

Boundaries of the Firm (Updated: Jan 25 2017)

The central question in this part of the literature goes back to Ronald Coase (1937): if markets are so great at coordinating productive activity, why is productive activity carried out within firms rather than by self-employed individuals who transact on a spot market? And indeed it is, as Herbert Simon (1991) vividly illustrated:

A mythical visitor from Mars... approaches Earth from space, equipped with a telescope that reveals social structures. The firms reveal themselves, say, as solid green areas with faint interior contours marking out divisions and departments. Market transactions show as red lines connecting firms, forming a network in the spaces between them. Within firms (and perhaps even between them) the approaching visitor also sees pale blue lines, the lines of authority connecting bosses with various levels of workers... No matter whether our visitor approached the United States or the Soviet Union, urban China or the European Community, the greater part of the space below it would be within the green areas, for almost all inhabitants would be employees, hence inside the firm boundaries. Organizations would be the dominant feature of the landscape. A message sent back home, describing the scene, would speak of “large green areas interconnected by red lines.” It would not likely speak of “a network of red lines connecting green spots.” ...When our visitor came to know that the green masses were organizations and the red lines connecting them were market transactions, it might be surprised to hear the structure called a market economy. “Wouldn’t ‘organizational economy’ be the more appropriate term?” it might ask.

It is obviously difficult to put actual numbers on the relative importance of trade within and between firms, since, I would venture to say, most transactions within firms are not recorded. From dropping by a colleague's office to ask for help finding a reference, transferring a shaped piece of glass down the assembly line for installation into a mirror, getting an order of fries from the fry cook to deliver to the customer, most economic transactions are difficult even to define as such, let alone track. But we do have some numbers. In what I think is one of the best opening sentences of a job-market paper, Pol Antras provides a lower bound: "Roughly one-third of world trade is intrafirm trade."

Of course, it could conceivably be the case that boundaries don't really matter—that the nature of a particular transaction and the overall volume of transactions is the same whether boundaries are in place or not. And indeed, this would exactly be the case if there were no costs of carrying out transactions: Coase's (1960) eponymous theorem suggests, roughly, that in such a situation, outcomes would be the same no matter how transactions were organized. But clearly this is not the case—in 1997, to pick a random year, the volume of corporate mergers and acquisitions was \$1.7 trillion dollars (Holmstrom and Roberts, 1998). It is implausible that this would be the case if boundaries were irrelevant, as even the associated legal fees have to ring up in the billions of dollars.

And so, in a sense, the premise of the Coase Theorem's contrapositive is clearly true. Therefore, there must be transaction costs. And understanding the nature of these transaction costs will hopefully shed some light on the patterns we see. And as D.H. Robertson also vividly illustrated, there are indeed patterns to what we see. Firms are "islands of conscious power in this ocean of unconscious co-operation like lumps of butter coagulating in a pail of buttermilk." So the question becomes: what transaction costs are important, and how are they important? How, in a sense, can they help make sense out of the pattern of butter and buttermilk?

The field was basically dormant for the next forty years until the early 1970s, largely because "transaction costs" came to represent essentially "a name for the residual"—any

pattern in the data could trivially be attributed to some story about transaction costs. The empirical content of the theory was therefore zero.

Williamson put structure on the theory by identifying specific factors that composed these transaction costs. And importantly, the specific factors he identified had implications about economic objects that at least could, in principle, be contained in a data set. Therefore his causal claims could be, and were, tested. (As a conceptual matter, it is important to note that even if Williamson's causal claims were refuted, this would not invalidate the underlying claim that "transaction costs are important," since as discussed earlier, this more general claim is essentially untestable, because it is impossible to measure, or even conceive of, *all* transaction costs associated with *all* different forms of organization.) The gist of his theory, which we will describe in more detail shortly, is that when contracts are incomplete and parties have disagreements, they may waste resources "haggling" over the appropriate course of action if they transact in a market, whereas if they transact within a firm, these disagreements can be settled by "fiat" by a mediator. Integration is therefore more appealing when haggling costs are higher, which is the case in situations in which contracts are relatively more incomplete and parties disagree more.

But there was a sense in which his theory (and the related work by Klein, Crawford, and Alchian (1978)) was silent on many foundational questions. After all, why does moving the transaction from the market into the firm imply that parties no longer haggle—that is, what is integration? Further, if settling transactions by fiat is more efficient than by haggling, why aren't all transactions carried out within a single firm? Williamson's and others' response was that there are bureaucratic costs ("accounting contrivances," "weakened incentives," and others) associated with putting more transactions within the firm. But surely those costs are also higher when contracts are more incomplete and when there is more disagreement between parties. Put differently, Williamson identified particular costs associated with transacting in the market and other costs associated with transacting within the firm and made assertions about the rates at which these costs vary with the underlying environment. The resulting

empirical implications were consistent with evidence, but the theory still lacked convincing foundations, because it treated these latter costs as essentially exogenous and orthogonal. We will discuss the Transaction-Cost Economics (TCE) approach in the first subsection.

The Property Rights Theory, initiated by Grossman and Hart (1986) and expanded upon in Hart and Moore (1990), proposed a theory which (a) explicitly answered the question of “what is integration?” and (b) treated the costs and benefits of integration symmetrically. Related to the first point is an observation by Alchian and Demsetz that

It is common to see the firm characterized by the power to settle issues by fiat, by authority, or by disciplinary action superior to that available in the conventional market. This is delusion. The firm does not own all its inputs. It has no power of fiat, no authority, no disciplinary action any different in the slightest degree from ordinary market contracting between any two people. I can “punish” you only by withholding future business or by seeking redress in the courts for any failure to honor our exchange agreement. This is exactly all that any employer can do. He can fire or sue, just as I can fire my grocer by stopping purchases from him or sue him for delivering faulty products.

What, then, is the difference between me “telling my grocer what to do” and me “telling my employee what to do?” In either case, refusal would potentially cause the relationship to break down. The key difference, according to Grossman and Hart’s theory, is in what happens after the relationship breaks down. If I stop buying goods from my grocer, I no longer have access to his store and all its associated benefits. He simply loses access to a particular customer. If I stop employing a worker, on the other hand, the worker loses access to all the assets associated with my firm. I simply lose access to that particular worker.

Grossman and Hart’s (1986) key insight is that property rights determine who can do what in the event that a relationship breaks down—property rights determine what they refer to as the residual rights of control. And allocating these property rights to one party or another may change their incentives to take actions that affect the value of this particular

relationship. This logic leads to what is often interpreted as Grossman and Hart’s main result: property rights (which define whether a particular transaction is carried out “within” a firm or “between” firms) should be allocated to whichever party is responsible for making more important investments in the relationship. We will discuss the Property Rights Theory (PRT) approach in the second subsection.

From a theoretical foundations perspective, Grossman and Hart was a huge step forward—their theory treats the costs of integration and the costs of non-integration symmetrically and systematically analyzes how different factors drive these two costs in a single unified framework. From a conceptual perspective, however, all the action in the theory is related to how organization affects parties’ incentives to make relationship-specific investments. As we will see, their theory assumes that conditional on relationship-specific investments, transactions are always carried out efficiently. A manager never wastes time and resources arguing with an employee. An employee never wastes time and resources trying to convince the boss to let him do a different, more desirable task.

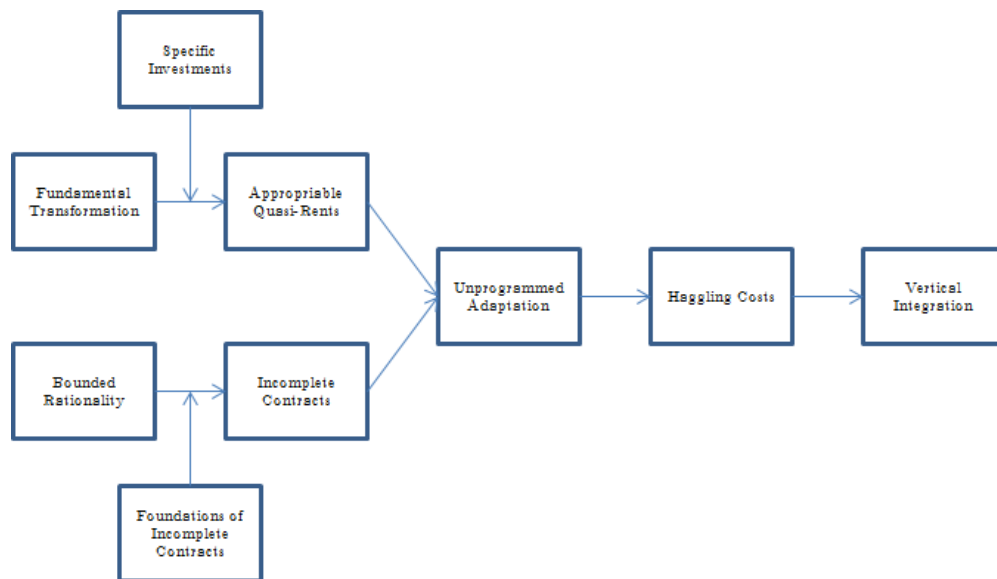
In contrast, in Transaction-Cost Economics, all the action takes place ex post, during the time in which decisions are made. Integration is chosen, precisely because it avoids inefficient haggling costs. We will look at some of the implications of this observation next week in the context of a model in which “haggling costs” are based on influence activities—costly activities aimed at persuading decision makers.

Finally, even the Property Rights Theory does not stand on firm theoretical grounds, since the theory considers only a limited set of institutions the players can put in place to manage their relationship. That is, they focus only on the allocation of control, ignoring the possibility that individuals may write contracts or put in place other types of mechanisms that could potentially do better. In particular, they rule out revelation mechanisms that, in principle, should induce first-best investment. We will address this issue in the sixth subsection.

As a research topic, the theory of the firm can be tricky to navigate. In contrast to many applied-theory topics, the theory of the firm takes a lot of non-standard variables as endogenous. Further, there is often ambiguity about what variables should be taken as endogenous and what methodology should be used, so the “playing field” is not well-specified. But ultimately, I think that developing a stronger understanding of what determines firm boundaries is important, since it simultaneously tells us what the limitations of markets are. I will try to outline some of the “ground rules” that I have been able to discern from spending some time studying these issues.

Transaction-Cost Economics (Updated: Jan 25 2017)

The literature on the boundaries of the firm introduces many new concepts and even more new terms. So we will spend a little bit of time sorting out the terminology before proceeding. The following figure introduces most of the new terms that we will talk about in this section and in the following sections.



As an overview, the basic argument of the Transaction-Cost Economics approach is the following.

Consider a transaction between an upstream manufacturer and a downstream distributor. Should the distributor **buy** from the manufacturer or should it buy the manufacturer and **make** goods itself? The **fundamental transformation** of ex ante perfect competition among manufacturers for the customer's business to ex post small numbers results *from specific investments* and results *in appropriable quasi-rents*—ex post rents that parties can potentially fight over, because they are locked in with each other. If the parties have written **incomplete contracts** (the **foundations** for which are that **bounded rationality** limits their ability to foresee all relevant contingencies), then they might find themselves in situations that call for **unprogrammed adaptation**. At this point, they may fight over the appropriate course of action, incurring **haggling costs**. These haggling costs can be reduced or eliminated if either the manufacturer purchases the distributor or the distributor purchases the manufacturer, and they become **vertically integrated**.

Every link in this figure is worth discussing. The **fundamental transformation** is, in my view, the most important economic idea to emerge from this theory. We had known since at least Edgeworth that under bilateral monopoly, many problems were possible (Edgeworth focused on indeterminacy) and perhaps inevitable (e.g., the Myerson-Satterthwaite theorem), but that competition among large numbers of potential trading partners would generally (with some exceptions—perfect complementarities between, say, right-shoe manufacturers and left-shoe manufacturers could persist even if the economy became arbitrarily large) push economies towards efficient allocations. Perfect competition is equivalent to the no-surplus condition (Ostroy, 1980; Makowski and Ostroy, 1995)—under perfect competition, you fully appropriate whatever surplus you generate, so everyone else in the economy as a whole is indifferent toward what you do. As a result, your incentives to maximize your own well-being do not come into conflict with others, so this leads to efficient allocations and nothing worth incurring costs to fight over. The underlying intuition for why large numbers of trading partners leads to efficient allocations is that a buyer can always play a seller and her competitors off each other, and in the limit, the next-best seller is just as

good as the current one (and symmetrically for buyers). Williamson's observation was that after a trading relationship has been initiated, the buyer and the sellers develop ties to each other (**quasi-rents**), so that one's current trading partner is always discretely better than the next best alternative. In other words, the beneficial forces of perfect competition almost never hold.

Of course, if during the ex ante competition phase of the relationship, potential trading partners competed with each other by offering enforceable, complete, long-term trading contracts, then the fact that ex post, parties are locked in to each other would be irrelevant. Parties would compete in the market by offering each other utility streams that they are contractually obligated to fulfill, and perfect competition ex ante would lead to maximized long-term gains from trade.

This is where **incomplete contracts** comes into the picture. Such contracts are impossible to write, because they would require parties to be able to conceive of and enumerate all possible contingencies. Because parties are **boundedly rational**, they will only be able to do so for a subset of the possible states. As a result, ex ante competition will lead parties to agree to incomplete contracts for which the parties will need to fill in the details as they go along. In other words, they will occasionally need to make **unprogrammed adaptations**. As an example, a legacy airline (say, American Airlines) and a regional carrier (say, American Eagle) may agree on a flight schedule for flights connecting two cities. But when weather-related disruptions occur, the ideal way of shifting around staff and equipment depends to a large extent on where both carrier's existing staff and equipment are, and there are simply too many different potential configurations for this. As a result, airlines typically do not contractually specify what will happen in the event that there are weather-related disruptions, and they therefore have to figure it out on the spot.

The need to make unprogrammed adaptations would also not be a problem if the parties could simply agree to bargain efficiently ex post after an event occurs that the parties had not planned for. (And indeed, if there was perfect competition ex post, they would not even

need to bargain ex post.) However, under the TCE view, ex-post bargaining is rarely if ever efficient. The legacy airline will insist that its own staff and equipment are unable to make it, so everything would be better if the regional carrier made concessions, and conversely. Such ex post bargaining inevitably leads either to bad ex post decisions (the carrier with the easier-to-access equipment and staff is not the one who ends up putting it in place) or results in other types of **rent-seeking** costs (time and resources are wasted in the bargaining process). These **haggling costs** could be eliminated if both parties were under the direction of a common headquarters that could issue commands and easily resolve these types of conflicts. This involves setting up a **vertically integrated** organization.

Further, vertically integrated organizations involve **bureaucratic costs**. Reorganization involves setup costs. Incentives are usually low-powered inside organizations. Division managers engage in **accounting contrivances** in order to alter decision making of other divisions or the headquarters. Finally, the contract law that governs decisions made by one division that affect another division differs from the contract law that governs decisions made by one firm that affect another—in essence, the latter types of contracts are enforceable, whereas the former types of contracts are not. This difference in contract law is referred to as **forbearance**.

When would we be more likely to see vertical integration? When the environment surrounding a particular transaction is especially complex, contracts are more likely to be incomplete, or they are likely to be more incomplete. As a result, the need for unprogrammed adaptations and their associated haggling costs will be greater. When parties are more locked in to each other, their ability to access the outside market either to look for alternatives or to use alternatives to discipline their bargaining process is lessened. As a result, there is more to fight over when unprogrammed adaptations are required, and their associated haggling costs will be greater. Additionally, integration involves setup costs, and these setup costs are only worth incurring if the parties expect to interact with each other often. Finally, integration itself involves other bureaucratic costs, and so vertical integration is more appealing if these

costs are low. Put differently, the integration decision involves a trade-off between haggling costs under non-integration and bureaucratic costs under integration. To summarize, the main empirical predictions of the TCE theory are:

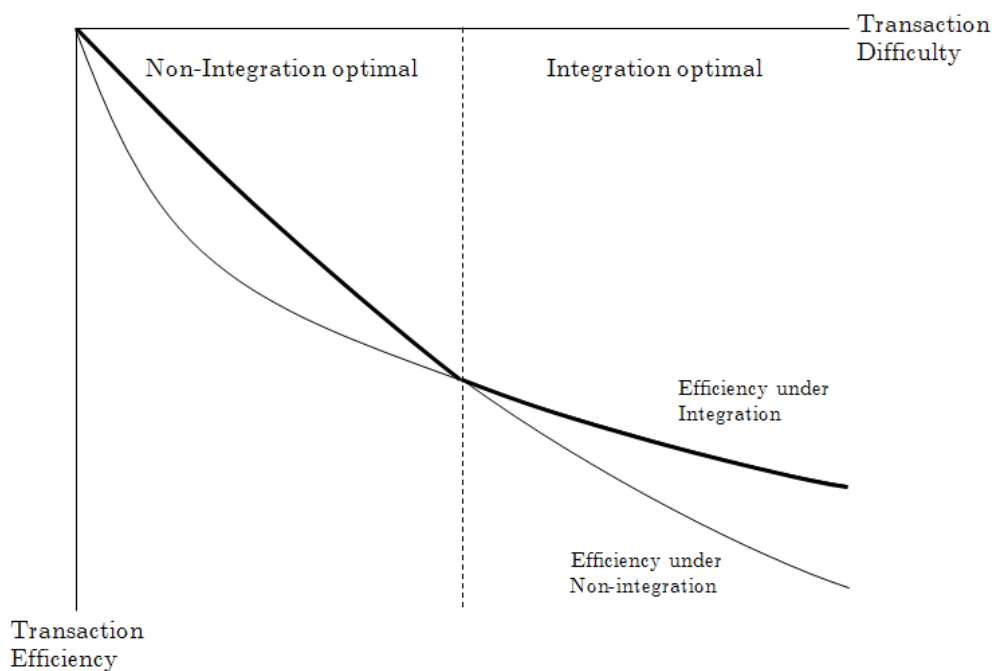
1. Vertical integration is more likely for transactions that are more complicated.
2. Vertical integration is more likely when there is more specificity.
3. Vertical integration is more likely when the players interact more frequently.
4. Vertical integration is more likely when bureaucratic costs are low.

This discussion of TCE has been informal, because the theory itself is informal. There are at least two aspects of this informal argument that can be independently formalized. When unprogrammed adaptation is required, the associated costs can either come from costly haggling (rent-seeking) or from inefficient ex post decision making (maladaptation). I will describe two models that each capture one of these sources of ex post inefficiencies.

There are a couple common themes that arise in the analysis of both of these models. The first theme is that when thinking about the make-or-buy or boundary-of-the-firm question, the **appropriate unit of analysis is at the transaction level**. Another theme is that these models consider **private-ordering solutions** rather than solutions imposed on the transacting parties by a third party such as a government—resolving conflict between parties need not involve government intervention. The question is therefore: for a given transaction, what institutions should *the interested parties* put in place to manage this transaction?

Since transactions differ in their characteristics, more difficult transactions will be more difficult no matter how they are organized. As a result, looking at the performance of various transactions and relating that performance to how those transactions are organized could lead one inappropriately to the conclusion that integration is actually bad for performance. Or one could dismiss the agenda, as one prominent economics blogger once did: “I view the Coasian tradition as somewhat of a dead end in industrial organization. Internally, firms

aren't usually more efficient than markets... .”



As the figure above (which Gibbons (2005) colorfully describes as “Coase meets Heckman”) shows, this is *exactly* the type of prediction that this class of theories predicts. And whether integrated transactions are less efficient, because integration is bad for transaction efficiency, or because transactions that are more complicated are more likely to involve integration and are likely to be less efficient matters. This difference matters, because these two different views have the opposite implications. Under the TCE view, discouraging firms from engaging in vertical integration (through, for example, strict antitrust policy) will necessarily be bad for firms’ internal efficiency. Under the alternative view, strict antitrust policy would serve not only to facilitate product-market competition, it would also *increase* firms’ internal efficiency.

Another theme that arises in these models is that there are many differences between transactions carried out between firms and those carried out within firms. An upstream division manager will typically be on a lower-powered incentive scheme than she would be if she were the owner of an independent upstream firm. Transactions within firms are subject

to different legal regimes than transactions between firms. Transactions within firms tend to be characterized by more “bureaucracy” than transactions across firms. There are two ways to look at these bundles of differences. Viewed one way, low-powered incentives and bureaucracy are the baggage associated with integration and therefore a cost of integration. Viewed another way, low-powered incentives and bureaucracy are also optimal choices that complement integration because they help solve other problems that arise under integration.

Finally, I will conclude with a description of one question that I do not think the literature has produced satisfying answers to. First, there are obviously more ways to organize a transaction than “vertical integration” and “non-integration.” In particular, the transacting parties could engage in simple spot-market transactions; they could engage in short-run contracting across firm boundaries in which they specify a small number of contingencies; they could engage in long-term contracting across firm boundaries in which perhaps decision rights are contractually reallocated (for example, an upstream firm may have some say over the design specifics for a product that the downstream firm is producing); or one party could buy the other party. The line between integration and non-integration is therefore much blurrier than it seemed at first glance.