

Theories of Firm and Market Organization: Focal Transactions, Empirical Testing, and Scaling Up

Steven Tadelis and Oliver E. Williamson
University of California, Berkeley

December 29, 2011

Abstract

Theories of economic organization often differ in three respects. First, and the most basic, is the choice of the focal transaction upon which the analysis is grounded – which is important in its own right and bears on whether other transactions can be interpreted as variations on a theme. The second is the extent to which the theory enables and is corroborated by empirical testing. And the third is whether the simple analytical model scales up to approximate real world entities. We conclude that as between the employment relation and the intermediate product market transaction, which are the two most widely used focal transactions, that the latter (1) is more tractable; (2) better permits other transactions to be interpreted as variations on a theme; (3) more readily enables empirical testing; and (4) better supports scaling up. We also observe that while the choice of the focal problem is basic, it is not fully determinative of the analytical framework by which the analysis is implemented. We argue in that many theories of economic organization, especially the more formal ones developed in the last 25 years, experience problems in one or more of these four ways whereas Transaction Cost Economics (TCE) compares favorably in all four respects. The full formalization of TCE, however, remains a work-in-progress.

Introduction

It has been 75 years since Ronald Coase (1937) challenged economic orthodoxy by pointing out that coordination of economic activity is sometimes done through markets using the price mechanism, while other times it is done in organizations like firms under the direction of managers, but that the economics of that time had failed to explain which activities go where and why. As Coase puts it, “[i]t is clear that these are alternative methods of coordinating production. Yet, having regard to the fact that if production is regulated by price movements, production could be carried on without any organization at all, well might we ask, why is there any organization?” (p. 388).

By posing this basic question, Coase launched a study of transaction costs and the economics of organization that would grow into a thriving area of research, which is very much thriving still. As, however, Coase observed in 1972, his 1937 paper was “much cited but little used” Coase (1972, p.63). It was much cited because many of his readers could see merit in the issues that Coase (1937) had raised. As we discuss below, however, it was not in the least obvious how to breathe operational content into the markets and hierarchies challenge that he had posed. This paper examines why it took so long for the concept of transaction costs and economic organization to take hold. We also examine theories of firm and market organization that appeared in the 1970s and later and observe that these differ both from Coase 1937 and among themselves in conceptual, empirical, and scaling up respects.

The paper is organized as follows. Section 1 is a brief background that highlights the main papers in the literature. Section 2 discusses Coase (1937) and his reasons for

focusing on the employment relation. We explain why the intermediate product market transaction, the focal choice in Transaction Cost Economics (TCE), has many advantages over the employment relation as a focal transaction, with special emphasis on how firms and markets differ in “interface management” respects. Section 3 articulates a version of Coase’s main theoretical arguments and argues that his theory, as well as other more recent theories such as the modern Property Rights Theory and Agency theory, lack the empirical “hook” needed to take the theory to data – which we contrast to the developments of TCE. Section 4 addresses the issue of scaling-up as a non-trivial challenge. Thus although there is widespread agreement that the first precept of all good theories is to “keep it simple” (Solow, 2001. p. 111), simplifications that do not scale up, hence lose contact with the phenomenon of interest (in this case, real firms), are not to be regarded on a parity with those that do scale up. We conclude in section 5 with a discussion of some the obstacles that we believe each of these approaches faces in the future.

1. Background

It is hard to overstate the importance and influence of Ronald Coase’s (1937) groundbreaking paper, which in his own words aims “... to bridge what appears to be a gap in 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.” (p. 389).

As fundamental as Coase’s challenge was, and despite the interesting issues that he discussed in his paper, the issues lay dormant until the 1960s. Part of that delay, we contend, is because the analytical approach that Coase recommended in 1937 could not be implemented and it was not obvious where to go from there.

Broadly in the spirit of Marshallian marginal analysis, Coase invokes “decreasing returns to the entrepreneur function” to explain limitations to firm size (1937, p. 394) and further asserts that “the whole of the ‘structure of competitive industry’ becomes tractable by the ordinary technique of economic analysis” (Coase, p.398). Had that been correct, transaction costs would have soon thereafter been incorporated into the neoclassical setup, but nothing of the sort transpired.

The problem is that the application of transaction cost reasoning across a wide range of transactions that differ in their attributes is not a matter for which the neoclassical concept of the firm is well-suited. As Harold Demsetz put it, 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).

Aside from Coase’s observation that, but for positive transaction costs, everything would be organized in one large firm, there were no refutable implications. Indeed, as Alchian and Demsetz would observe, the expansive concept of transaction costs was so elastic as to permit any anomaly to be explained after the fact – which is to say that it was tautological (1972, pp. 783-784).

The concept was nevertheless important and was given a new life when Coase (with respect to externalities (1960)) and Arrow (with respect to vertical integration (1969)) showed that externalities and vertical integration would vanish if the then-standard assumption of zero-transaction costs was maintained. That is a preposterous state of affairs and invited a series of conceptual papers by Williamson (1971, 1979), Alchian and Demsetz (1972), and Klein Crawford and Alchian (1978) which brought the economics of governance to the forefront in the 1970s. It took another decade and more before a new generation of economists moved beyond words and diagrams to bring mathematical methods to bear on the issue of governance. Most influential among these efforts are the fundamental contributions of Grossman and Hart (1986) and Hart and Moore (1990), who launched the modern Property Rights Theory of the firm (PRT), as well as Holmstrom and Milgrom (1991, 1994), who developed what we refer to as the Agency approach to organizational design.

The formal modeling of TCE is challenging. After 25 years in the making, the theoretical literature that aims to uncover the costs and benefits of different modes of organization has not always led to clear connections between theory and practice. Contributing factors include one or more of the following: the choice of employment as the “focal transaction” by Coase and others, the lack of empirically measureable variables, and the challenge that the theories have in “scaling up” to correspond with the real-world organizations that we observe. It is useful to expand a bit on each of these three neglected factors.

First, “firm” and “market” are often interpreted as “employee” and “contractor,” respectively. Thus, Coase (1937) took the employment relation to be focal, it being his

view that “A firm is likely therefore to emerge ...in the case of services – labor – than it is in the case of the buying of commodities” since, for the latter, “the main items can be stated in advance and the details which will be decided later will be of minor significance” (1937, p. 392). Albeit an important and interesting transaction, taking the employment relationship to be the “focal transaction” limits the scope of analysis and precludes a more general understanding of the costs and benefits of different modes of organization as they relate to non-labor transactions. Awaiting a deeper examination, Coase’s dismissal of the intermediate product market transaction is premature. (Indeed, as discussed below, the intermediate product market transaction has proven to be a more instructive focal transaction.)

Second, much of the literature has emphasized methods of analysis to the neglect of empirical relevance. As discussed above, Coase (1937) contends that his analysis shows that “a definition of a firm may be obtained which is not only realistic in that it corresponds to what is meant by a firm in the real world, but is tractable by two of the most powerful instruments of economic analysis developed by Marshall, the idea of the margin and that of substitution, together giving the idea of substitution at the margin.” (p. 386-7). So described, Coase (1937) has not led to empirical analysis that sheds light on the costs and benefits of different modes of organization. TCE, in contrast, treats firm and market organization as alternative modes of governance that differ in discrete structural ways, which can be taken to real world data in a rather straightforward way. Similarly taking other formal theories that have emerged in the past 25 years to the data has often encountered obstacles (see, e.g., Whinston (2001, 2003)).

Third, by and large, the analytical models that have been developed in the past 25 years do not “scale-up” to include multiple such stages in real world organizations. As Holmstrom (1999) put it, “The problem is that there really are no firms in these [formal] models, just representative entrepreneurs.” (p. 100). TCE, by contrast, focuses not on entrepreneurs but considers “mini-firms” – namely a series of stages, of within each of which the activity is technologically non-separable but between which stages the activities are technologically separable, and asks how to organize the interfaces between them. As we suggest later, this approach leads quite naturally to scaling up in ways that relate to real world organizations.

2. The Choice of a Focal Transaction

Inasmuch as simple market exchange between anonymous traders for well-defined goods and services can be presumed usually to work well, longer term contracts for which actions, decisions and modifications take place over time is the class of activity for which the choice of governance is material. For simplicity, we will focus on markets and hierarchies, without reference to intermediate levels of control.¹ Whereas TCE views the relation between buyer and seller symmetrically, Coase put the purchaser in control: when the “details of what the supplier is expected to do is not stated in the contract but is decided later by the purchaser ...[then] that relationship which I term a ‘firm’ may be obtained” (1937, p. 392).

¹ See Williamson (1983, 1991) for more on hybrid transactions and intermediate levels of control.

The two transactions that Coase considered for explicating this relationship are the employment relation and the intermediate product market transaction.² As discussed above, Coase took the employment relation to be focal.

By way of elaboration, Coase writes that “it is important to note the character of the contract into which the factor enters that is employed within a firm. The contract is one whereby the factor, for a certain remuneration (which may be fixed or fluctuating), agrees to obey the directions of an entrepreneur *within certain limits*. The essence of the contract is that it should only state the limits to the powers of the entrepreneur.” (p. 191, emphasis in original text.) He furthermore states in his conclusions that “We can best approach the question of what constitutes a firm in practice by considering the legal relationship normally called that of ‘master and servant’ or ‘employer and employee.’” (p. 403). He concludes this comparison with the observation that “it is the fact of direction which is the essence of the legal concept of ‘employer and employee,’ just as it was the economic concept which was developed above.” (p. 404)

Interestingly, Simon (1951), in what at the time would be considered state-of-the-art formal modeling (in an era that is pre Arrow-Debreu securities and pre-dates agency theory), also focused on the employment relation as the alternative to a sales contract, which is the common one in market exchange. Although Coase (1937) is not cited, Simon’s definition of employment is very much in congruence with that of Coase and also relates to the “zone of acceptance” as described by Chester Barnard (1938). Simon

² Coase (1937) clearly indicates that he tried to encompass a general view of organization and considered many kinds of transactions. For example, he touches on both vertical and horizontal integration when he argues that “the definition of a firm that was given above can be used to give more precise meaning to the terms “combination” and “integration”... ” (p. 397), and later makes the clear connection between “combination” and what is now referred to as horizontal integration, and between “integration” and what is now referred to as vertical integration. He nevertheless chose to focus attention primarily on the employment relation as the defining essence of a firm.

considers a worker, W , a Boss, B , and a decision that needs to be made, x , and writes that “ W accepts authority when his behavior is determined by B 's decision. In general, W will accept authority only if x_o , the x chosen by B , is restricted to some given subset (W 's “area of acceptance”) of all the possible values. This is the definition of authority that is most generally employed in modern administrative theory.” (1951, p. 294).

Issues of firm and market organization and the roles of authority and hierarchy therein remain mainly dormant in the economics literature for another 20 years, possibly because, once again, of this emphasis on the employment relation. Be as it may, when Williamson (1971) first laid out the rudiments of TCE, as further developed in Williamson (1975, 1976, 1979, 1985) and in Klein, Crawford, and Alchian (1978), the intermediate product market transaction was made focal. Applications grew, an empirical literature progressively took shape, and a more formal literature appeared in the early 1980's. We contend that focusing on the intermediate product market transaction led to a more fruitful analysis of organization and governance for four reasons: the employment relation is too encompassing; the intermediate product market transaction is simpler; the intermediate product market transaction focuses especially on transactions as exchanges between technologically separable stages; and the employment relation, interesting and important as it is, never led to an interpretation of other transactions as variations on a theme.

Encompassing. If a firm makes rather than buys components, then laborers that are hired by “the firm” presumably make components. But Coase claimed that it is easy to contract out for components. Do workers instead assemble components? But that sounds mundane. Note, moreover, that the supplier firms to which components are

outsourced also produce those components with hired labor. Whether, therefore, a firm produces to its own needs or buys the component in question from an independent supplier, the supply stage hires labor. The questions of which components are taken out of the market and are produced internally and of the tradeoffs that accrue thereto never get directly posed if supply is described in this way. As a concrete example, consider the case of legal services. Large companies will have several, sometimes many dozens of lawyers working in-house as employees. At the same time these firms will contract for other legal services which are performed by law firms who hire employees to perform the work. Hence, employees are used in both of these scenarios, so this does not explain when the transaction is conducted in the firm and when through the market.

Simplicity. Economic organization being very complex, the choice of a tractable focal transaction is especially important. As between the employment relation and the intermediate product market transaction, it is our judgment that the latter is simpler and better directs our attention to the “central forces.”

As an example, consider the case of Boeing’s Dreamliner aircraft. Many delays have been experienced in the construction of the Dreamliner.³ One of the more important components that were outsourced – which TCE would have predicted would result in bilateral dependency and become the source of delays if outsourced – was the aircraft’s fuselage, which was designed and manufactured by Vought Industries. With the benefit of hindsight, in July of 2009 Boeing acquired Vought in order to achieve better coordination of the production process. Taking the intermediate product market

³ These delays are responsible for its first aircraft being delivered three years behind schedule. See http://money.cnn.com/2011/10/25/news/companies/boeing_787_dreamliner/index.htm for a recent article. Many more can easily be found online.

transaction to be focal for applying the logic of make-or-buy to the fuselage decision would have been instructive.

Stages. Transactions in the intermediate product market take place not between individuals but between technologically separable stages of production. That directs attention to interface management within and between firms. Rather than deal with individuals, the intermediate product market transaction highlights how to organize the interface (as between make-or-buy) between technologically separable *stages* of production (the “mini-firms” to which we referred earlier within each of which employment relations are operative). Posed in this way, the questions to be addressed are which transactions are better mediated at the interface by markets, which by hierarchies, and why. As we describe in more detail in Section 3, attention is thereby focused on the defining attributes of transactions and governance structures, and not on employee versus contractor, or other elements of human capital.

As an illustration, consider the wave of Information Technology (IT) outsourcing that became quite ubiquitous in the early 2000’s, where many large firms outsourced their IT services to specialized providers. Firms would then let go of the employees they hired themselves, and the work was done by employees of the IT service firms, implying that regardless of the firm they are in, employees are the ones performing the job. Considering IT services as a form of an intermediate good (a support service that is essential for the business) makes it possible to pose the question of when is it that IT services are better procured through the market or within the firm, ignoring the fact that it will be performed by one kind of employment or another.⁴

⁴ See Tadelis (2007) for a discussion of the IT outsourcing wave and the problems that several large firms faced after outsourcing these transactions.

Experience. As mentioned earlier, Coase(1972, p.63) stated that his 1937 paper had been “widely cited but little used” over the intervening 35 years.⁵ It was cited because the lapse in orthodoxy to which he referred was both real and serious. Partly it was little used because it was ahead of its time, to which the application of zero transaction cost reasoning by Coase (1960) and Arrow (1969) are pertinent. But it was also little used because the employment relation was a less tractable focal transaction than the intermediate product market transaction – which would become apparent in the 1970s and beyond. Indeed, after the intermediate product market transaction was made focal, it became apparent that the kind of costs and benefits described by TCE applied not only to vertical integration, the focus of Williamson (1971), but also to horizontal integration (especially with respect to organization form, to which Chandler (1962) is relevant), franchise bidding for natural monopoly (Williamson (1976)), regulation (Goldberg (1976)), corporate finance (Williamson (1988); Woodward(1988)), corporate strategy (Nickerson and Silverman (2009)), marketing (Gatignon and Gatignon (2010), John and Reve (2010)), political economy (Dixit (1996), Levy and Spiller (1994)), procurement (Bajari and Tadelis (2001)) and privatization (Levin and Tadelis (2010)). Indeed, Jeffrey Macher and Barak Richman (2008) report that there have been more than 900 empirical tests of TCE in economics, business, and the contiguous social sciences. These applications posed new questions of their own but were also variations on themes that would arise from taking the intermediate product market transaction to be focal.

⁵ The curious reader can look at the citation pattern of Coase’s paper. It actually received few citations until the 1970’s, while citation counts started rising dramatically in the 1980s and beyond. That Coase (1937) was published in the 1952 AEA Readings in Price Theory (1952) is nonetheless noteworthy.

3. Firm Boundaries and Empirical Traction

As mentioned earlier, Coase not only uncovered a lapse in economic theory by asking why are there firms and what determines their size, but he suggested that the classical economic reasoning of substitution at the margin can be applied to answer this question, thus offering a method for addressing these questions. Summarizing in his own words, “A firm, therefore, consists of the system of relationships which comes into existence when the direction of resources is dependent on an entrepreneur. ... a firm becomes larger as additional transactions (which could be exchange transactions coordinated through the price mechanism) are organized by the entrepreneur and becomes smaller as he abandons the organization of such transactions. The question which arises is... [w]hy does the entrepreneur not organize one less transaction or one more?” (p. 393-4).

Coase addresses three important issues. First, there is a clear definition of a firm: transactions in a firm are directed by an entrepreneur and not by the price mechanism. Second, there is a clear outline of the boundaries agenda: the size of the firm is measured in the number of transaction that is organized in the firm, as a substitute to organizing them through the market. Last but not least, Coase poses the question of comparative statics: what factors will cause a firm to grow or contract? It is this last point to which economic methodology comes to bear on organization: what is the measureable margin that will determine the size of a firm?

Interestingly, and as we described briefly in Section 1, Coase writes in the spirit of Marshallian marginal analysis that “as the firm gets larger, there may be decreasing

returns to the entrepreneur function”, which he uses to justify limitations to firm size (1937, p.394). He furthermore asserts that “the whole of the ‘structure of competitive industry’ becomes tractable by the ordinary technique of economic analysis” (Coase, p. 398). If correct, transaction costs would have been easily incorporated into the neoclassical setup and would have become textbook orthodoxy. Instead, it was mainly ignored for most of the next 30 years.

The problem is that the purported ease of incorporating transaction costs with orthodoxy was never demonstrated – in 1937 or in 1970 or 2012. Aside from the observation that, but for positive transaction costs, everything would be organized in one large firm (Coase, 1937, p. 294), there were no refutable implications that could be taken to data.⁶ To the contrary, as Alchian and Demsetz (1972) observe, the expansive concept of transaction cost was so elastic as to permit any anomaly to be explained after the fact – which is to say that it was tautological. Indeed, they write: “We do not disagree with the proposition that, *ceteris paribus*, the higher is the cost of transacting across markets the greater will be the comparative advantage of organizing resources within the firm; it is a difficult proposition to disagree with or to refute.” (1972, pp. 783-784).

We see therefore that a closer examination of Coase’s arguments reveals that there is a serious missing piece. Namely, there is little one can say about which transactions should be organized in the firm and which through the market. Coase did recognize, however, that “quite apart from the question of diminishing returns, the costs of organizing certain transactions within the firm may be greater than the costs of carrying out the exchange transactions in the open market.” (1937, p. 396)

⁶ For example, Coase’s decreasing returns story implies that a reason for firms to vary in size is the talent of their entrepreneurs in reducing transaction costs. But talent for reducing transaction costs cannot be measured independently.

Nonetheless, Coase does not offer that empirical “hook” that is needed to answer the question of which transactions should be organized within the firm and which not. Moreover, Coase does not give a clear explanation of how firms and markets differ in the way each mediates exchange, making it quite hard to identify the characteristics of transactions that would make one mode of organization preferable to the other.⁷ Realistically, however, there is only so much that can be expected of a paper as original as Coase (1937), and, later, his demonstration that economic organization is uninteresting without positive transaction costs (Coase (1960)).

What TCE does, in effect, is to relieve the lapse in economic theory to which Coase 1937 called attention by breathing operational content into positive transaction costs, working up the logic of “discriminating alignment” (to be explained below), displaying the public policy ramifications, and undertaking empirical testing.⁸

Williamson (1971) recognized that only “when the need to make unprogrammed adaptations is introduced does the market versus internal organization issue become engaging.” (1971, p. 113). Williamson begins with the observation that when contracts, be they simple or complex, are complete then there is no advantage of internal organization over market transactions. He furthermore observed that even when long-term contracts are incomplete, a series of short-term contracts can in principle deal with the need for adaptation. The rub is that “Short term contracts... [in practice] pose problems ... if either (1) efficient supply requires investment in special-purpose, long-life

⁷ Coase lists several issues that impose costs of transacting such as price discovery, marketing, and more, yet none of these gave way to empirical predictions.

⁸ Coase generously concurs when he stated that “My discussions with researchers in this field has made it clear to me that that Williamson’s influence has been immense. In a real sense, transaction cost economics, in his writing and teaching, is his creation.” (1993, p. 98). We would put it somewhat differently: Williamson was there at the creation, but many others have also contributed to its success.

equipment, or (2) the winner of the original contract acquires a cost advantage, say by reason of ‘first mover’ advantages (such as unique location or learning, including the acquisition of undisclosed or proprietary technical or managerial procedures and task-specific labor skills).⁹ He later continues, “In consideration, therefore, of the problems that both long and short-term contracts are subject to, vertical integration may well be indicated... [since] differences between successive stages can be resolved more easily by the internal control machinery.” (1971, p. 116)

To summarize, Williamson (1971) contributes to the the development of TCE in three important ways. First, the intermediate product market transaction is taken to be focal and the management of adaptation is the central problem to be solved. Second, he observed that market and hierarchy each deal with adaptation in different ways, which is consequential for the choice between them. Last, operational life was breathed into the TCE project: given a non-trivial degree of contractual incompleteness, intermediate product market transactions that that are characterized by a higher degree of specificity, either due to specific assets or specific knowledge or skills, are candidates for governance by hierarchy (inside the firm).

Two things are worth expanding upon: the issue of managing the interface of technologically separable stages, and the resulting comparative statics.¹⁰ First, taking the intermediate product market transaction as focal directs attention to interface management within and between firms. Figure 1 is a heuristic display how the make and buy alternatives differ in this respect. Market mediated exchange, shown in the upper

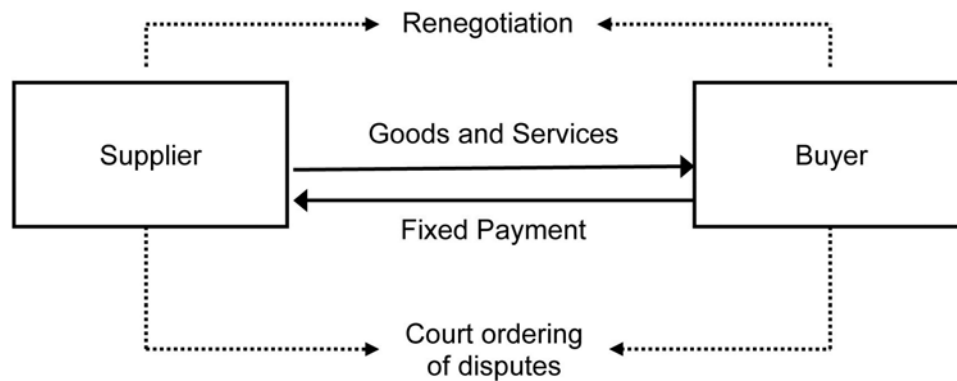
⁹ Note that point (1) is exactly the issue that Klein, Crawford and Alchian (1978) elaborate on. They indeed write that their “attention to appropriable specialized quasi rents is not novel” and attribute it to “Williamson's path breaking work in the area” (p. 301-2).

¹⁰ The ideas sketched out below are elaborated more carefully and in more detail in Williamson (1985), and summarized in Tadelis and Williamson (2011).

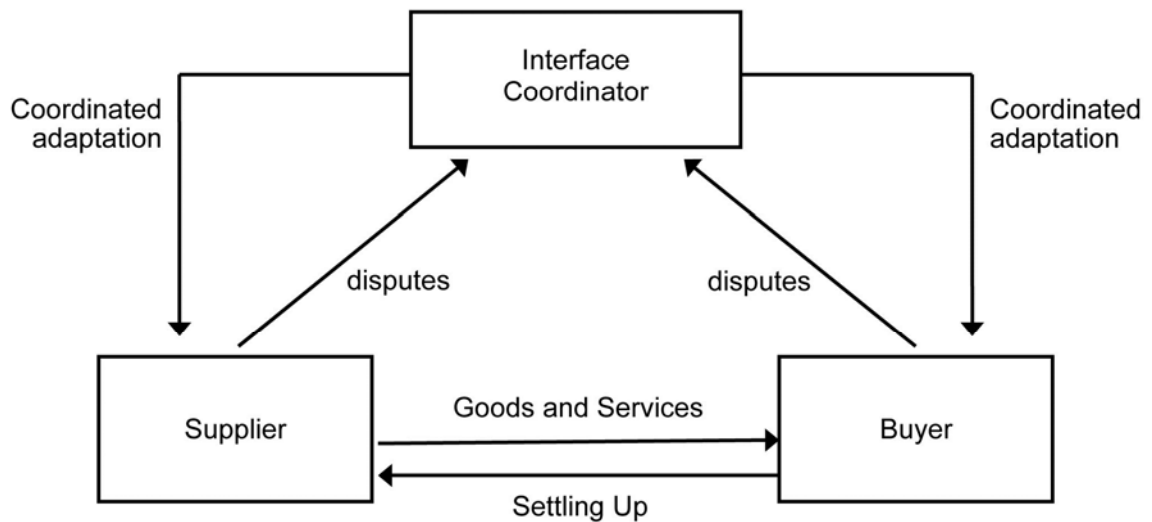
panel of figure 1, takes place between independently owned and operated stages. The parties to market mediated exchange will negotiate the initial contract in a hard-headed way, each will appropriate the stream of net receipts that accrues to it from autonomous adaptation during the execution of the contract (what is referred to as “high-powered incentives” in the agency literature), and either party can appeal irreconcilable disputes to the courts, which in turn will conduct hearings under the applicable rules of evidence and apply the appropriate legal rules to award money damages. By contrast, coordinated adaptation under hierarchy, as shown in the lower panel of Figure 1, is promoted by unified ownership of the two stages coupled with the creation of a new actor, the *interface coordinator*, to which each stage reports and, as needed, is subject to administrative direction and control. Low-powered incentives within stages (akin to cost-plus reimbursement), considerable administrative control by the interface coordinator, and the absence of outside legal appeal (which reinforces the authority of the interface coordinator) are thus the interface instruments by which hierarchy operates. We return to this distinction again in Section 4.

Second, the comparative statics suggested by Williamson (1971) offer a natural way to order transactions, and explains why different kinds of transactions are better mediated through the form or the market. Transactions with high levels of specificity are those for which organizing inside the firm is less costly than through the market. As transactions become less specific, the benefit of organizing the transaction inside the firm drops, and for transactions with a low enough level of specificity, it is less costly to organize through the market.¹¹ This was referred to earlier as “discriminating alignment,”

¹¹ This holds the level of contractual incompleteness fixed. See Riordan and Williamson (1985) for a formal treatment of this result. A complementary comparative static argument holds when the degree of



(Simple) Market Mediated Exchange



(Complex) Hierarchy Mediated Exchange

Figure 1: Interface Mediation

where transactions with different characteristics are aligned with governance mechanisms that are best suited to coordinate the required adaptations that are anticipated.

specificity is held fixed and the level of incompleteness varies. See Tadelis (2002) for a formal treatment of this statement, and Tadelis and Williamson (2011) for a treatment of both comparative static results.

TCE therefore offers measureable characteristics of transactions, as well as measureable characteristics of modes of governance, and offers clear comparative statics that can be taken to data. It is not surprising, therefore that a “very large empirical literature exists lending support to the TCE” (Whinston, 2001 p. 184).¹²

Coase’s original effort to understand the costs and benefits of firm versus market is not alone in making path-breaking conceptual contributions while, at the same time, lacking in empirical applicability. Another example of such an influential contribution is the seminal paper by Grossman and Hart (1986), and the important follow up work of Hart and Moore (1990), which together lay the foundations of PRT. They start like Coase by offering a definition of the firm: “We define the firm as being composed of the assets (e.g., machines, inventories) that it owns.” (1986, p. 692). They then model the interaction between two stages of production engaged in a (vertical) transaction, where the crucial issue is to decide, as between two independent stages of production, which stage has residual rights of control over the assets that it owns. By contrast with TCE, which assumes that integration implies unified ownership and the use of hierarchy to implement coordinated adaptation at the interface, the Grossman-Hart-Moore setup contemplates *directional* integration. Whether A acquires B or B acquires A thus matters. The two main empirical predictions are that first, joint ownership of assets is not optimal, and second, the deciding factor on whether A should acquire B or B should acquire A is which of the two individuals operating stages A and B has more important ex ante investments in specific human capital.

Three weaknesses have been noted regarding the empirical content of the PRT approach. First, as Whinston (2001, 2003) clarifies, the comparative statics of PRT

¹² See, e.g., the surveys by Shelanski and Klein (1995) and by Macher and Richman (2008).

concern marginal variations in which party's investment is more important, which are difficult to measure directly and proxy measures are hard to come by. Holmstrom (1999) states this problem succinctly when he states that it "is challenging to think of ways to measure these effects, since they involve investments and associated margins that according to the theory cannot be observed or priced." (1999, p. 87).

Second, Holmstrom observes that the "prediction of the Hart-Moore model that joint production never is optimal is counterfactual. Joint ventures (and shared ownership, more generally) have always been an important part of the corporate landscape." (1999, p. 86). This critique of Holmstrom is quite important. When a firm is created, then that is the stage at which organizations choices are made and boundaries are defined.¹³ If two stages are then to be put under unified ownership, then why would not one manager be in charge of both stages to maximize joint output? This is ruled out by PRT's assumption that two separate managers are needed for the two stages. But that does not preclude de novo entry into both stages by an entrepreneur who assigns to himself the role of "interface coordinator" and creates separate managers to operate each stage. The directional inefficiency to which PRT refers would then be relieved, albeit at the cost of an interface coordinator.

Last, in a point to which we return below in section 4, the action in PRT is at the individual level, and assets are owned by these individuals. Holmstrom takes aim at this core feature of PRT: "There are neither firms nor workers. The same critique that was directed at Alchian and Demsetz' vision of the firm, that organizational affiliations did not matter for transactions, could be directed at the Hart-Moore model just as well.

¹³ Of course, changes in circumstance after a firm is founded may cause changes in organizational form. This is not critical for what follows.

Individual ownership of assets does not offer a theory of organizational identities unless one associates individuals with firms.” (1999, p. 87).

A noteworthy feature of PRT is that directional integration is treated symmetrically, in that inefficiency (manifested as weakened incentives to invest in ex ante specific human capital) is associated with vertical integration whatever the directional choice, putting the costs and benefits of integration versus market intermediation on equal footing. Indeed, the modeling technology that Grossman and Hart brought to bear on the problem of organization, where discrete institutional choices affect the marginal responses of well-defined functions, has proven as a useful vehicle to investigate a host of related problems.¹⁴ As Holmstrom puts it, “Thanks to this theory we have an improved understanding of how market incentives operate and, importantly, a modeling technology with which to investigate them.” (1999, p.100).

Finally, we observe that the empirical content of Agency theory as a theory of governance and organizations is also limited. Agency theory is silent about how transactions differ with respect to measureable attributes such as specificity or contractual incompleteness. It instead investigates the consequences of incomplete information, in that some aspects of the transaction have zero costs of contracting while others have infinite costs. Specifically, measures of the transaction (say the component's functionality, quality, or a signal of these) can be specified at no cost, while the effort and actions of people who work towards achieving the desired goals cannot be specified at any cost. With few exceptions, of which the contributions of Holmstrom and Milgrom (1991, 1994) with respect to multi-task agency and its applications are noteworthy,

¹⁴ E.g., see Aghion and Bolton (1992) for corporate finance and Aghion and Tirole (1997) on the nature of authority.

Agency theory has offered little insights on the function of hierarchy and governance. Instead, even the work of Holmstrom and Milgrom is most valuable in understanding the employment relation, as it is mostly concerned with how individuals ought to be compensated, and how tasks should be allocated to individuals.

4. Scaling Up

The ability to take a theory to data is one critical feature upon which all would-be theories are judged, regardless of the discipline. There is, however, another virtue that is no less important but that is not as often considered, which we refer to as the “scaling-up test”. Economic models are, by and large, caricatures of reality for good reason. Given that economics (and the social sciences more generally) are inordinately complex (Simon, 1957, p.89; Wilson, 1999, p.183), the student of complexity needs to simplify by focusing on key features. Any good theory abstracts from many real issues, thereby to highlight the central forces of action.

We need to ask, therefore, after understanding the core forces that the theory exposes, whether it can be taken back or is limited to the caricature that it depicts. With respect to vertical integration, the question is this: Does repeated application of the simple model of vertical integration (or not) of two successive stages scale up to approximate a firm where a succession of such interfaces are addressed?

This question is rarely addressed by the new theories of economic organization referred to above. Interestingly, the most prominent of Holmstrom’s criticisms of PRT is the inability of PRT to scale up as we suggest. As we quoted him earlier, “The problem is

that there really are no firms in these [formal] models, just representative entrepreneurs.” (1999, p. 100). This critique is echoed as well by Holmstrom and Roberts (1998) who write that “firms are poorly defined in property rights models and it is not clear how one actually should interpret the identities of [the parties].” (p. 79).¹⁵ In our view though, Holmstrom’s critique is equally valid towards much of the Agency Theory models of organization because these models are concerned more with incentives and task allocations than with how to coordinate transactions. Hence, it is hard to see how one takes the results and conclusions of PRT to real world firms, either to explain or to prescribe the multitude of organizational design choices that are made by private sector firms and public sector bureaucracies. Similarly, as informative as Agency theory might be for the design of incentive schemes and task allocations to employees, it sheds little light on how to organize transactions within and between firms.

As for TCE, the importance of scaling up is expressly admitted and simple efforts to describe the boundary of the firm in TCE terms have been made (Williamson, 1985, pp. 96-98). Unlike PRT or Agency theory, where individuals are at the center of attention, TCE deals with stages of production. TCE asks first that technologically separable boundaries be identified, and with such identification, TCE considers which of a series of interfaces should be controlled within a firm’s hierarchy by a series of interface coordinators (some of whom might be combined) , and which should be

¹⁵ Holmstrom and Roberts (1998) take aim at both TCE and PRT as not being able to explain a host of anecdotes that they bring forth, mostly as examples of hybrid governance that are neither integrated nor simple market exchanges. Two comments are in order to correct some misconceptions in Holmstrom and Roberts (1998). First, they take TCE’s main objective to be one that solves the hold-up problem, which is not the case. As we observe earlier, TCE is mainly considered with managing adaptation. Second, they focus mostly on ownership of assets as the vehicle of control for both PRT and TCE. We have illustrated that there is much more to it than mere ownership. In fact, TCE recognizes that there are many more instruments than ownership, since financial incentives, administrative control, and legal rule all play a role in TCE’s characterization of market verses hierarchy.

procured through the market. This logic does not only apply to production stages, such as different subcomponents of the firm's primary products, but it applies also to backward integration into raw materials, forward into sales, and marketing, and to services as well.

5. Concluding Remarks

We have tried in this paper to offer some perspective into the development of research on the economics of organization starting with Coase's seminal paper, through the ideas of TCE that were put forth in the 1970s, and ending with the formal mathematical approaches offered by PRT and Agency Theory. Our critical take on the progression of this still thriving field of study is that a theory's success primarily rests on three elements of its construction: the choice of its focal transaction, the way in which it lends itself to empirical testing, and its ability to scale up to explain reality.

We have argued that among the theories that have been put forth, TCE compares favorably on all three fronts. We believe that TCE's success is due in part to its focus on the intermediate product market transaction. But there is more: TCE delivered both empirical testing and scaling up advantages precisely because it went beyond boundaries by prominently featuring the problem that organization sets out to address, namely, the management of adaptation.

Holmstrom's observation is pertinent: whereas PRT "focuses all the attention on the question of boundaries, which admittedly has been the outstanding puzzle to solve ever since Coase raised it more than 50 years ago ..., we also have to ask why Coase thought (or we think) that the question of boundaries is such an interesting and essential

one to address.” (p. 100). The boundaries question for TCE mainly arises in conjunction with scaling up.

TCE, is still a work-in-progress and stands in need of more formal mathematical modeling that will help refine its messages and bring it to a more mature stage in the “natural progression” of theory, a point that has been made by Williamson (2005). As he writes, “the economics of governance has moved through a natural progression from informal theory (where the early intuitions reside) to pre-formal theory (where the basic logic is set out) to semi-formal theory (of a reduced-form kind) to fully formal theory. Ideally, value is added at each step.” (2005, p.11). It is our conviction that more formal theories of TCE are called for to continue the desired progression.

The issues discussed herein are intended to provide added perspective for understanding complex economic organization. Our treatment is suggestive rather than definitive and, as such, invites follow-on work on these issues. We nevertheless conjecture that similar issues to those described above are posed when assessing the efficacy of other would-be theories within economics and the contiguous social sciences more generally.

References

Alchian, Armen, and H. Demsetz. (1972) "Production, Information Costs, and Economic Organization," *American Economic Review*, **62(3)**: 777-795.

Arrow, Kenneth J. (1969) "The Organization of Economic Activity: Issues Pertinent to the Choice of Market versus Nonmarket Allocation." In *The Analysis and Evaluation of Public Expenditure: The PPB System*, pp. 59–73. Washington, DC: US Government Printing Office.

Bajari, Patrick and Steven Tadelis (2001) "Incentives versus Transaction Costs: A Theory of Procurement Contracts." *Rand Journal of Economics*, **32(3)**:287–307.

Barnard, Chester. (1938) *The Functions of the Executive*. Cambridge: Harvard University Press (fifteenth printing, 1962).

Brown, Trevor L. and Matthew Potoski (2003) "Transaction Costs and Institutional Explanations for Government Service Production Decisions," *Journal of Public Administration Research and Theory*, **13(4)**:441-468

Coase, Ronald H. (1937) "The Nature of the Firm," *Economica* **4**:386-405.

Coase, Ronald H. (1960) "The Problem of Social Cost," *Journal of Law and Economics*, **3(1)**:1–44.

Coase, Ronald H. (1972) "Industrial Organization: A Proposal for Research," in V. R. Fuchs, ed., *Policy Issues and Research Opportunities in Industrial Organization*. New York: National Bureau of Economic Research, pp. 59-73.

Coase, Ronald (1993) "Coase on Posner on Coase," *Journal of Institutional and Theoretical Economics*, **149(1)**:96-98.

Demsetz, Harold. (1983) "The Structure of Ownership and the Theory of the Firm," *Journal of Law and Economics*, **26(1)**:375-390.

Dixit, Avinash. *The Making of Economic Policy: A Transaction Cost Politics Perspective*. Cambridge, MA: MIT Press, 1996.

Gatignon, Aline and Hubert Gatignon (2010) "Erin Anderson and the Path-Breaking Work of TCE in New Areas of Business Research: Transaction Costs in Action," *Journal of Retailing*, **86(3)**:232-247.

Goldberg, Victor P. (1976) "Regulation and Administered Contracts," *Bell Journal of Economics*, **7(2)**:426-448.

Grossman, Sanford and Oliver D. Hart (1986) "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy*, **94(4)**:691-719.

Hart, Oliver D. and John Moore (1990) "Property Rights and the Nature of the Firm," *Journal of Political Economy*, **98(6)**:1119-1158.

Holmstrom, Bengt and Paul R. Milgrom (1991) "Multitask Principal-Agent Analyses: incentive Contracts, Asset Ownership, and Job Design," *Journal of Law, Economics and Organization*, **7(0)**:24-51.

Holmstrom, Bengt and Paul R. Milgrom (1994) "The Firm as an Incentive System," *American Economic Review*, **84(4)**:972-91.

Holmstrom, Bengt and John Roberts (1998) "The Boundaries of the Firm Revisited," *Journal of Economic Perspectives*, **12(4)**:73-94.

John, George and Torger Reve (2010) "Transaction Cost Analysis in Marketing: Looking Back, Moving Forward," *Journal of Retailing*, **86(3)**:248-256.

Klein, Benjamin, Robert Crawford and Armen Alchian (1978) "Vertical Integration, Appropriable Rents, and the Competitive Contracting Process," *Journal of Law and Economics*, **21(2)**:297-326.

Levin, Jonathan D. and Steven Tadelis (2010) "Contracting for Government Services: Theory and Evidence from U.S. Cities," *Journal of Industrial Economics*, **58(3)**:507-541

Levy, Brian and Pablo Spiller (1994) "The Institutional Foundations of Regulatory Commitment: A Comparative Analysis of Telecommunications Regulation." *Journal of Law, Economics and Organization*, **10(1)**:201-46.

Macher, Jefferey T. and Barak D. Richman (2008) "Transaction Cost Economics: An Assessment of Empirical Research in the Social Sciences" *Business and Politics*, 10(1)

Jackson A. Nickerson and Brian S. Silverman (Eds.) *Economic Institutions of Strategy*, Advances in Strategic Management (Vol. 26), Emerald Group Publishing Limited, 2009

Riordan, Michael and Oliver E. Williamson (1985) "Asset Specificity and Economic Organization," *International Journal of Industrial Organization*, 1985, **3(4)**:365-78.

Tadelis, Steven (2002) "Complexity, Flexibility and the Make-or-Buy Decision," *American Economic Review Papers and Proceedings*, **92(2)**:433-437

Tadelis, Steven (2007) "The Innovative Organization: Creating Value Through Outsourcing," *California Management Review*, **50(1)**:261-277

Whinston, Michael (2001) "Assessing the Property Rights and Transaction-Cost Theories of Firm Scope," *American Economic Review Papers and Proceedings*, **91(2)**:184-188.

Williamson, Oliver E. (1975) *Markets and Hierarchies: Analysis and Antitrust Implications*. New York: Free Press.

Williamson, Oliver E. (1979) "Transaction-Cost Economics: The Governance of Contractual Relations." *Journal of Law and Economics*, **22(2)**:233-61.

Williamson, Oliver E. (1983) "Credible Commitments: Using Hostages To Support Exchange," *American Economic Review*, **73(4)**:519-540.

Williamson, Oliver E. (1985) *The Economic Institutions of Capitalism*. New York: Free Press.

Williamson, Oliver E. (1999) "Public and Private Bureaucracies," *Journal of Law, Economics, and Organization*, **15(1)**:306-342.

Williamson, Oliver E. (2005) "The Economics of Governance," *American Economic Review Papers and Proceeding*, **95(2)**:1-18.