the student of transaction cost economics, had he been consulted, could have predicted many of these from the outset.)

The separation of ownership from control in the modern corporation has been a matter of grave concern since Adolph Berle and Gardiner Means published their famous book on The Modern Corporation and Private Property in 1932 (and Adam Smith had similar concerns with joint stock companies much earlier). The problem ascribed to the modern corporation is that full-time managers have captured control from absentee owners. This being contrary to the intent, control should be restored to these owners. But what does that entail?

My examination of corporate governance (Williamson, 2008) discloses that this is a case where there are no good choices: admonishing the board to serve as a diligent monitor is pointless if the board lacks the requisite information; expecting the management to disclose the

relevant information and display the ramifications in an unbiased way collides with the realities of managerial discretion; providing the board with its own staff to uncover and interpret the data is costly, will never close the information gap, and invites conflict; and the active participation of board members that is often observed in leveraged buyouts and start-ups with venture capital financing is not a realistic option in ongoing, mature corporations.

Where delegation is the only realistic option, the separation of ownership from control should be accepted as a given and attention focused on ways by which to infuse integrity into the delegation process.[12]

Rainmakers. The Oxford International Dictionary defines rainmaker as a "medicine man who uses incantations and magic rituals for the purpose of producing rain." Interestingly, the term has come to be applied more generally to individuals who are perceived to generate considerable income for the firms that employ them – especially lawyers in law firms, fund managers in financial firms, and some consultants.

That appears to be far removed from the intermediate product market transaction. The focal transaction nevertheless relates to commercial rainmakers in two respects. The first and most important is with reference to the absence of asset specificity. The second is to the "magical powers" that are ascribed to some rainmakers.

The main types of asset specificity that arise in conjunction with the intermediate product market transaction are physical assets of specialized design for the use of a particular client, site specificity, where two stages are located cheek-by-jowl (to reduce transportation and inventory costs; sometimes in the service of thermal economies), dedicated assets (where significant investments in generic plant are made in response to a specific long-term contract, the output of which could be sold only at distress prices if the order were to be prematurely terminated), and human asset specificity, where the human assets acquire deep knowledge from learning-by-doing, possibly to include a specialized vocabulary (code) for communicating and through associational gains.

Human asset specificity differs from physical, site, and dedicated assets in that the latter three investments are all owned by the firm whereas (by reason of laws prohibiting servitude) human assets are not.[13] If, therefore, a buyer (client) is not significantly tied to the supply firm by physical, site, or dedicated assets, an unusually productive employee (and his team) could take the clients with them should they resign. (Such a leader and his team are especially mobile if the human asset skills that they acquire during their employment are not specific to the employer.)

Transaction cost reasoning thus interprets rainmakers as individuals (teams) that possess the following attributes: (1) the individual (team) is perceived to be exceptionally competent and has little human specific capital in the firm; (2) the firm is a professional shell with little physical, site specific, or dedicated assets in supplying services to clients; and (3) the rainmakers, the incumbent firm for which the rainmakers work, and their rivals can all make a good estimate of the stream of earnings that each rainmaker group generates. Rainmakers, in such circumstances, can demand to be paid a large fraction of their respective earnings streams under the credible threat they can leave, take their clients with them, and be paid "full value" elsewhere. Settling up by the payment of large bonuses at the end of the accounting year is thus understandably observed for rainmakers who have developed reputations for excellence with their clients (and interested potential clients).

But there are caveats. The basis for rainmaker claims may vary from objective to mystical. Thus whereas some claims may be directly observable (as where clients observe the court room performance of a lawyer and his team) others may be inferred from reports that may be real or contrived. Real reports of investment successes, moreover, do not distinguish between success that is due to deep knowledge and good practice and that which is explained by a run of good luck. So clients may be more cautious about ascribing causality.

More interesting are contrived reports, as where the fund manager owns the firm and plays his cards close to his chest – by refusing to reveal the portfolio that clients buy and the

mechanisms by which the fund is operated. Bernard Madoff appears to have been a master of such a Ponzi scheme in selling to his "favored" (naïve) clients. Blatant opportunism in combination with Phineas Taylor Barnum's confidence that "there's a sucker born every minute" explain this last class of activity. [14]

To be sure, the foregoing could be dismissed as an exercise in common sense. Any concerted effort to understand rainmaking would quickly reveal these regularities. But how then to explain the widespread outrage with rainmaker compensation? I submit that the pervasive role played by asset specificity is not widely appreciated. The preconditions for rainmakers are the absence of transaction specific assets of physical, site, and dedicated kinds coupled with actual or putative responsibility for success, which comes readily to the fore only to those with a "prepared mind."

I conjecture, moreover, that the incentives of rainmakers to expend entrepreneurial energies would be attenuated by insisting that they be paid for "long run" performance (five years) rather than on an annual basis. The reasons are several. First and foremost, this puts the firm in a strong negotiating position with those who propose to leave prior to the expiration date, hence acts as a mobility barrier: the rainmaker is now tied to the firm. Second, there will be a loss of value if holding the team together becomes more difficult under these constraints. And third, rainmakers could become concerned that accounting or other business practices will be changed to their disadvantage, hence have less confidence in the process. This is not to say that to base bonuses on the current year's income stream is best. Looking ahead, however, is vital in assessing alternative compensation schemes.

3.2 Puzzles

Among the puzzles that beset the study of economic organization to which transaction cost economics has been applied are (1) limits to firm size, (2) scaling up, and (3) the criterion by which to assess "failure." Consider each.

Limits to firm size. The limits to firm size problem was noted early by Frank Knight (1921) and Ronald Coase (1937) and more recently by Tracy Lewis (1983). Of these three, Lewis formulates it the most succinctly. As he puts it, since an established firm can always "use the input exactly as the new entrant would have used it … [and can furthermore] improve on this by coordinating production from his new and existing inputs," the large firm will always realize greater value (Lewis, 1983, p. 1092). Transaction cost economics queries this conclusion by examining whether the mechanisms on which Lewis implicitly relies -- replication and selective intervention – reliably deliver the described result.

Thus suppose that two successive stages of production are combined with the understanding that (1) the acquired stage will continue to behave in the same autonomous manner as it had in the pre-acquisition status except as (2) the acquiring stage intervenes, always but only, when expected net gains can be ascribed to coordinated adaptations. In that event, the combined firm can never do worse (by replication) and will sometimes do better (by selective intervention). Accordingly, more integration is always better than less. Repeated application of this logic leads to everything being organized in one large firm.

Albeit a tedious exercise microanalytic exercise (Williamson, 1985, pp. 132-156), serious incentive and information problems (of which failures of due care are an example of the first; and accounting distortions are an example of the second) beset both. More generally, the apparent attractions of combining replication with selective intervention notwithstanding, implementation problems proliferate.

Scaling up. Most theories of economic organization, transaction cost economics included, work out of toy models. The typical phenomenon of interest is the modern corporation. Does the toy model scale up?

The scaling-up assumption is often ignored (possibly out of awareness that scaling up cannot be done) or is sometimes scanted (possibly in the belief that scaling up can be accomplished easily). The influential paper by Michael Jensen and William Meckling on “Theory

of the Firm: Managerial Behavior, Agency Costs, and Capital Structure" (1976) is an exception. The authors work out of a simplified setup where an entrepreneur (100% owner-manager) sells off a fraction of the equity of the firm, as a result of which his incentive intensity is reduced and cost-effective monitoring arises as a response. What is of real interest to the authors, however, is not the entrepreneurial firm but the "modern corporation whose managers own little or no equity" (1976, p. 356). Investigating the latter was beyond the scope of their paper, but they express belief that "our approach can be applied to this case … [These issues] remain to be worked out in detail and will be included in a future paper" (1976, p. 356). Although Jensen and Meckling never produced the follow-up paper, the authors nevertheless confront the need for scaling up.

Scaling up issues relevant to the modern corporation are also posed by the theory of the firm as team production (Alchian and Demsetz, 1972), the modern property rights theory of directional integration (Grossman and Hart, 1986; Hart, 1995), and transaction cost economics. The first of these ascribes team production to a condition of technological nonseparability, as illustrated by manual freight loading: "Two men jointly lift heavy cargo into trucks. Solely by observing the total weight loaded per day, it is impossible to determine each person's marginal productivity" (1972, p. 779).

For this argument to scale up presumably implies an inclusive condition of nonseparability: all parts are continuously coordinated as a unit. However, Simon's examination of the "architecture of complexity" (1962) reveals otherwise: viable systems (of biological, physical, symbolic, and organizational kinds) are those where the composite system is made up of nearly-decomposable subsystems, within which interactions are numerous and frequent and between which they are much more attenuated.[15] This is fundamental. It is because "many complex systems have a nearly decomposable hierarchical structure … [that we are able] to understand, to describe, and even to 'see' such systems and their parts" (Simon, 1962, p. 477). Comprehensive nonseparability is the very antithesis of decomposability.

What about "directional integration"? The property rights theory of vertical integration examines the organization of successive stages A and B and contends that it matters whether stage A acquires stage B or B acquires A. Directional integration introduces unexamined ordering and asset ownership complications into scaling up (as consider a series of N related stages, for which there is no natural order, for which a series (sequence) of make-or-buy decisions needs to be made). If a coherent theory of an integrated firm emerges from such a setup that has yet to be shown.

How does the transaction cost economics setup fare in scaling up respects? Does successive application of the make-or-buy decision, as it is applied to individual transactions, scale up to describe something that approximates a multi-stage firm? Note in this connection that transaction cost economics assumes that the transactions of interest are those that take place between technologically separable stages. This is the “boundary of the firm” issue as described elsewhere (Williamson, 1985, pp. 96-98). Upon taking the technological “core” as given (possibly by reason of site specific investments), attention is focused on a series of separable make-or-buy decisions − backward, forward, and lateral – to ascertain which should be outsourced and which should be incorporated within the boundaries of the firm. So described, the firm is the inclusive set of transactions for which the decision is to make rather than buy – which does not depend on the sequencing and does implement scaling up, or at least is a promising start. Expressed in relation to Alfred Chandler's description of the multidivisional form (1962), scaling up, as herein described, can be interpreted as the coherent set of stages that make up an operating division within the modern corporation.

Failure. Avinash Dixit’s examination of The Making of Economic Policy (1996) opens with a discussion of normative public policy analysis in which the government is assumed to maximize a social welfare function. Policymaking, so described, is viewed “as a purely technical problem. The implicit assumption is that once a policy that maximizes or improves social welfare has been found and recommended, it will be implemented as designed, and the desired

effects will follow" (Dixit, 1996, pp. 8-9). This is tantamount to doing black box welfare economics in which transaction costs are assumed to be zero.

It is Dixit's judgment (and mine) that applied welfare economics, like the theory of the firm, should open up the black box and examine "the actual workings of the mechanism inside" (Dixit, 1996, p. 9) in a comparative institutional way. What I have referred to as the Remediableness Criterion is an effort to restore perspectives.

The Remediableness Criterion holds that an extant practice or mode of organization for which (1) no feasible superior alternative can be described and (2) implemented with expected net gains is (3) presumed to be efficient. Insistence on feasibility disallows hypothetical ideals. Superior feasible alternatives that cannot be implemented because the costs of implementation exceed the projected gain likewise fail the remediableness test. There is, however, a caveat: the presumption of efficiency is rebuttable. If, for example, the main obstacle in the implementation of a feasible superior alternative is politics, then the possibility that this has its origins in unacceptable (prior or current) political practices comes under scrutiny. If a feasible superior alternative could be implemented with expected net gains but for entrenched antisocial interests (e.g., discrimination against a minority), then the status quo could be condemned and the instruments of reform (e.g., civil rights) given social legitimacy – to which democratic politics must eventually accede.[16]

5.3 Challenges

Challenges of many kinds await. I mention two: disequilibrium contracting and full formalization.

Disequilibrium contracting. Transaction cost economics predominantly deals with equilibrium transactions of a mature kind. Such are to be distinguished from start-ups, especially high-technology firms, which often arise out of perceived opportunities to provide something altogether new (Shane, 2001). These latter are high-risk undertakings that combine venture capitalists with entrepreneurial, technical, and legal talent in a race to be first. High-

powered incentives apply, and real-time involvement by all of the critical actors (as managers or directors) is practiced. If and as the start-up succeeds, the big rewards are realized when the firm goes public. Thereafter, the firm progressively takes on the characteristics of a business-as-usual enterprise.

Startups thus differ from mature corporations in that the former are evanescent forms of organization for which real-time responsiveness is of the essence and to which concentrated ownership and high-powered incentives are well suited. Presenting, as they do, "disequilibrium" issues, the efficient alignment of mature (equilibrium) transactions with governance structures, as set out above, does not uncritically carry over. The concept of "best feasible alignment" is nevertheless robust and applies to transactions of both equilibrium and disequilibrium kinds.

Full formalization. As described in Appendix I, the fourth stage in the natural progression involves full formalization. Usually, deeper understanding accompanies moves across successive stages. Yet there are precautions. Loss of contact with the "key interactions" to which Solow refers is a continuous concern: a "model can be right in … [a] mechanical sense, but still rather unenlightening because in some way imperfectly suited to the subject matter. It can obscure the key interactions, instead of spotlighting them" (2001, p. 112).

To be sure, this does not preclude work going on at several stages simultaneously. It is furthermore noteworthy that the scientific criteria for evaluating fully formal theory are exceptionally demanding. If and as, however, full formalizations purport to deal with real phenomena, the test of prediction and empirical testing becomes operative.

It is not for me to say whether David Kreps is correct or not in the quote that follows. I nevertheless conjecture that the continuing dialog to which he refers will be instructive (Kreps, 1999, p. 123):

> It is incontestable that mathematics-based theory has become the common language of economics (see Debreu, 1990 or Kreps, 1996). Most economists, and especially and most critically, new recruits in the form of graduate students,

learn transaction-cost economics as translated and renamed (incomplete) contract theory. I hope it is not taken amiss if I claim that (for the prototypical graduate student) it is harder to read Markets and Hierarchies or The Economic Institutions of Capitalism than the classics-illustrated versions, written in the comfortable language of middle-brow theory. If the classics-illustrated versions miss subtleties connected with, say, bounded rationality and truly unforeseen contingencies, or they do not give any indication of how social embeddedness or internal consistency/simplicity affects a relational contract, then the consumer of the classics-illustrated editions has missed important pieces of the message.

Fully formal research opportunities reside in responding to the above described puzzles and challenges and subtleties – to which transaction cost economics as a science of organization will be the beneficiary.[17]

Incentive Mode	Administrative Intensity	Contract Control	Regime	
I	(+)	(+)	(+)	
II	(+)	(+)	(0)	
III	(+)	(0)	(+)	:Market
IV	(+)	(0)	(0)	
V	(0)	(+)	(+)	
VI	(0)	(+)	(0)	:Hierarchy
VII	(0)	(0)	(+)	
VII	(0)	(0)	(0)	

Table 1. Alternative Modes of Governance

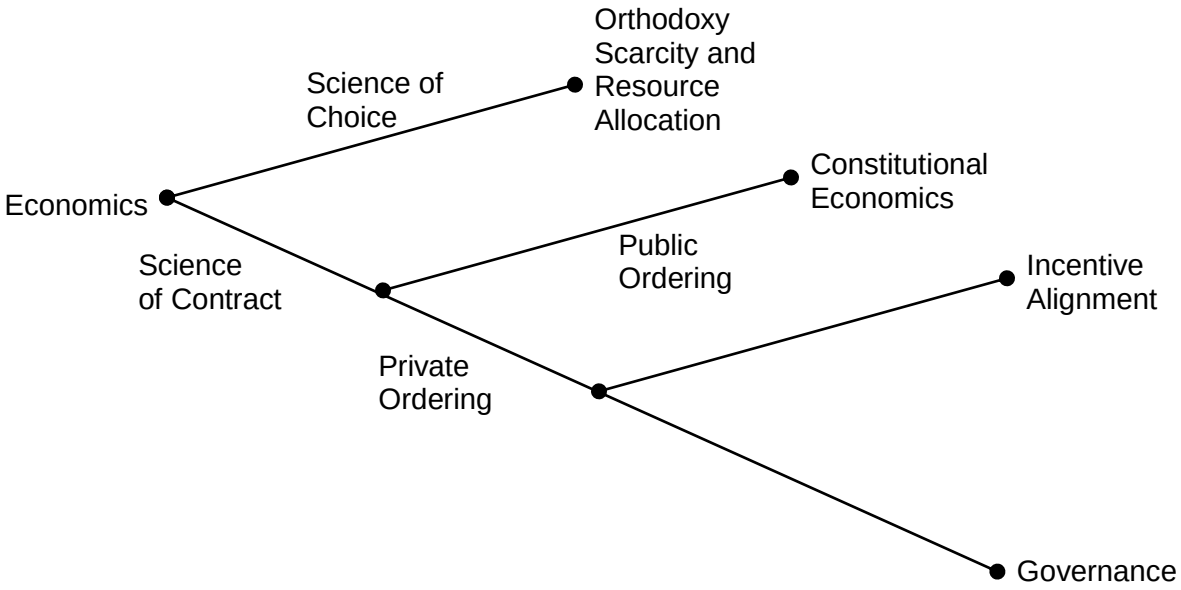
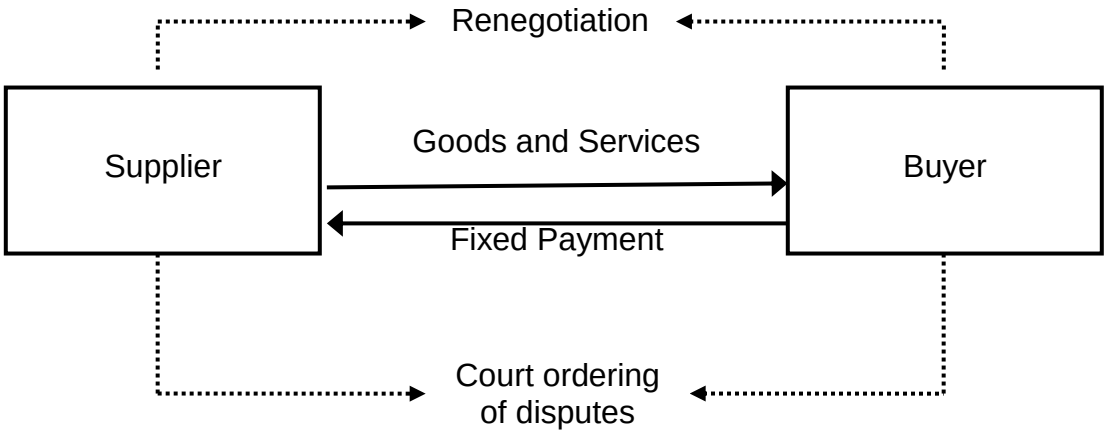
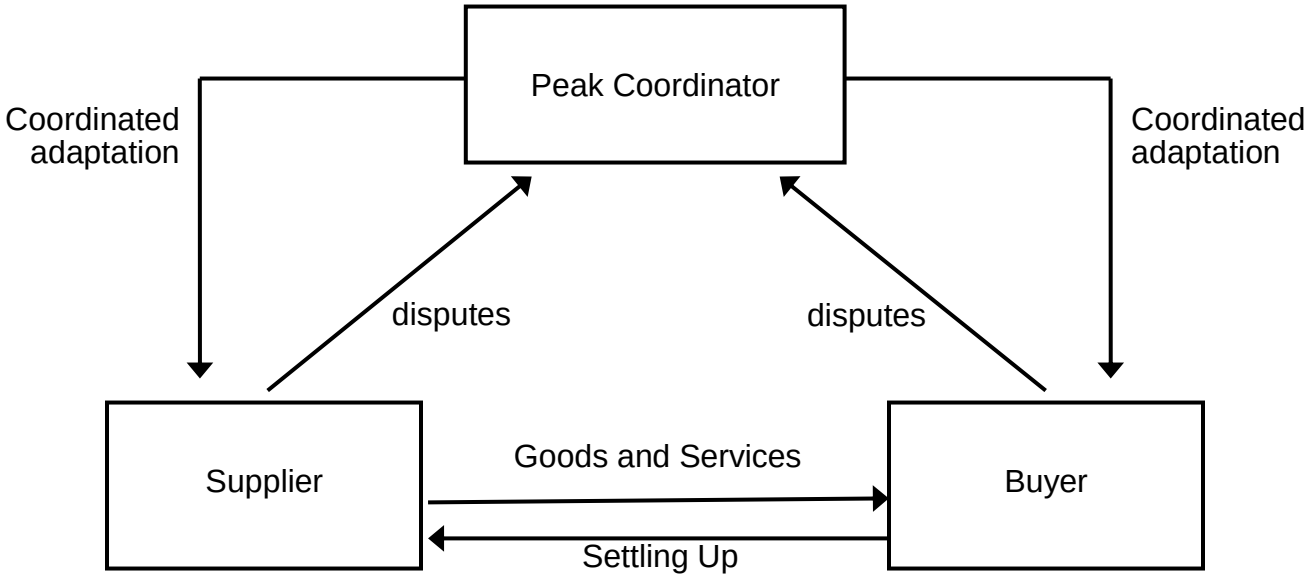


Figure 1. The Sciences of Choice and Contract



Market Mediated Exchange



Coordinated Adaptations and Dispute Settlement
by the Office of the Peak Coordinator

Hierarchy Mediated Exchange

Figure 2.

Footnotes to text

* The author is Edgar F. Kaiser Professor Emeritus of Business, Economics, and Law at the University of California, Berkeley.

This paper was prepared for the conference celebrating Paul Joskow's career in Economics at MIT – as Paul moves on to his next assignment as President of the Alfred P. Sloan Foundation. As everyone is aware, Paul is an accomplished applied microeconomics. Not everyone is aware, however, of the extent of Paul's contributions to transaction cost economics. I have prepared Appendix II as a reminder.

1. A growing interest among economists was taking shape in the 1970s and would give rise to new theories of economic organization – to include mechanism design (see the June 2008 issue of the American Economic Review for the Nobel Prize lectures of Leonid Hurwitz, Roger Myerson, and Eric Maskin), information economics, to which information asymmetries were a recurrent theme (Arrow, 1964, 1974; Akerlof, 1970; Spence, 1971; Rothschild and Stiglitz, 1976), team theory (Marshak and Radner, 1972); agency theory, of formal (Ross, 1973; Holmstrom, 1979) and informal (Alchian and Demsetz, 1972) kinds; property rights theory, also of informal (Alchian, 1961; Demsetz, 1967) and (post-1970) formal (Grossman and Hart, 1986) kinds. Evolutionary Economics (Nelson and Winter, 1973) also deals with related issues.

2. I was a student at Carnegie from 1960-1963, where my experience was very much like that of Jacques Dreze, who observed that "Never since have I experienced such intellectual excitement" (1995, p. 123).

3. The comparative contractual study of firm and market opens the door to an examination of organization more generally, to include the hybrid mode (Williamson, 1991) and the organization of government bureaus, nonprofits, cooperatives, medical care,

educational, and the list goes on. The key move is to examine each of these activities through the lens of contract/governance.

4. Herbert Simon advises that "Nothing is more fundamental in setting our research agenda and informing our research methods than our view of the nature of the human beings whose behavior we are studying" (1985, p. 303).

5. Bounded rationality has been defined as "behavior that is intendedly rational but only limitedly so" (Simon, 1958, p. xxiv).

6. This is a benefit that George Shultz ascribes to economics more generally: "Our discipline makes one look ahead, ask about indirect consequences, take note of variables that may not be directly under consideration" (1995, p. 1).

7. Robert Michels (1962) and Irving Goffman (1969) were keenly aware of the hazards, mainly in contexts different from commercial contracting.

8. What I mean by implementation is "the act of implementing" by breathing operational content into a theory by (1) expressly naming the basic attributes that are ascribed to human actors; (2) explicating the key mechanisms, in both feasibility and procedural respects, and (3) deriving refutable implications to which empirical tests can be applied. Implementation thus be thought of "operationalization" – which word, however, appears neither in my dictionary nor on my spellchecker.

9. The lack of accounting objectivity is a chronic issue for the study of contract and organization.

10. Indeed, W. R. Fisher had reached a similar conclusion at the turn of the 20th century (Fisher, 1907, pp. 39-46):

    Regulation does not end with the formulation and adoption of a satisfactory contract, in itself a considerable task. If this were all, a few wise and honest men might, once in a generation, supervise the framing of a franchise in proper form, and nothing further would be necessary. It is a current fallacy and the common practice in American public

life to assume that a constitution or a statute or a charter, once properly drawn up by intelligent citizens and adopted by an awakened public, is self-executing and that the duty of good citizens ends with the successful enactment of some such well matured plan. But repeated experience has demonstrated – what should have been always apparent – the absolute futility of such a course, and the disastrous consequences of reliance upon a written document for the purposes of living administration. As with a constitution, a statute, or a charter, so with a franchise. It has been found that such an agreement is not self-enforcing. … [Moreover, the] administration may ignore or fail to enforce compliance with those essential parts of a contract entrusted to its executive authority; and legal proceedings … are frequently unavoidable long before the time of the franchise has expired.

11. It is noteworthy that the absence of asset specificity was implicit in the finance literature from Modigliani and Miller (1958) through the 1980s.

12. I observe in this connection that "there may be no more basic function for the board to perform than keeping the gate open so that owners of equity can sell their shares to organized interests that will vote the rascals out" (2008, p. 263). The report that the Securities and Exchange Commission plans to prohibit brokerage firms from voting their clients' shares –which have typically been voted "in favor of standing managements and boards" – holds the prospect of making it easier change the composition of unduly compliant boards (Scannell and Fitzpatrick, 2009, p. C1).

13. There is more to it than the law, in that human assets exercise volition while other assets do not.

14. It is elementary that regulatory commissions (such as the Securities and Exchange Commission) should be especially wary of investment funds that work out of unobserved investment mechanisms and provide suspect earnings reports.

15. Simon interprets the nearly decomposable architecture of complex systems in evolutionary terms – as a response to the overwhelming complexity that would otherwise defeat such systems. The basic regularity in nearly decomposable systems is that "(a) the short-run behavior of each of the component subsystems is approximately independent of the other components; (b) in the long run, the behavior of any one of the components depends only in an aggregate way on the behavior of the other components" (Simon, 1962, p. 474). Modular designs have these properties.

16. An antidemocratic polity, by contrast, can persist much longer. A growing international consensus that an authoritarian government should be socially sanctioned may take a long time to prevail yet be the best reform alternative.

17. On the fully formal modeling of transaction cost economics (or variants thereon), see Grossman and Hart (1986), Hart and Moore (1991), Hart (1995), Bajari and Tadelis (2001), Tadelis (2002), Whinston (2003), Gibbons (2005), Tirole (2009), and Tadelis and Williamson (2009).

Appendix I

Methodology

I sketch here what I take to be the overarching framework out of which transaction cost economics works. I begin with a discussion of pragmatic methodology, after which I discuss the natural progression through which theory development frequently works.

Pragmatic methodology. My brief discussion of pragmatic methodology is a combination of Robert Solow (2001) with Milton Friedman (1953) and my experience as an applied microeconomist.

The first three precepts of pragmatic methodology as set out of Solow are these: keep it simple; get it right; make it plausible – to which list is added: derive refutable implications and submit these to the data.

Friedman observes with respect to the first precept that if "most phenomena are driven by a very few central forces, … [then] what good theory does is to simplify; it pulls out the central forces and gets rid of the rest" (Friedman in Snowdon and Vane 1997, p. 196). Keeping it simple entails naming and explicating the "main case" – which, for transaction cost economics, entails economizing on transaction costs. Indeed, "a fundamental hypothesis of science is that appearances are deceptive and that there is a way of looking at or interpreting or organizing the evidence that will reveal superficially disconnected and diverse phenomena to be manifestations of a more fundamental and relatively simple structure" (Friedman, 1953, p. 33). But whereas Friedman puts it in the singular – "a way" and "a … simple structure" – economic organization is very complex and I subscribe to pluralism: there are "ways" and there are "simple structures." Sometimes these will be complementary, in which case a richer understanding will be realized by combining them. Sometimes they will be conflicting, in which case we can run a competition. For that we need a cutting edge.

Getting it right "includes translating economic concepts into accurate mathematics (or diagrams or words) and making sure that further logical operations are correctly performed and verified" (Solow, 2001, p. 112). Although full formalism (mathematics) is the ultimate objective, often that is reached gradually, through the "natural progression," as discussed below.

Plausible simple models of complex phenomena are expected to "make sense for 'reasonable' or 'plausible' values of the important parameters" (Solow, 2001, p. 112). Also, because "not everything that is logically consistent is credulous" (Kreps, 1999, p. 125), fanciful constructions that lose contact with the phenomena are suspect – especially if alternative and more veridical models yield refutable implications that are congruent with the data. Invoking the implausible assumption of a zero transaction cost is a red flag if the challenge is to "study the world of positive transaction costs" (Coase, 1992, p. 717).

The fourth precept of pragmatic methodology holds that all would-be theories should yield refutable implications. Nicholas Georgescu-Roegen had a felicitous way of putting it: "The purpose of science in general is not prediction, but knowledge for its own sake," yet prediction is "the touchstone of scientific knowledge" (1971, p. 37). This is especially important in the social sciences where advocates of particular "viewpoints" (if not theories) often have strong ideological predilections. Faced with an impasse, prediction serves touchstone purposes.[1]

Most economists know in their bones that theories that are congruent with the data are more influential. Friedman's reflections on a lifetime of work are pertinent: "I believe in every area where I feel that I have had some influence it has occurred less because of the pure analysis than it has because of the empirical evidence that I have been able to organize."[2] There is no question that transaction cost economics is more influential because of the empirical work that it has engendered (Joskow, 1988; Whinston, 2002; Macher and Richman, 2008).

The natural progression. Theories, especially in the social sciences, rarely appear full blown but typically undergo a natural progression from informal to preformal to semiformal to fully formal analysis. Many of the good ideas in Coase (1937) were of an informal kind.

Preformal analysis got underway in the 1970s (Williamson, 1975, 1979; Klein, Crawford, and Alchian, 1978) and evolved into semiformal work in the 1980s (Klein and Leffler, 1981; Williamson, 1983; Riordan and Williamson, 1985) and later. The first fully formal model was the paper by Sanford Grossman and Oliver Hart (1986), which stimulated a large number of follow-on efforts in the years since. Although I have taken the position that their "directional" approach to vertical integration is problematic (Williamson, 2002), this has been deeply influential work.[3]

It is furthermore noteworthy that the theory and the data are often interactive as this progression unfolds (Newell, 1990, p. 114):

> Theories cumulate. They are refined and reformulated, corrected and expanded. Thus, we are not living in the world of Popper … [Theories are not] shot down with a falsification bullet…. Theories are more like graduate students – once admitted you try hard to avoid flunking them out…. Theories are things to be nurtured and changed and built up.

The predisposition to nurture and build up to which Allan Newell refers does not, however, go on indefinitely. All would-be theories eventually need to stand up and be counted, especially in the social sciences, where the phenomena are uncommonly complicated and "any direction you proceed in has a very high a priori probability of being wrong" (Simon, 1992, p. 21).

Note with respect to this last that would-be theories that fail to yield predictions that are corroborated by the data are not the exception but the rule. Yet theories that "fail" are often informative. A better understanding of the phenomenon is sometimes realized in the process, as a result of which future research can be redirected in more promising ways. Additionally, new tools of analysis are sometimes developed by fully formal modeling efforts that find applications in the study of other phenomena.

Footnotes to Appendix 1

1. I attended two remarkable conferences in which the need for empirical testing was contested. The first of these was an NBER conference on Money and Macroeconomics that was held at Carnegie when I was a graduate student. The most controversial exchange was over the Friedman-Meisselman empirical study of the simple Keynesian model, which pitted Friedman against "most of the rest." Aside from being a formidable debater, Friedman had the data and had run the tests. That mattered to the outcome and had a lasting impression.

   The second was a conference in the 1980s in Europe, which pitted transaction cost practitioners against transaction cost critics. What most distinguished these two was that transaction cost economics made predictions that were submitted to the data while the opponents came up empty in predictive/empirical respects. (Paul Joskow was at the second conference and, I hope, remembers it similarly.)

2. Personal communication from Milton Friedman to the author, February 6, 2006.

3. Hart and Moore's recent paper on "Contracts as Reference Points" concedes that the Grossman-Hart-Moore tradition placed too much emphasis on noncontractible ex ante investments, is poorly suited to studying the internal organization of firms, and poses some foundational concerns (2008, pp. 2-3).

Appendix II

Joskow Contributions to Transaction Cost Economics

"Vertical Integration and Long Term Contracts: The Case of Coal Burning Electric Generating Plants", Journal of Law, Economics and Organization, 1:1, Spring, 1985, pp 33-80.

"Long Term Vertical Relationships and the Study of Industrial Organization and Government Regulation", Journal of Institutional and Theoretical Economics, December 1985, 587-593.

"Contract Duration and Relationship Specific Investments: Empirical Evidence from Coal Markets", American Economic Review, 77:1, March 1987, pp 168-185.

"Price Adjustment in Long Term Contracts: The Case of Coal", Journal of Law and Economics, XXXI, April 1988, pp 47-83.

"Asset Specificity and the Structure of Vertical Relationships: Empirical Evidence," Journal of Law, Economics and Organization, IV:1, Spring 1988, pp.95-117.

"The Performance of Long-Term Contracts: Further Evidence from Coal Markets," Rand Journal of Economics, 21:2, Summer, 1990, pp 251-274.

"The Role of Transactions Cost Economics in Antitrust and Public Utility Regulatory Policies", Journal of Law, Economics and Organization, 7, 1991, pp 53-82.

"The New Institutional Economics: Alternative Approaches", Journal of Institutional and Theoretical Economics, 151:1, March 1995, 248-259.

"Introducing Competition into Regulated Network Industries: From Hierarchies to Markets in Electricity", Industrial and Corporate Change, 5:2, 1996, pp 341-382.

"Asset Specificity and Vertical Integration", The New Palgrave Dictionary of Economics and Law, Peter Neuman, Editor, MacMillan, 1998.

"Transaction Cost Economics, Antitrust Rules and Remedies," Journal of Law, Economics and Organization, 18:1, April 2002, pp 95-116.

“Electricity Sector Restructuring and Competition: A Transaction Cost Perspective,” The Economics of Contracts: Theories and Applications, Eric Brousseau and Jean-Michel Glachant (eds.)., Cambridge University Press, 2002.

“Vertical Integration,” Handbook of New Institutional Economics, (C. Menard and M. Shirley, eds.), 2005, Springer.

“Regulation and Deregulation after 25 Years: Lessons for Research,” International Review of Industrial Organization, 26: 169- 193, 2005.

“Competitive Electricity Markets and Investment in New Generating Capacity,” The New Energy Paradigm (Dieter Helm, editor), Oxford University Press, 2007.

"Regulation of Natural Monopolies," Handbook of Law and Economics, A.M. Polinsky and S. Shavell, editors, Elsevier, 2007.

"Incentive Regulation in Theory and Practice: Electric Distribution and Transmission Networks," in Economic Regulation and its Reform: What Have We Learned? (N. Rose, ed.), University of Chicago Press, forthcoming.

"New Institutional Economics: A Report Card," New Institutional Economics, J-M Glachant and E. Brouseau, eds. Cambridge University Press, 2008.

"Vertical Integration," Issues in Competition Law and Policy, Wayne Dale Collins (ed.), Volume 1, Chapter 11, American Bar Association Section on Antitrust Law, August 2008.