chapter twenty

Transaction cost economics and the firm

This chapter continues the discussion initiated in chapter 19 about models of the firm, with consideration given to models where the firm is something more like a market than a consumer. That is, we think of the firm as an institution within which transactions between individuals take place — an alternative to transactions that take place in a market. To make sense of this, we have to see how a transaction placed within the context of a firm is different from the "same" transaction placed within a market, which is the subject of *transaction cost economics*.

This chapter also continues to prepare for our critique of the methods that have been developed throughout the book. To accomplish this second purpose, we take things in somewhat roundabout fashion. To begin, we recount the theory of transaction cost economics as told by Williamson (1985).[1] This recounting will seem more verbal and less mathematical than topics discussed earlier in the book, and after laying out the theory along the lines of Williamson (1985), we go back and see how we can match pieces of that theory with things we did earlier.

20.1. Transaction cost economics and firms

Transaction costs

When undertaking a transaction, parties to the transaction must incur several sorts of costs. *Ex ante costs* are incurred before the transaction takes place. If the transaction is to be governed by a written contract, the contract must be drafted. Whether governed by a contract or simply by verbal commitments, the terms of the transaction must be negotiated. *Ex post costs* are incurred in consummating and safeguarding the deal that was originally struck.

---

[1] The rendition given here is not precisely as in Williamson (1985) in emphasis or, in some minor respects, in organization. And the rendition here is not nearly as rich in examples and elaborations as is the work being abridged.

In some cases these costs are negligible. But in other cases they can be quite substantial. Insofar as transactions can be arranged in different ways (making use of different legal and social institutions, providing more or less detail in a contract, reserving rights to the transacting parties in one fashion or another), these different ways will have distinct costs. The basic notion of transaction costs economics is that *transactions tend to be "placed" in a way that maximizes the net benefits they provide, including the costs of the transaction.* In particular, a transaction whose (transaction) costs outweigh the benefits of completion will not be undertaken at all.

To get a better fix on these costs, it is helpful to look at factors that give rise to them. Williamson compiles a list of these factors in two parts: factors pertaining to the individuals who undertake the transaction, and factors specific to the particular transaction.

Human factors

Human beings are *boundedly rational* and *opportunistic*. Bounded rationality is important because, in the first place, it means that it will be costly for individuals to contemplate and contract for every contingency that might arise over the course of the transaction; this adds to the ex ante cost of drafting the contract. These costs may be so high that the individuals fail to provide for the contingency in the contract or fail to undertake the contemplation necessary to foresee the contingency. And there may be contingencies that the individuals cannot foresee at all.[2] Contingencies that are not provided for ex ante may add to ex post costs, since parties to the transaction may be required to negotiate further whenever such contingencies arise. And, insofar as it is understood that some contingencies are not foreseen or provided for ex ante, the parties may build into the original contract specific means by which the contract will be amended as required. These "means" or *governance structures* bring on costs of administration and the like.

To say that individuals are *opportunistic* means that they are self-interested with guile. If it will further his own weal, an opportunistic individual will break any of the commandments. Note well the conditional *if* here. Institutions can be sometimes arranged so that when an individual gives his "word of honor," then it will not be in his interests to go back on that word to extract some short-term gain. (Think back to chapter 14 or see below if it isn't obvious to you how this can happen.)

---

[2] We will distinguish between contingencies that are unforeseeable and those that could be foreseen but that are not foreseen because of the costs of contemplation. From a theoretical point of view, these two categories could be combined by saying that the cost of foreseeing the first sort is infinite.

To distinguish simple self-interest from opportunism, think of a completely honest individual who would never break her word or misrepresent what she knows, but who still seeks to maximize her own welfare. This is self-interest, as compared to an opportunistic individual who would break his word or engage in misrepresentation under the right circumstances. Moreover, our use of the term "opportunism" is stretched to mean that it is opportunistic to refuse to divulge information that you hold and another lacks when the other person asks you to give up that information. That is, if you are selling a used car whose quality you know and a buyer asks you what the quality is, even to withhold this information would be opportunistic behavior.ᵃ

Classify individuals according to a three-by-three scheme, where the first dimension is the individual's degree of rationality and the second is the individual's self-interest orientation.

For degree of rationality, our three categories are complete rationality, bounded rationality, and behavioral. A completely rational individual has the ability to foresee everything that might happen and to evaluate and optimally choose among available courses of action, all in the blink of an eye and at no cost. A boundedly rational individual attempts to maximize but finds it costly to do so and, unable to anticipate all contingencies and aware of this inability, provides ex ante for the (almost inevitable) time ex post when an unforeseen contingency will arise. A behavioral individual acts according to some specified behavioral pattern that doesn't (except by coincidence) correspond to the maximization of any specific utility or welfare function.

For the individual's self-interest orientation, to the opportunist and the completely honest but self-interested individuals discussed, add the utopian, someone who has a sense of the social good and seeks to maximize it.

Various pieces of economic theory can be thought of as concerning societies comprising one of the nine classes of individuals so created. If the individuals in a society are behavioral, then their self-interest orientation is irrelevant once their behavior is specified, since (by definition) they don't act out of self-interest at all.³ Also irrelevant to such individu-

---

ᵃ This stretches the notion of opportunism very far indeed, far beyond its most "natural" usage. To amplify, suppose A owns a piece of land on which is buried some treasure. A doesn't know where the treasure is, and it is too expensive to dig up the entire plot to find it. B knows where the treasure is, however, and with that knowledge it would be worthwhile for B (or, if A had the knowledge, for A) to dig. According to this definition of opportunism, it would be opportunistic of B to refuse to tell A where the treasure is, if A asks.

³ While correct as stated, this is a bit misleading. Reasonable models of behavior will reflect the self-interest orientation that is supposed of the individual. For example, firms

als is all of orthodox economic theory, which is based on consumers who maximize. Evolutionary theories, such as in Nelson and Winter, but applied to the actions of consumers instead of the actions of firms, would seem to be appropriate to the economic analysis of a society of such individuals.

Most economic theory concerns completely rational individuals. If they are utopian in their interest orientation, then one employs *team theory*, in which (it is assumed) all individuals have the same utility function (which can be thought of as social welfare). They may not have common information, which makes the subject interesting, but they all act in a way that anticipates perfectly what their fellows will do contingent on the information their fellows possess, and they act to maximize some single criterion of overall welfare.[4]

Completely rational individuals who are self-interested but without guile populate the economies of general equilibrium. There is no mention of deceit or of private information. Since self-interested individuals who are not opportunistic will not withhold pertinent information, and since each side to the transaction is completely rational and so anticipates the possibly pertinent information that the other side possesses, there will be no pertinent private information at all (after a round of questions and honest answers). In undertaking a transaction, therefore, parties know just what they are getting, and any allowance for future contingencies is complete and completely understood by all.[b]

Completely rational and opportunistic individuals populate the models of moral hazard and incentives and adverse selection and market signaling; i.e., the stuff of part IV. One is able to work through how others will act and react given appropriate incentives. And one is able to work through the distribution of qualities that will be brought to market, or the distribution of qualities that accompanies some given signal or other, or the appropriate (equilibrium) action to take when faced with a menu of contracts. But the incentives that are put in place (either to take a particular action or to send a particular signal) must respect the ability of the individuals to lie, cheat, and steal if so doing would be beneficial.

---

in the simulation model of section 19.3 are "behavioral," but their behavior is tied to their "self-interest" as measured by relative profitability. Similarly, we could imagine a model of individual behavior where the individual responds to prompts that are tied to her "self-interest" such as her levels of consumption of various goods relative to those of her neighbors, and so on.

[4] For an exposition of team theory, see Marshak and Radner (1972).

[b] In general equilibrium, different firms are allowed to have different production functions. So, as a matter of the formal theory, the fact that one firm can undertake a particular production plan that another cannot is *not* based on proprietary information held by the first.

Self-interest orientation

	Utopian	Self-interested without guile	Self-interested with guile
Behavioral	Evolutionary methods(?)		
Degree of Bounded rationality	Team theory	Temporary equilibrium	Transaction cost economics
Complete		General equilibrium	Information economics

Figure 20.1. Matching economic theories to types of individuals.

This leaves boundedly rational individuals. For boundedly rational individuals who are utopian, some pieces of team theory (Marshak and Radner, 1972) apply. For boundedly rational individuals who are self-interested but without guile, the literature of *temporary equilibrium* applies. (See Grandmont [1988].) And, finally, boundedly rational individuals who are self-interested with guile populate the world of transaction cost economics.

Figure 20.1 (loosely adapted from Williamson [1985, fig. 2-1]) presents all this schematically. (The question mark attached to evolutionary methods records the relative lack of work along these lines. The usefulness of evolutionary methods for exploring these categories seems clear, but more evidence is desirable.)

Qualities of transactions

Bounded rationality and opportunism don't come to much when the sort of transaction that is contemplated is something like an exchange of apples for oranges or for money. At least, this is so if we imagine that the buyer of the apples/oranges is sufficiently well versed in these matters to be able to tell the quality of a piece of fruit upon quick inspection. It is the conjunction of the human factors discussed above and various aspects of the specific transaction that lead to significant transaction costs. Williamson identifies three aspects or qualities of a transaction that bear on the level and nature of transaction costs: *asset specificity, extent of uncertainty,* and *frequency*.

A transaction has high levels of *asset specificity* if as the trade develops one side or the other or both becomes more tied to and in the "power" of the other side. A simple example is a company that makes glass bottles locating a plant adjacent to a bottler. Before the bottle maker puts his plant next to the bottler, each side can (potentially) deal with many alternate traders. If there is negotiation between bottle manufacturer and bottler

before any plants are built, we would expect (if there are many of each type) that the deal struck will reflect other opportunities that each side has. The technology of the two production processes is such that there are efficiency gains if the two locate side by side, and we expect the two to divide, according to some bargaining scheme, the efficiency gains that come from side-by-side plants. Note well that in ex ante negotiations, if a bottler doesn't like the deal that a specific bottle manufacturer proposes, she can (typically) turn to many other bottle manufacturers. But once the bottle manufacturer puts his plant next to the bottler and the bottler puts her bottling lines next to the bottle manufacturing plant, each side has specific assets at risk. Now each side has a degree of monopoly power against the other; opportunism has scope to operate. The bottler, for example, might tell the bottle manufacturer that despite the contract signed earlier she wants a lower price per bottle. The bottle manufacturer, held up in this fashion, doesn't have any as-good alternative trading partners. So, anticipating this possibility, the bottle manufacturer (and, for that matter, the bottler) may expend resources to negotiate a very rigid contract ex ante and to have the ability to enforce that contract ex post. Or, to take a slightly less oportunistic example, the bottler and bottle manufacturer may not foresee, ex ante, the impact that plastics will have on their two industries. But the two will have to adapt their transactions as the technology for plastic containers develops, and their proximity will mean that each is somewhat a captive of the other in any ex post (re)negotiations.[5]

In the story above, asset specificity is at work on both sides of the transaction. There will also be cases in which asset specificity binds only or mainly on one side. To take an example, consider the plight of the graduate student. After a year or two of study, the student has invested enormously in assets that are specific to the department at which she studies, such as meeting specific requirements, passing specific exams, and so on. Some of this may be transferable to another department, but much is not. The department, on the other hand, has much less invested in the particular student. (When the student proceeds to the dissertation stage, this balance is redressed to some extent.) What began as an exchange in which each party had many alternative trading partners becomes one in which one side has much more at risk. In such cases, relatively more resources will be expended in the form of transaction costs to safeguard the individual student from being exploited by professors and administrators; e.g., formal structures are set up by which the student can appeal unjust

---

[5] For this to be an example of asset specificity, it is important that the old assets put in place earlier are not rendered completely obsolete by the new technology.

decisions to department chairpersons, deans, and the like; and both within the university and in the courts the burden of proof rests more on the university than on the student. [c]

The second quality of a transaction that bears on its costs is the extent of *uncertainty* in the transaction. This goes hand in hand with bounded rationality. Indeed, uncertainty is the major complexity that gives rise to bounded rationality. Note that uncertainty here is defined very broadly. It includes uncertainty about contingencies that can be anticipated, insofar as it is costly to anticipate them or to include provisions for dealing with them within a contract. It includes uncertainty about contingencies the nature of which can only roughly (or not even roughly) be divined ex ante. And it includes uncertainty of the sort where one party has information that the other lacks.

Finally, there is the *frequency* of the transaction. This aspect of the transaction does not bear on the absolute magnitude of its costs, as with the previous two aspects, but rather on the relative costs of various means for dealing with the transaction. When a transaction between two parties recurs frequently, the two parties can construct special governance structures for the transaction, even if these special structures are costly, since the cost of the structure can be amortized over many transactions. But when the transaction is a one-time-only transaction or recurs only infrequently, then it is generally more costly to put into place specialized mechanisms for this particular transaction and relatively less costly to make use of "general purpose" governance structures, which are, perhaps, less than ideally tuned to the specific transaction.

Classifying transactions by governance provisions

Transaction cost economics lays great stress when classifying different forms of transaction on the way in which the terms of the transaction are adapted to circumstances as they arise. These features of a transaction are called the *terms of governance*. Terms of governance can be explicitly and rigidly specified within a contract that governs a transaction, for example, an explicit and formalized procedure for arbitration, as in major league baseball. Or the terms of governance can be implicit, arising from common practice and law. For example, ownership of an asset generally

---

[c] In relations between the courts and universities, private universities are usually allowed to pursue consistent, so-called "dean's justice," where the decision of the dean is without serious appeal on procedural grounds. The notion is that if the university follows this practice consistently, then the student should know of it at the inception of the relationship, and the courts have no reason to interfere in a private transaction. Of course, this leaves the question why students would ever trust to dean's justice, a point we will return to in the next section.

confers upon the owner the right to command the use of that asset, within limits. Hence, when a bottler purchases from a bottle manufacturer the bottle manufacturing plant and simultaneously hires the manufacturer to staff the facility, the bottler has different "rights" under law than if she simply contracts to buy bottles from a manufacturer who owns the bottle manufacturing plant. The "contract" between the two explicitly specifies ownership of the assets, and the pattern of ownership together with the law implicitly specifies the resulting governance provisions.

Williamson (1985, chap. 3) gives the following classification scheme:

Transactions within the framework of *classical contracting* are those in which the terms of the transaction are completely specified ex ante. This includes the textbook exchange of apples for oranges, but goes beyond this to include any contract where adaptation beyond the explicit terms of the contract is not expected. For example, a complex purchase and sale agreement for a given piece of real estate would be included within this category.[6] Such P&S agreements typically include very specific clauses pertaining to liquidated damages: Party 1 receives such and such if party 2 fails to perform in such and such a manner. This sort of explicit, ex ante provision for nonperformance is typical of classical contracts whenever nonperformance is an issue. Of course, enforcement of the contract remains a problem; one can rely upon bonded third parties, who act with very little discretion (such as escrow agents), and in the end (and often with this sort of contract, from the beginning) the courts are called in to adjudicate disputes.

As third parties who act with discretion are added, we move into the realm of *neoclassical contracting* or *trilateral relationships*. The contract no longer says what damages are due for various sorts of breach or, more generally, what sort of adaptation will be made in various contingencies, but instead prescribes a third party who will determine appropriate damages/adaptation, according, perhaps, to some specified procedure. For example, in arbitration one sometimes sees a scheme in which both parties to the dispute submit their "claims" between which the arbiter is compelled to choose; the arbiter may not propose a compromise.

---

[6] A difficulty in providing illustrative examples in this chapter is that social customs often decide the form that a transaction takes. Real estate transactions, for example, are conducted quite differently in different states of the United States, and it is easy to envisage a society in which the form of real estate transactions is trilateral (see below), with the lending authority taking the role of the adjudicating third party. The examples that are used throughout this chapter are drawn from customs and practices that are prevalent in the United States, and they may not be descriptive of other countries. The obvious challenge in such cases is to explain in terms of the theory why there are differences. In general, cross-society comparisons of institutional practices is a very fertile field for empirical tests of the theory described here.

When the two parties to the transaction have no formal agreement about how the arrangement will be adapted to circumstances but instead rely upon their own ability to work things out as they go along, we have a *bilateral relationship*. Successful bilateral relationships remind one of cooperation in the repeated game version of the prisoners' dilemma: Each side is willing to cooperate with the other in order to preserve cooperation. (This connection will be developed more fully in the next section.)

Just as many sorts of trilateral arrangements can be made, so there are many forms of bilateral contracting. One extreme form deserves special attention. In a *hierarchical transaction* one of the two parties retains, by law or by custom, most of the authority to determine how the contract will be fulfilled. The second party will retain some explicit rights, such as to abrogate the contract, perhaps at some specified cost. And certain rights are implicitly retained under law. But up to such limits, the first party or *hierarchical superior* determines how matters will proceed. The chief example of this form, and the reason for its importance, is the classic labor contract: A worker earns his wages by carrying out the demands of his boss, retaining (in economies where slavery and indenture are illegal) the right to quit the job if these demands become too onerous or distasteful (cf. Simon, 1951).

When one party to the transaction takes command of the assets of the second, effectively internalizing the transaction, we have a *unified governance structure*. Here the focus is on the fact that "ownership" connotes control according to common practice. Note that where slavery and indenture are illegal, one party cannot buy the human capital of a second party; employment of a worker by a firm is not a transaction with a unified governance structure, but is instead a hierarchical transaction (if not some other form of transaction).

Matching transactions to governance structures

What sort of governance structure will minimize transaction costs in specific instances? Williamson advances the following scheme: Transactions are classified according to the specificity of the assets involved (from nonspecific through intermediate specificity to highly specific) and the frequency of the transaction (from occasional to frequent).

If assets are nonspecific, then there is no need for any fancy form of governance, since there is no need for complex, long-term contracts. Even if the relationship between the two parties is enduring, the relationship can be governed by a sequence of short-term contracts, since competition from the marketplace will prevent either individual from taking too great an advantage of the other. Classical contracting, then, goes with nonspecific

assets. It is when assets become more specific that the fancier forms come into play, if at all.

> Two caveats to this assertion should be offered. First, this presumes that the extent to which the two parties face competition in the relationship is unchanging with time. We can imagine a situation in which A has a lot of competition initially for B, but that competition disappears through time, even though no physical specific assets are put in place. To keep the simple conclusion just given while accommodating such a case, we would need to expand the definition of specific asset to include "market power."
>
> Second, we earlier mentioned a purchase and sale agreement for the purchase of real estate as an example of classical contracting. But when buying and selling real estate, the parties involved commit quite significant specific assets to the transaction, namely their opportunity to find in timely fashion an alternative partner should the current deal go awry. What keeps this within the domain of classical contracting is that the transaction is one with which there is a lot of experience. A real estate lawyer may be required to explain to you all the contingencies for which you must make allowance, but the extent of those contingencies is fairly well known; "boilerplate" contract forms can often be used. From the point of view of the match being developed here, what goes wrong in this case is that we do not have a third axis for "uncertainty" or "complexity." The usual real estate transaction may be very occasional (for the parties involved) and may involve substantial specific assets (for a short period of time), but the range of possible contingencies is well known; it is not complex in the relevant sense. (Or, alternatively, you can think of this as a case in which the extent of rationality is substantial, at least if the parties take the prudent course of action and employ specialists to help draft the contract.) Hence this sort of transaction can be handled by a classical contract.
>
> Similar qualifications can be made to almost every piece of the match between transactions and governance structures that we lay out. We will refrain hereafter from making these qualifications, but be clear that this is a very simplistic scheme. It gives general tendencies and it should be qualified in many ways.

When assets are moderately specific, relational contracting (i.e., trilateral, bilateral, hierarchical) comes to the fore. It is the pattern of the frequency of the transaction that plays the principal role in determining which relational-contracting form is best.[7] If the given transaction between the two parties is repeated between them frequently, then full bilateral contracting can work well. If one party of the two engages in this sort of transaction relatively frequently (and the second party engages in

---

[7] For frequency of the transaction, you should have in mind the frequency of the specific transaction or of transactions of a roughly similar nature. Anticipating what will come later, if two parties never repeat a precise transaction but frequently repeat transactions sufficiently similar to one another so that parties can carry an overall reputation for behavior, then we would essentially have a case of frequent transactions.
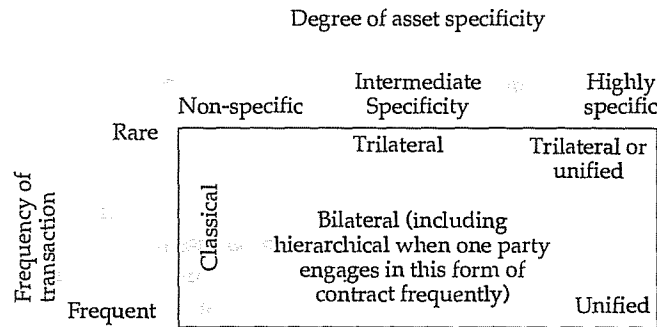
Degree of asset specificity

		Non-specific	Intermediate Specificity	Highly specific
Rare			Trilateral	Trilateral or unified
Frequent	Classical		Bilateral (including hierarchical when one party engages in this form of contract frequently)	Unified

Figure 20.2. Matching governance structures to characteristics of the specific transaction.

(y-axis label: Frequency of transaction)

the transaction with less frequency), then we may be able to get by with a bilateral form that gives most of the discretion to the first party; i.e., a hierarchical form may be appropriate.[8] But as the frequency of the transaction for either of the two parties decreases, bilateral forms in general begin to suffer, and it may become necessary to enlist the aid of a third party who can act in an adjudicary role.

When assets become very specific to the transaction, the costs of relational contracting rise: Each party has more at risk, and so must engage in more and more pretransaction planning, during-transaction monitoring, and posttransaction enforcement. A point is reached, then, where the costs of relational contracting become quite high, and the parties consider, if the laws permit, a unified governance structure, where one party buys out the other and takes full command of and responsibility for the transaction.

All these considerations lead to the picture in figure 20.2, adapted with minor modifications from Williamson (1985, fig. 3-2).

Firms and markets

The connection from the theory just developed to the theory of the firm is relatively straightforward. A firm corresponds to unified governance; a firm is a legal entity that commands an array of assets and in whose name various transactions are consummated (with other firms and with individuals).[9]

---

[8] We will see one reason why in the next section.

[9] We will not be very precise here about the differences between corporations, proprietorships, partnerships, and so on. These differences can be significant, but for a first cut at the theory they are best ignored.

Why does this make a firm into something in the category of a market? After all, the preceding sentence has firms acting in markets, playing the sort of role that is played by the individual consumer. Think of it in terms of the following example: Party A and party B wish to undertake an exchange. Party A will supply expertise concerning the design of some product, and B will supply expertise in manufacturing the product. This transaction requires certain tools — computers to aid in the design, and industrial lathes and various jigs to produce the product. We could imagine A owning the computers and B owning the lathe and jigs, in which case we would think of the exchange between them as a market-mediated exchange. Alternatively, we can imagine (say, because once B owns the jigs, he doesn't need many of the design capabilities of A) that A owns the computers, the lathes, and the jigs; she then controls a firm, which employs B for his labor services. The exchange of labor services for money is a firm-mediated transaction; as distinct from (say) the market-mediated exchange of a product design for money where B continues to own the lathes and jigs, or the market-mediated exchange where A obtains from B finished pieces produced according to plans given to B by A, which A then sells to the consumer. If A controls all the assets and employs B, it isn't important at a first level of analysis whether A owns the assets directly or controls the assets through some form of corporation; it is A's control of the assets that changes the nature of the transaction.[10]

> In this example, if the design and/or the production technique is proprietary then perhaps the most important asset of all will be the design or the technique or both. And ownership of this asset is the most complex thing of all. If A owns all the physical capital and also the right to block others from using the design, so she controls the production amounts, and if B supplies labor services and is disbarred from taking the knowledge he acquires about the production technique to form another firm or to some competitor, then we have a very clean case of a firm (controlled by A) employing B's labor services. An equally clean case would be where B owns all the physical capital and the design, and he has the right to employ another designer who can modify A's original design, while A is barred from using the design or a modification of it. Now imagine a case where the two work across a market interface, with A owning the computer and B the machine tools. Who owns the design? Who controls the amount produced of the good? (If B purchases the design from A, then presumably B. But if B acts as a subcontractor to A, then A has this control.) Who controls the ability to modify the design? (If A sells "turn-key" jigs, then this could be A; if A sells the details of the design outright, then perhaps both have this right, depending on the details

---

[10] The legal form of A's control may have important consequences for legal responsibilities, taxes, and the like and, at a second level of analysis, these considerations would enter into the calculus of transaction cost economics.

of the contract between them.) Even a very simple exchange like this one can become quite complex very quickly.

Inefficiencies in unified governance

This example suggests a question that brings to bear the final piece of theory that will be recounted from Williamson: What are the relative *inefficiencies* of unified governance? When telling the tale of figure 20.2, it was asserted that we move to unified governance when the level of specialized assets is high, presumably because the costs of consummating, monitoring, and enforcing the transaction at arms length, given the amount at risk and the relative infrequency of the transaction, become prohibitive. But if one avoids these costs through unified governance, then why not use unified governance exclusively?[11]

As our example of A and B points out, the move to unified governance doesn't avoid transactions all together; instead it changes the nature of the exchange that must be made. In place of an exchange of plans for money or finished pieces for money, we have an exchange of labor services for money. This new exchange comes with an array of transaction costs itself, and it is the *relative* costliness of the various sorts of transaction that determines which form of governance we use. Note well: We see some cases of our prototypical transaction in which B buys plans from A in the market (A is an independent industrial designer), cases where A buys pieces from B (B is a subcontractor), cases where A holds all the equipment and buys labor services from B (the typical machine shop), and cases where B holds all the equipment and buys labor input from A (automobile manufacturers who employ designers). Of course, real-world examples will usually involve more than two individuals and so will be vastly more complex than the simple example we have given, but in each case you should expect to see how the relative costs and benefits of various patterns of exchange determine which pattern of exchange is used.

So what are the costs attendant to unified governance? Williamson points to the differences in what he calls *high-powered* market incentives and *low-powered* internal incentives. That is, with a full market interface between A and B, B faces strong incentives to produce efficiently, to care for his lathe, and so on. There are many things that dull these

---

[11] A related question can be asked at a more macroeconomic level. As there are such things as market power and externalities in the world, we know that a market economy will not necessarily reach a Pareto-efficient outcome. Doesn't this make a centrally planned and run economy superior? What can the market do that a centralized economy cannot? What are the inefficiencies in centralization? These questions, which have answers somewhat similar to the answers we will give for unified governance of a single transaction, are known as the Lange-Lerner controversy in the literature.

incentives when A owns all the assets and employs B. Williamson's list (1985, chap. 6), which is longer and more detailed than we can give here, but a selection follows. If A buys out B's assets and employs B, then A is unable to match such strong incentives in the employment contract she gives to B. She has a difficult time monitoring the effort B expends, and she has an especially difficult time seeing *how* B expends whatever effort he does expend. For example, A may give B a piece-rate contract, which gives B strong incentives to produce parts but very weak incentives to maintain what are now A's lathe and jigs. If B has private information about how hard a particular task is, A may have a difficult time wresting that information from him. It might be imagined that any incentives the market can provide, A can provide for B as well. But there are measurement problems in this case. If B is responsible for maintaining the lathe in good condition and if his contract gives him a financial incentive to do so, the cost of inspecting the lathe must be borne by the two parties. If B's contract incorporates some measure of the profitability of the entire enterprise, B must be concerned that A will manipulate the accounting system to B's detriment. On a more humane note, if A and B form a long-term employer-employee relationship, then A may "forgive" B if certain bad outcomes occur. But this removes the hard edge of market incentives that B would face in an arms-length, market-mediated transaction.[12]

20.2. Mathematical models of transaction cost economics

The preceding discussion, while (I hope) clear, does not conform to the style adopted throughout the book. By now you are probably used to seeing ideas exposited with models that begin with consumers who have specific utility functions, living in a particular (highly stylized) environment, who go on to maximize their way to whatever they can get. How much of Williamson's story can be cast in such terms?[13]

---

[12] There is another side to this last point. Sociologists are usually aghast at the way economists predicate most behavior on base self-interest, and they are quick to point out that in a "good" organization workers internalize the interests of the organization and act at least in part for it. Presumably, the social forces that lead boss A to forgive and protect subordinate B also lead subordinate B to act against his own narrow interests and in favor of the interests of A. Forgiveness may be a cost of unified governance (or any form of long-term relationship), but it may accompany substantial benefits.

[13] Why should we attempt to recast Williamson's story in this fashion? The maintained hypothesis throughout this book is that formal mathematical modeling promotes comprehension and clarity of thought, and it is especially valuable for checking the internal consistency of the stories told.

Williamson locates transaction cost economics in the domain of individuals who are boundedly rational and opportunistic. After part IV (and even part III), you should have no problem with opportunism. But all the consumers we have analyzed in this book have been "rational," at least in the sense that their choice behavior has conformed to the basic models of chapters 2, 3, and (when there were dynamics) 4. How are we to model boundedly rational behavior? What does boundedly rational behavior mean?

As noted in part I, there are no generally accepted answers to these questions. There has been some movement towards developing theories of boundedly rational behavior, but nothing yet has taken center stage as the standard model. In retelling pieces of the preceding story in terms familiar from earlier parts of this book, we must rely on models of individual behavior that are rational in all the standard ways. In effect, we are pushing the theory of transaction cost economics into the domain of individuals who are unboundedly rational and opportunistic. It is important to think through how this colors the analysis and, as importantly or more so, what is missed by this.

With all this as prologue, we mention (in varying degrees of detail) a few of the pieces of the grand scheme of the previous section that have been scrutinized in more mathematical terms.

Incomplete contracts and ownership

Grossman and Hart (1986) analyze a variation on the basic story of the designer A and craftsman B. Their focus is on who should own which assets. The story runs roughly as follows: A and B must, in a two-period model, decide how much to invest in their particular assets, A in a computer system that will aid the design and B in a lathe and jigs. We let a_A and a_B be these levels of investment. The investment levels are chosen, and then further decisions q_A and q_B must be taken. The benefit to party X ($X =$ A or B) depends on the initial decision a_X that X takes as well as the two subsequent decisions and is given by a function $V_X(a_X, q_A, q_B)$.[d]

The crucial assumptions made in the analysis are:

(1) At the outset, before the decisions a_X are taken, the two parties can negotiate over ownership of the two types of assets. This is done in a competitive market. Ownership of an asset gives its owner the right to choose q_X. So, for example, if B sells his assets to A, then in the second period A chooses both q_A and q_B.

---

[d] Grossman and Hart use a particular functional form for V_X. As we won't attempt to reproduce their analysis but only to relate the form of their story, we won't go into these details.

(2) Irrespective of who owns the assets, party X must choose a_X. Think of a_X as a maintenance decision, made in an earlier period while the assets were being used.

(3) Payments made for the exchange of ownership rights to the assets cannot be made contingent on the a_X, q_X, or the values of V_X.

(4) Binding contracts concerning the levels of the a_X cannot be written.

(5) Binding contracts concerning the levels of the q_X can be written, but only after the a_X decisions have been made. It is assumed that efficient levels of the q_X will be chosen. If each party retains ownership of his own asset, then the gains from choosing efficient levels of the q_X (over the levels of benefits obtained if the q_X are chosen noncooperatively) are split 50-50 between the two, coming from some (unmodeled) bargaining process.

Assumptions 3, 4, and 5 drive the model and deserve comment. Assumption 4 corresponds to a standard story of moral hazard with no observability at all. This also explains why it is that the prices charged for ownership rights cannot depend on the a_X. It is further assumed that the V_X are not observable or, if observable, cannot be made the basis of an enforceable contract. (The latter possibility can arise if the levels of the V_X cannot be verified to a standard required by a court of law or some other third party who would be called in to adjudicate any disputes.) As for the q_X, the notion is that q_X comprises a lot of detail about what to do with asset X, depending on contingencies that ex ante cannot be foreseen, or that are too expensive to foresee, or that are foreseen but too expensive or (because of problems of verifiability) impossible to provide for contractually. Ex post binding contracts concerning the q_X can be made, however, since once the contingencies have been realized, a simple, single decision is taken.

When Grossman and Hart turned to analysis, they assumed that the two individuals are well aware ex ante what levels of q_X will be chosen (depending on the pattern of ownership rights). That is, there is no uncertainty about the level of q_X in the model. Why is it infeasible to contract over the levels of the q_X in the first period, but feasible to specify these levels contractually in the second period? Grossman and Hart suggested a model in which q_X is an overall level of investment or effort, which must be directed in particular ways ex post, and one cannot specify ex ante which of the many ways investment or effort should be allocated because of an inability to foresee later contingencies. But then why not write a contract giving the overall level of investment or effort in the first period?

A consistent story is that while an enforceable contract *can* be written concerning the precise allocation of investment or effort (so much to task 1, so much to 2, and so on), one cannot write an enforceable contract on the overall level (so much effort to be distributed in a manner later to be determined). And, in the first period, a detailed contingent contract is either impossible or too expensive to write.

Because ownership confers the right to choose the q_X, different patterns of ownership yield different initial decisions concerning the a_X. If A sells her assets to B, she can anticipate that q_A and q_B will be chosen with B's interests in mind, and this will affect her choice of a_A. A similar story results if A owns the assets. And if each party owns his own asset, then the a_X decisions are made with a view towards strengthening the bargaining position of each, in the bargaining envisioned in (5). Grossman and Hart show how different parameterizations of the model lead to different patterns of ownership being optimal; in effect, they give cases where it is optimal for A to buy out B, for B to buy out A, and for the two to maintain an arms-length, market-based relationship.

This model gives a mathematical look into several of the aspects of transaction cost economics crucial to theories of the firm. The opportunism of individuals is represented by the choices of the a_X. Asset specificity is captured by the fact that each side is locked into a relationship with the other once we move past the stage of choosing the a_X. Governance ex post is determined by ownership; the contract is "filled out" in the second period according to who owns the assets. If we make a rough identity between ownership of assets, unified governance, and firms, then we have a theory that, depending on the particular circumstances, specifies what is the transactionally efficient scope of the firm or firms.

The importance of bargaining costs

Milgrom and Roberts (forthcoming) offer an important commentary on the basic scheme of transaction cost economics. Specifically, they reinforce the importance of the human factors to the story, and they suggest that emphasis in the context of market-mediated transactions should be placed on *short-run bargaining costs*. Their argument runs as follows: Imagine two parties to a transaction are fully rational (in particular, they can foresee all future contingencies and how those contingencies will be met, even if a contract doesn't specify these things) and are able to execute binding short-term contracts. Assume as well that the parties are risk neutral, so that any redistribution of wealth between them has only distributional consequences; there are no income effects.

In general, if the two parties undertake a long-term transaction and do not write (for whatever reason) a long-term contract covering all contingencies, they will have to engage in bargaining and renegotiation as events unfold. In general, bargaining can be costly; time is taken, and (although we didn't see this in the simple settings of chapter 15) in cases where parties have private information, inefficient outcomes can result. It is clear, therefore, that bargaining costs are an important part of transaction costs.

Milgrom and Roberts argue that in this context bargaining costs are the *only* part of transaction costs. Even if there is high asset specificity, much uncertainty, and any specification of frequency, if short-term bargaining costs are zero, all transaction costs will be zero; the parties will reach efficient arrangements as time unfolds. What does it mean for short-term bargaining costs to be zero? That binding contracts governing short-run actions can be made at zero cost, and these contracts always result in short-run efficient actions (holding fixed subsequent actions).

Consult Milgrom and Roberts (forthcoming) for the formal argument, but a sketch is easy to give. In the final stage of any transaction, the assumptions guarantee that an efficient arrangement will be reached. Move back to the penultimate stage. Because the two parties are farsighted, they know what arrangements will arise in the final period, and they know those arrangements will be efficient. Because they are risk neutral, any redistribution of wealth that will take place in the final round of arrangements can be "undone" at the current stage. [e] Applying the no-bargaining-costs assumption again, they achieve an efficient arrangement concerning the penultimate round of actions, and so on. They can work back to the start of the transaction, and their short-term agreements will guarantee efficient actions all the way along.

All of Milgrom and Roberts' assumptions play a role in this argument, but two deserve special highlighting. The assumption that the two parties are rational (in particular, farsighted) is crucial, and it reinforces the importance Williamson attaches to human factors. The assumption that enforceable short-term contracts can be reached is crucial. Compare the conclusion reached by Milgrom and Roberts with that reached by Grossman and Hart. Grossman and Hart had rational, farsighted individuals. But in their story, these individuals cannot sign a binding agreement concerning the initial round of decisions a_X. As both Grossman and Hart and Milgrom and Roberts noted, if the a_X could be made the subject of a

---

[e] If they discount, they would seemingly have to discount at the same rates or have available simple loan contracts between themselves so that they could take advantage of any efficient trading of wealth across time.

binding contract, then (in their setting) the ownership of the assets would be irrelevant.

Having put bargaining costs into the spotlight, Milgrom and Roberts went on to study the costs of bargaining and their sources. In so doing, they made use of the formal literature on bargaining — what we discussed in chapter 15, what we omitted (i.e., bargaining with incomplete information) — and also work on "bargaining via particular mechanisms," such as auctions. And they discussed how both asset specificity and measurement problems play a role in bargaining costs.[14]

Chapter 14 and figure 20.2

A third way we can set the discussion of the first section to mathematics is to interpret figure 20.2 using the theory of repeated play from chapter 14. We saw in chapter 14 that cooperative behavior can arise in situations of repeated interaction where each side to the transaction cooperates because each has a stake in the future sufficient to outweigh any gains that could be obtained by acting opportunistically in the short run. The theory of chapter 14 suggests the following:

(a) Each side must have a stake in a maintained relationship. Or, more accurately, if one party is at risk from opportunistic behavior by the second, then the second must be at risk in the future if he engages in opportunism. Moreover, if the short-run gains for the second party from opportunism are quite large, the current value of the long-run gains from a continued relationship must be large as well. Either the relationship must be repeated fairly frequently (so that, effectively, the discount factor is close to one), or the value of the continuing relationship must be high.

(b) It is enough that the party who puts the other party at risk repeats the transaction fairly frequently, has a substantial amount at stake in the execution of any single transaction, and has her performance observed by potential future trading partners. Then this party's behavior can be tied to her reputation for behaving in a particular way.

(c) The noisier the environment, the less well such constructions work.

(d) There are many possible "equilibria" in repeated interactions, so that definite predictions can become difficult.

Figure 20.2, at least in one respect, appears to fit well with this theory. As the frequency of the transaction is lowered, the nature of relational

---

[14] They went on to study the sources of "bargaining costs" in unified governance, to which we will come in a bit.

contracting passes from bilateral and hierarchical to trilateral. This can be explained in some cases very naturally and directly from (a) and (b) above.[15] Think, for example, in terms of the sale of securities. An individual seller of securities may sell with sufficient frequency to particular buyers (or a particular community of buyers) so that no intermediation is necessary. An example might be AT&T directly marketing its securities to its shareholders (although even here there is some intermediation). But in most cases a third party, an *exchange*, is brought into the picture; the exchange provides certain guarantees that facilitate trade. For example, in futures contracts, the exchange acts as guarantor that the contract will be made good. The exchange polices the trades that take place, monitoring the actions of professional traders, and the exchange (often) offers arbitration services when disputes arise. For this the exchange is paid some sum of money. It is the stream of those payments (or, rather, the amount by which this stream exceeds the direct costs of the service) that gives the exchange an incentive to police trades made in its name, at least as long as investors monitor the reputation of the exchange and are willing to take their business elsewhere if the exchange doesn't fulfill its part of the bargain.[16]

Another example of the applicability of (a) and (b) to the study of governance of transactions is given by Simon (1951). He considers the basic labor exchange relationship and, in particular, the fact that the employer is typically a hierarchical superior in this relationship. That is, the employer specifies what tasks the worker will undertake as those tasks arise, subject only to some general "rules" and to the worker's right to quit. Discretion is reserved largely to the employer, and so the reputational glue that holds this together (that allows the worker to trust the employer) is the general reputation of the employer.

---

[15] Other explanations can be given as well. For example, the lower the frequency, the more the cost of a specialized institution must be amortized over a single transaction, and hence the greater the incentive to move to some form of "general purpose" trilateral relationship. See Milgrom and Roberts (forthcoming) for a detailed development of this argument.

[16] Reputation is not the full story here. (1) Informational efficiency plays a role. Because many trades take place on the exchange, the exchange is able to centralize information about the traders — who is honest and who is not. Even if individual investors can obtain information about one another and deal with each other with sufficient frequency that bilateral relationships are possible, the relatively lower costs of the information for the exchange may still push the transaction from bilateral to trilateral form. (2) Securities trading is really a good deal more complex than a simple exchange between two individuals. Typically, each individual has an agent, a broker, who executes the trade for the individual. As individuals deal relatively infrequently with their brokers, and brokers deal with each other constantly, the possibility of collusion between the brokers against the interests of the individuals is quite high. Hence the exchange has a role to play in monitoring the actions of brokers.

Point (c) fits very neatly into the commentary that accompanies figure 20.2 as well. Quoting from Williamson, successful bilateral relationships are found in cases where the conditions to which parties adapt are "exogenous, germane, and easily verifiable ... and ... consequences must be confidently related thereto" (Williamson, 1985, 77). In other words, it must be clear to each party that circumstances call for an adaptation, and the nature of the required adaptation must also be clear.

Consider in this context the trilateral relationship between a publicly held firm, investors, and external auditors. The firm, to raise capital provides investors with information about its financial health. Since a good deal of moral hazard is associated with the provision of this information, a third party, an auditor, is brought in to bless the reported information (or not). These auditing firms are well paid for their services, as long as they maintain a reputation for giving honest audits. The question is, On what is this reputation pegged? Because each audit engagement is unique and because the number of decisions an audit firm must make is very large, if each audit was tailored to very specific circumstances, it would be quite difficult to decide ex post whether an audit firm was diligent and honest in what it did. Accordingly, when the quality of an audit is questioned, what is important is whether the auditing firm followed well-established guidelines (in the United States called Generally Accepted Auditing Standards) that are set by the industry, even if following those guidelines does not provide the most informative audit possible in the particular circumstance. Moreover, auditors provide very coarse signals of the information they receive. In most cases, they say only that the information provided by the firm is "basically okay" by giving an "unqualified" positive opinion. By having fixed procedures and very coarse reports, it is possible to check ex post whether the auditing firm did what it is supposed to, and so it is possible to maintain a useful reputation as an independent third party.

> It is worth noting that chapter 14 doesn't apply precisely to this story. Chapter 14 concerns simple repeated game situations where one party (or both) plays the same game over and over. In most applications, a transaction is never repeated precisely. But you shouldn't have a hard time seeing how the theory in chapter 14 adapts to cases in which the sequence of "repeated transactions" is really a sequence of similar but not precisely identical transactions. As long as one can find some rule that can be applied (in observable fashion) to the sequence of similar transactions, reputation and folk-theorem style constructions can be made to work.

While the theory of chapter 14 helps us to understand parts of figure 20.2, it does not explain all of that figure. In particular, for very high levels of asset specificity, figure 20.2 calls for unified governance. How do we

explain this? (1) Perhaps it has something to do with the requirement that the value of a continued relationship must rise with increases in the short-run inducements to act opportunistically. With higher levels of asset specificity, short-run inducements can certainly be higher. (2) As assets become more specific, losses from the loss of cooperation increase. Insofar as there are noisy observables, so that relational contracting will require some periods of "noncooperation," we might expect to find a move towards unified governance as the cost of those periods increases.^{f} (3) There is also intuitive appeal to the notion that "strategic risk aversion" plays a role here. Parties to a relational contract may be afraid of finding an inappropriate trading partner, of a relation gone permanently sour, or of long periods of time spent sorting out how the benefits of cooperation will be divided; higher levels of asset specificity mean an increase in the cost of such things, and thus an increase of the overall costs of a relational transaction. Some of these things can be found in chapter 14, but others play no role there; it remains to refine the theory of repeated interaction to encompass the missing effects.

Firms and reputation

Firms have no role to play in the previous subsection, but it is not difficult to provide one for them. In all the constructions of chapter 14, at least one party must be long lived. There are long-lived individuals for whom these constructions can work, but the reach of these constructions is lengthened considerably if we allow firms to carry reputations. A stylized model of the basic idea runs as follows. We imagine a sequence of pairs of individuals who play the game in figure 20.3. That is, at date t, for $t = 1, 2, \ldots$, an individual A_t plays this game against B_t. Note that the unique Nash equilibrium has B choosing x', which gives both sides a payoff of 0.

$$\begin{array}{ccc} B & x & A \quad y \\ \circ & \longrightarrow & \circledast \longrightarrow (1,1) \\ \downarrow x' & & \downarrow y' \\ (0,0) & & (2,-1) \end{array}$$
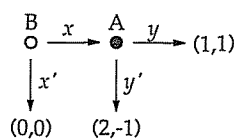
Figure 20.3. An extensive form game.
In this game, the payoff to player A is listed first and that of B is listed second.

---

^{f} Note that these factors would bind relatively less strongly on the third party to a trilateral arrangement, and so we do not completely lose this form of governance for highly specific assets and low frequency.

If we had one player A playing a sequence of Bs, and if this A discounted her payoffs with a discount factor close to one, in the spirit of chapter 14 we could construct an equilibrium in which B_t always chooses action x and A chooses action y, giving both sides a payoff of 1: A chooses y because if ever she chooses y', all subsequent Bs choose x'. In effect, A maintains a reputation for choosing y.

But if each A_t plays only once, this construction will not work, and the mutually beneficial transaction x followed by y cannot be supported (owing to A_t's opportunism and inability to guarantee that she won't select y').

We can get back the reputation construction if we imagine that each A_t lives and consumes for two periods. In period t, A_t plays the game shown in figure 20.3 against B_t. In period $t + 1$, A_t retires and lives off her savings. Assume that A_t's utility is given by $c_t + .95c_{t+1}$, where c_k is A_t's level of consumption of the numeraire good in period k $(k = t, t+1)$.

Suppose then that a "firm" is formed by A_1, called *Honest A, Inc.* This firm carries a reputation with the B_ts as follows: B_t is willing to trust A_t (that is, play x) if A_t owns Honest A, and no previous A_s who owned Honest A ever chose y' when trusted by B_s. That is, if some A_s ever chooses y' instead of y given the chance, and if this A_s owns Honest A, then the reputation of Honest A is irretrievably besmirched.

If the B_ts act in this fashion, each A_t has a positive incentive to purchase Honest A from A_{t-1} if the reputation of Honest A is still good. Suppose the purchase price of Honest A is 10 as long as its reputation is good. Will A_t purchase this firm from A_{t-1}? If A_t believes she can sell the firm to A_{t+1} for \$10 as long as she doesn't hurt its reputation, the answer is yes. Her choices are:

(1) Don't purchase the firm from A_{t-1}. Then B_t will choose x', and A_t will obtain 0 in each of periods t and $t+1$, for a discounted payoff of 0.[17]

(2) Purchase the firm, and then, when B_t chooses x, choose y'. This gives $2 - 10$ in period t (the payoff from the game less the purchase price) and 0 in period $t+1$ (her ownership rights are worthless, since she sullied the reputation of the firm), for a discounted payoff of -8.

(3) Purchase the firm, and then, when B_t chooses x, choose y. This gives $1 - 10$ in period t and then 10 in period $t+1$, the latter being the proceeds from selling the firm to A_{t+1}. This gives a discounted payoff of .5.

---

[17] We do not include in these computations the utility A_t derives from any endowment she might have. Since we have assumed a linear utility function, this is okay. We should assume that A_t has endowment of at least 10 in period t, so she can afford to buy Honest A from A_{t-1}.

Hence (3) is best for A_t, given the posited behavior by A_{t+1} (a willingess to buy) and B_t (a willingness to choose x). This in turn justifies the presumed strategies of A_{t+1} and B_t; we have a Nash equilibrium.

There is nothing complicated here. This is just like the reputation constructions of chapter 14, except that Honest A carries the reputation, and each owner of this "firm" is willing to preserve that reputation in order to recoup the cost of buying the reputation. If Grossman and Hart's story revolved around the notion of a firm as the owner of capital, here the firm is something less tangible; it comprises only its reputational capital. The key is that the individual who makes decisions in the name of the firm has a stake in the consequences for the firm's reputation of those decisions, obtained here by tying the decision maker's second period consumption to the capitalized value of the firm's reputation.[9]

Even by the standards of stylized models in this book, this model is extraordinarily stylized. It should certainly come as no surprise that firms and other organizations carry reputations, or even that for some firms their reputation is as valuable an asset as they have. (Public accounting firms and investment banks come to mind immediately.) These reputations enable all manner of transactions to take place that without the force of reputation would either require extraordinarily complex contracts or would become impossible. Return to the example of the "contract" between a graduate student and the department in which she studies. The contract, such as it is, specifies almost nothing except governance procedures. New students are given a handbook of requirements (subject to change more or less without notice) and instructions how to appeal to the department chairperson and/or the dean. The specific details of a student's program, advisor, financial aid, and so on, are all determined by the parties as time passes and contingencies arise, with the bulk of discretion given to the faculty of the department and sundry deans and administrators. Students can appeal capricious behavior on the part of

---

[9] We don't provide any problems at the end of this chapter, but if we did, one would be: In fact, this setting permits a few things not altogether possible in chapter 14. Suppose, for example, there is noise in the observable action of A. That is, suppose that A may intend to choose y, but there is some chance that by a slip of the hand, y' is observed instead. Can you construct an equilibrium of the following form? If some A_t is observed to choose y', then the value of the firm owned by A_t falls to zero. But A_{t+1} is then able to form a new Honest A, which starts with a fresh reputation. The value of the reputation of this firm doesn't start at 10 immediately, but it grows to 10 as long as its owners are never observed choosing y'. Can you construct an equilibrium where each A_t intends to choose y and each B_t chooses x, if the probability of "y' by mistake" is small? For which patterns in the growth of the value of Honest A is this sort of behavior sustained in an equilibrium? Must the value of Honest A's reputation fall to zero if A_t is observed choosing y'? And what happens if A_{t+1} has the ability to slip some money surreptitiously to A_t?

professors to the department chairperson and capricious behavior on the part of the department chairperson to the dean, but why would a student ever trust the dean? The student trusts in a combination of goodwill and concern for the department's and university's reputation with prospective future students. The stylized model shows how this sort of reputation works in equilibrium, at least as long as retired faculty and administrators continue to take pleasure in the distinction of their old department or university.

The question naturally arises whether reputation really adheres to the individual manager or to the firm. Common sense suggests that the answer is some of both. The stylized model shows how, in an extreme case where the manager (party A_t) is so short lived that reputation cannot adhere to her, reputation adhering to the firm can work instead. But this model, because of its simple structure, misses important pieces of this story. In particular, it may be informationally efficient for reputation to adhere to the organization and not to the individual. It is easier to keep in mind that Honda produces excellent cars than to keep in mind the names of all the engineers and mechanics who together make the cars with the Honda logo. Although more difficult to tell mathematically (because it involves a formal specification of information processing costs or some other model where consumers are not hyperrational), firms and other organizations provide an efficient means for enlisting the power of reputation in promoting transactions.

> Of course, there is a potentially substantial free-rider problem here. Engineer Yamazaki may have to work quite hard to provide his portion of the excellent design of a Honda. He bears all the costs of this effort, yet would (probably) see only a tiny dimunition in the reputation of Honda if he slacked off. So the force of Honda's reputation in keeping him hard at work is probably quite low. But this free-rider problem is met by peer-group and supervisory pressure. (If you are unhappy that there are no problems at the end of this chapter, you could try to flesh this out into a full-fledged model.)

Ex ante versus ex post incentives

Turning to Williamson's discussion of the inefficiencies of unified governance, we note two sorts of contributions.

A major piece of Williamson's discussion concerns the inability of a hierarchical superior to match market incentives because the hierarchical superior retains the ability to "reset" the terms of a contract, an ability that she will enlist in her own interests ex post. The problem is that the knowledge that this is how she will behave ex post prevents otherwise sharp-edged incentives from being put in place ex ante. The problem is

an inability to commit a priori to a particular incentive scheme and/or intervention rule.

We have seen this theme (the inability to commit) in several places in the book, and we will not go into much detail here. But, to refresh your memory, recall our discussion of the ratchet effect. We imagine a principal, an agent, and a task the difficulty of which is unknown to the principal. This task must be performed twice (say), and (we suppose) the principal cannot commit to a two-period incentive scheme. If in the first period the principal learns from the agent's performance precisely how difficult the task is, then she will use this information in designing second period incentives. Hence it will be in the interests of the agent to make the task look harder than it really is in the first period. Note that the root of this problem is the principal's inability to commit to a two-period incentive contract. If the principal could commit at the outset to a compensation package (and to refrain from manipulating the accounts, intervening whenever she wishes to ex post, forgiving unlucky agents, and so on), then she can (theoretically) match high-powered market incentives.

We can take these ideas a step further with an analysis of influence costs. In a unified governance structure, the authority to make decisions is centralized, and individuals affected by these decisions will wish to influence the central authority to whatever extent is possible. If the central authority can be corrupted, attempts will be made to corrupt her. If activities intended to corrupt the central authority take time away from more productive activities, this is a cost of unified governance. These costs increase if the corruption causes the central authority to make decisions that are inefficient. Of course, corruption across a market interface is not unheard of, but the more centralized the authority, the greater will be the level of energy devoted to corrupting activities.

It might seem that this depends on a corruptable centralized authority. But even an "uncorruptable" central authority may be amenable to influence. Insofar as the central authority lacks information relevant for making decisions and looks to subordinates (or to their performances) for this information, then those subordinates will attempt to influence the central authority by manipulating this information. The central authority may well be aware of these attempts at manipulation, but the only way to stop them entirely is to shut off the flow of information, which may be far from optimal. In equilibrium, even with an uncorruptable central authority, wasteful influence activities may go on. See Holmstrom (1982), Holmstrom and Ricart i Costa (1986), Milgrom and Roberts (1988), and Tirole (1986) for formal models.

20.3. Bibliographic notes

Repeated reference has been made to Williamson (1985) as the basic source of transaction cost economics; this is an important summary statement of the subject that should be read by all. While Williamson's contributions have been many and important (not the least of which is the unified treatment of the subject in his book), the subject really originates with Coase (1937). Chandler (1977) and Klein, Crawford, and Alchian (1978) provide other important perspectives.

As for the more mathematical analyses of section 20.2, besides the already referenced Grossman and Hart (1986), Holmstrom (1982), Holmstrom and Ricart i Costa (1986), Milgrom and Roberts (1988, forthcoming), Simon (1951), and Tirole (1986), see Kreps (forthcoming) for the notion of a firm as the carrier of a reputation that passes from one generation of owners to the next and Wilson (1983) on the role and nature of reputation in auditing. For a survey of other work along these lines, see Holmstrom and Tirole (1989).

References

Chandler, A. 1977. *The Visible Hand*. Cambridge, Mass.: Harvard University Press.

Coase, R. 1937. "The Nature of the Firm." *Economica* n.s. 4:386–405..

Grandmont, J.-M., ed. 1988. *Temporary Equilibrium*. Boston: Academic Press.

Grossman, S., and O. Hart. 1986. "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration." *Journal of Political Economy* 94:691–719.

Holmstrom, B. 1982. "Managerial Incentive Problems — A Dynamic Perspective." In *Essays in Economics and Management in Honor of Lars Wahlbeck*. Helsinki: Swedish School of Economics.

Holmstrom, B., and J. Ricart i Costa. 1986. "Managerial Incentives and Capital Management." *Quarterly Journal of Economics* 101:835–60.

Holmstrom, B., and J. Tirole. 1989. "The Theory of the Firm." In *Handbook of Industrial Organization*, R. Schmalensee and R. Willig, eds. Amsterdam: North Holland.

Klein, B., R. Crawford, and A. Alchian. 1978. "Vertical Integration, Appropriable Rents, and the Competitive Contracting Process." *Journal of Law and Economics* 21:297–326.

Kreps, D. Forthcoming. "Corporate Culture and Economic Theory." In *Positive Perspectives on Political Economy*, J. Alt and K. Shepsle, eds. Cambridge: Cambridge University Press.

Marshak, J., and R. Radner. 1972. *The Theory of Teams*. New Haven, Conn.: Yale University Press.

Milgrom, P., and J. Roberts. 1988. "An Economic Approach to Influence Activities in Organizations." *American Journal of Sociology* 94:S154–79.

––––––––. Forthcoming. "Bargaining Costs, Influence Costs, and the Organization of Economic Activity." In *Positive Perspectives on Political Economy*, J. Alt and K. Shepsle, eds. Cambridge: Cambridge University Press.

Simon, H. 1951. "A Formal Model of the Employment Relationship." *Econometrica* 19:293–305.

Tirole, J. 1986. "Hierarchies and Bureaucracies." *Journal of Law, Economics, and Organization* 2:181–214.

Williamson, O. 1985. *The Economic Institutions of Capitalism*. New York: Free Press.

Wilson, R. 1983. "Auditing: Perspectives from Multi-Person Decision Theory." *Accounting Review* 58:305–18.