Design Rules, Volume 2: How Technology Shapes Organizations
Chapter 2 Transactions in a Task Network

Carliss Y. Baldwin
Design Rules, Volume 2: How Technology Shapes Organizations
Chapter 2 Transactions in a Task Network

Carliss Y. Baldwin
Harvard Business School

Working Paper 21-030
Design Rules, Volume 2: How Technology Shapes Organizations

Chapter 2 Transactions in a Task Network

By Carliss Y. Baldwin

Note to Readers: This is a draft of Chapter 2 of Design Rules, Volume 2: How Technology Shapes Organizations. It builds on prior chapters, but I believe it is possible to read this chapter on a stand-alone basis. The chapter may be cited as:


I would be most grateful for your comments on any aspect of this chapter! Thank you in advance, Carliss.

Abstract

From the 1930s through today, many economists have conceived of large technical systems for the production of goods and services as a series of transactions. This point of view has led eminent economists to assert that transactions are the fundamental unit of analysis in the economic system.

This conceptualization has been very powerful, but it is also limiting. To truly understand the relationship between technology and organizations we must look “beneath” transactions at the full set of tasks and transfers that must take place to design and produce useful goods and services in an efficient way. The purpose of this chapter is to describe how transactions fit into a larger network of tasks and transfers and to identify the technological determinants of transaction costs.

Introduction¹

Economic theory as it developed in the 20th Century increasingly suppressed the complex technical recipes used to produce goods and services and focused on the equilibrium determination of prices and quantities as settled by transactions. Thus from the 1930s through today, many economists have conceived of large technical systems for the production of goods and services as a series of input and output transactions.² This point of view has led eminent economists to assert that transactions are the fundamental

¹ This chapter is based on my paper “Where Do Transactions Come From?” (Baldwin, 2008).
² Between the transactions, an abstract production function is used to convert priced inputs into priced outputs. See, for example, Hart and Moore (1990) and Baker, Gibbons and Murphy (2002). Empirical specifications often use the Cobb-Douglas family of functions.
unit of analysis in the economic system.3

This conceptualization has been very powerful, but it is also limiting. One cannot see how technology shapes organizations by looking at transactions alone. To truly understand the relationship between technology and organizational design we must look “beneath” transactions at the full set of tasks and transfers that are needed to design and produce useful artifacts in an efficient way. From a technological perspective, transactions are only a small subset of all the transfers of material, energy and information that must take place in the economic system as a whole.

The purpose of this chapter is to describe where transactions may be cost-effectively located in a larger network of tasks and transfers determined by technology. Transactions and contracts establish the boundaries between legally separate organizations. What happens within such organizations is left open for now, but will be addressed in Chapters 3 and 4.

In the first part of the chapter, consistent with standard practice in organizational economics, the underlying technologies are assumed to be exogenous and unchanging. At the end of the chapter, I show how technologies can be modified to reduce the cost of transactions in specific locations.

2.1 Technology Viewed through the Lens of Transactions

Ronald Coase conceived of production as a chain of invariant “transactions” which could take place within a firm or across firms.4 Implicitly, he viewed technologies as a fixed series of “transactions,” coordinated by “entrepreneurs.” Starting from some point in the system, each entrepreneur decides whether the next transaction is less costly if performed inside his firm or as an “exchange transaction” in the market. “Transaction costs” are the costs of using a market. We can represent this concept of the technical system as follows:

\[
\{TTTT\} T^* \{TTT\} T^* T^* \{TTTT\} T^* \\
\]

Ts refer to transactions, which are fundamental (and exogenous). Ts within brackets are transactions inside firms, T*s outside brackets are exchange transactions subject to costs of using the market. Coase specifically allows for diseconomies of scope (long strings of Ts), perhaps caused by distance from the entrepreneur. The last T* in the sequence indicates the termination of production in an exchange transaction with a consumer.

Following Coase’s seminal contribution, economic theories of organization split into three separate, but overlapping branches: transaction cost economics, property rights theory, and agency theory. Each branch had a different starting point in terms of

3 Commons (1934), as quoted by Williamson (1985) p. 3.

4 Coase (1937).
assumptions, but all were concerned with the optimal design of transactions and contracts between technologically related parties. However, as with Coase, the technologies were characterized in a very simple way.

Oliver Williamson defined transactions as transfers: “A transaction occurs when a good or service is transferred across a technologically separable interface. One stage of activity terminates and another begins.” Drawing on the work of John Commons, he went on to claim that transactions are the basic unit of economic analysis. However, Williamson separated Coase’s indivisible transactions into three parts as shown in Figure 2-1. Each “transaction,” he said, comprises an earlier stage, a later stage, and a “technologically separable interface” in between.

**Figure 2-1  Linear Model of Production**

![Linear Model of Production](image)

According to Williamson, “technological separability between successive production stages is a widespread condition … .” However, even though the stages of production are separable, the assets on either side of the interface may be specialized to one another. *Asset specificity* in turn gives rise to transaction “hazards”—given the opportunity, one side may attempt to hold up the other in order to claim a larger share of the surplus.

In common with Coase, Williamson takes the underlying technical system as exogenous and unchanging. The technological facts are given *ex ante*. An organizational framework consisting of firms and markets is built on top of them. The idea that technological separability and asset specificity might be matters of choice does not enter into the analysis. Hence there is no interplay between transaction cost economics and technology.

In a separate line of research, property rights theorists, including Armen Alchian, Harold Demsetz, Benjamin Klein and Robert Crawford began to look how the allocation of property rights over non-human assets affects “relationship-specific” investments in the presence of significant technological complementarities. Oliver Hart, Sanford

---

6 *ibid.* p. 3.
7 *ibid.*
Grossman, and John Moore, formalized the setup, clarified the theoretical underpinnings, and derived new propositions. But these authors also assumed that the underlying technology was fixed. Given the technology, an organizational framework consisting of contractible rights and residual control rights (ownership) was designed to elicit an optimal amount of relationship-specific investment from each party.\textsuperscript{8}

Agency theorists took a somewhat different perspective on the technical system, by considering the requirements of delegated work. Stephen Ross, Michael Jensen and William Meckling, Bengt Holmstrom, Paul Milgrom and others posited a technology in which the owner of a technology cannot perform all necessary actions herself, but must contract with one or more agents to get the job done. They investigated the impact of different modes of compensation on the agents’ productive behavior. The principal was assumed to control the agents’ contracts: her primary task was to find an “incentive-compatible” contract that maximized her returns net of payments to the agents. Again technology was taken as given, and an organizational framework consisting of incentive contracts was designed to elicit the optimal amount of effort from each agent and the maximum surplus for the principal.\textsuperscript{9}

In all three branches of organizational economics—transaction cost economics, property rights theory, and agency theory—technology is modeled as a mathematical function relating agents’ actions to rewards (a so-called payoff function). The concept of technology implicit in all three branches of organizational economics is much simpler than any technology used to produce real goods and services.

Abstracting away from the complexity of real technologies led to new and powerful insights about the structure of organizations and their role in the economic system. However, one can only go so far with this simplified view of technology and technical systems. To better understand how real technologies affect organizations, we must look within production and design processes to see what happens inside the “black box” of payoff functions. Once we take this step, we can no longer view transactions as the “basic unit of analysis.” Instead, following engineering practice, we must look at tasks and transfers.

2.2 The Task Network

The basic unit in the design of any production process is a task.\textsuperscript{10} A task changes the material world by transforming inputs into outputs in a goal-directed fashion.\textsuperscript{11} Tasks

\textsuperscript{8} Demsetz (1967); Alchian and Demsetz (1972); Klein, Crawford and Alchian (1978); Grossman and Hart (1986); Hart and Moore (1990). Relationship-specific investments are optimal in the sense of being second-best.

\textsuperscript{9} Ross (1973); Jensen and Meckling (1976); Holmstrom and Milgrom (1991; 1994). Again contracts are optimal in the sense of second-best.

\textsuperscript{10} Thompson (1967); J.R. Galbraith (1977); Pahl and Beitz (1995); Puranam, Raveendranam and Knudsen (2012); Puranam (2018).

\textsuperscript{11} Though tasks change the material world, the inputs and outputs of a task may not be material.
must be carried out by agents, but, because of physical and cognitive limitations, no single agent is capable of carrying out all tasks. Thus it is necessary to transfer various things—material, energy and information—from agent to agent in a productive system.\(^\text{12}\)

In many cases, it is useful to distinguish between production tasks and design tasks. The goal of design tasks is to create or change a technical recipe. The goal of production tasks is to carry out a known recipe. In some cases, production and design tasks may be combined as when a cook changes a recipe on the fly. In other cases, they may be performed by specialists working in separate organizations.

Taken as a whole, the tasks, the agents, and the transfers make up a vast network of activity, in which tasks-cum-agents are the nodes and transfers are the links. In a well-functioning task network, agents perform the tasks (including design tasks) needed to produce goods and services. Agents are also matched to tasks and are linked via transfers in such a way that the desired goods are obtained, and no agent has to carry out tasks beyond his or her ability.\(^\text{13}\)

The tasks and transfers in the network are designed, but not by a single person. The designers of the network are generally people with local knowledge, local authority, and local incentives. Because of intrinsic cognitive limits, a single individual, team or company can only work on a subset of the network.\(^\text{14}\)

The network model of production proposed here is more microscopic than the linear model of production adopted by Coase, Williamson and the property rights theorists.\(^\text{15}\) It seeks to model production as an “activity system” described in terms of nodes and links.\(^\text{16}\) Representing production as a network of tasks allows us to model new patterns of dependency and interaction, including parallel flows of material and information, backward flows (rework and feedback), and iterative and uncertain flows (trial-and-error learning). These more complex patterns cannot be modeled as a simple

---

Thinking through the solution to a problem has immaterial thought as both an input and an output. However, the solution will have the capacity to be implemented, and when implemented, will change the material world.


\(^\text{13}\) Increasingly work done by machines is becoming a substitute for work done by humans. Transfers from machine to machine and between machines and humans are thus part of the technical architecture of any modern technical system. However, organizations as I’ve defined them are made up of human beings, thus machines are not part of the systems’ organizational architecture. Similarly, 150 years ago, horses and other animals were an essential part of the technical architecture of many production and transportation systems, but not part of the organizational architecture of the enterprises running the systems.


\(^\text{15}\) Grossman and Hart (1986); Hart and Moore (1990); Baker, Gibbons and Murphy (2002).

\(^\text{16}\) This practice was pioneered by Porter (1996) and has been utilized by Rivkin (2000), Siggelkow (2001), Ethiraj and Levinthal (2004), and Rivkin and Siggelkow (2007). Others, e.g. Powell (1987), Langlois and Robertson (1992), and Sturgeon (2002), have modeled production as a network of firms, but not as a network of tasks.
“sequence of stages,” but they do arise—frequently—in real production and design processes.

At this new, more microscopic level of observation, transactions and contracts are not the basic unit of analysis as Williamson suggested, but are embedded in a more complex network structure. Transactions and contracts are but a small subset of all transfers that need to take place.

### 2.3 What Gets Transferred?

**Material and Energy**

What gets transferred in a task network? First of all, *materials and objects* get transferred from agent to agent through the great chain of production. For example, an automobile starts out as ores, petroleum, silicon, wood, leather, and trace elements. Through a series of tasks and transfers, these raw materials are transformed into components, which are then assembled into a complex artifact.

Likewise, *energy* in various forms—human, heat, mechanical, electrical—gets transferred from generators of energy to those points where the energy is needed.

**Information**

Information also must be transferred among agents within the network. In fact, it is useful to distinguish three types of information: *data, designs, and “tags.”* \(^{17}\)

*Data* includes such things as physical and biological facts, preferences, demands and prices. Whereas materials can be thought of as flowing “down the chain” of production, data often flows “up the chain.” For example, in a modern automobile assembly plant, an order for “a green sedan with a sunroof” may be transmitted from a customer to a salesperson, and thence to a production scheduler. The data in the order can then be used to modify the “downstream” tasks of making a particular automobile. \(^{18}\)

*Designs* are technical recipes, including algorithms, procedures, instructions, and chemical formulas. Designs are the instructions that turn resources into things that people use and value. They are based on technical knowledge and provide the means to solve specific human problems.

---

\(^{17}\) Economists recognize the centrality of information to the functioning of modern economic systems. However, the literature of information economics usually conceives of information as a “signal” arriving from the outside world. Often it is assumed that some agents receive the signal, whereas others do not, hence the information is “asymmetric.” Because their conceptual focus is on signals, economic models tend to concentrate on data and data management and to ignore designs and tags. See, for example, Vickrey (1961); Arrow (1962); Akerlof (1970); Spence (1973); Stiglitz (1975); and Aoki (2001). A notable exception is Simon (1969), Chapter 1, who focuses squarely on designs.

\(^{18}\) This stylized example has been informed by the work on “build-to-order” systems, flexible supply chains, and mass customization by Fujimoto (1999), Gilmore and Pine (1999), and Spear (2002).
To carry out a given technical recipe and especially to create a new design, it may be necessary to transfer complex information and knowledge between and among actors. The design of effective information transfers within an enterprise is the goal of the “information processing view” of firms pioneered by Jay Galbraith. However, to be transferred, information and knowledge must be codified to some degree. Tacit information and knowledge resides “within the heads” of individuals. According to proponents of the “knowledge-based view” of firms, a key challenge for many organizations is to integrate the tacit knowledge of specialists while minimizing the transfer costs.

Tags are the third and final type of information transferred in the network. Tags provide information about which agents can perform particular tasks or transfers. For example, the consumer who bought the green sedan first had to locate an auto dealer. She could do so by looking at the yellow pages, by using an electronic search engine, or by remembering that she had seen a dealer’s sign on her way to work. Yellow pages listings, search engine links, and signs are all tags. Advertisements, telephone numbers, email addresses, domain names, and URLs are also tags.

Decision rights and property rights are a special form of tag. They establish who or what has the right to direct the network at a particular point. For example, there is an upper bound on the number of automobiles that a given assembly line can manufacture in week. If orders exceed the line’s capacity, some agent must decide what will be produced. Two schedulers for one line will create chaos, hence it is reasonable to give one scheduler (probably a computer) the decision rights over a particular line. Which agent has those rights is determined in two steps: first, socially binding property rights determine who gets to select the scheduler; second, a particular scheduler with appropriate training (if human) or programming (if a computer) is designated for a particular line at a particular time.

Money and Credit

Last but not least, in market economies, money and credit must be transferred from point to point in the task and transfer network. Like data, transfers of money and credit generally flow “upstream.” Historically, such transfers involved the movement of material objects, e.g., coins or bullion. But over time, money and credit have become dematerialized, so that today, most such transfers involve only information: an entry in two accounting ledgers stored as data within computers.

In summary, transfers of material, energy, information, and money take place throughout a vast network of productive activity. Transfers are needed because there are limitations on the physical and cognitive capacities of both human beings and machines.

---

19 J.R. Galbraith (1974)
20 M. Polanyi (1966); Nonaka (1994); Grant (1996).
21 For a discussion of the role of tags in complex systems generally, see Holland (1996).
Such transfers must take place in a complex, but logical order, in order to turn components like sheet metal and bolts, plastic shapes, glass, paint, and electronic equipment into complex but useful artifacts like a green sedan with a sunroof. The actual task network is far more complex than indicated by the linear model of production assumed by Coase and Williamson: it involves forward transfers, backward transfers, as well as cycles caused by iteration and trial-and-error.

2.4 Transactions

Within the task network, I define a transaction to be a mutually agreed-upon set of transfers between two or more parties with compensatory payment. It is a reciprocal exchange based on some degree of mutual understanding.

For example, let us assume that Bob needs a screwdriver to perform some task in the network, and Ann has an object that can function as a screwdriver (perhaps a Swiss Army knife). In a transaction-free setting, Ann may simply hand the object to Bob. In this case, (1) Ann and Bob do not have to count the fact that a transfer has occurred; (2) Ann and Bob do not have to define what the object is (Bob may see a screwdriver, while Ann sees a knife); and (3) Bob does not have to pay Ann for the object. The transfer of the object can be effective and productive even if none of these conditions is satisfied. However, under my definition, the transfer cannot be a transaction unless all three conditions are satisfied. (By this definition, unilateral transfers, including gifts, inheritances, thefts, and advertisements, are not transactions. This accords with commonsense usage, as well as the common law definition of a contract.)

Meeting the three conditions adds tasks of defining, counting, and compensating to the network plus transfers of information so that Ann and Bob can arrive at a common understanding of what each is getting and giving up. Thus a transaction is more than a plain transfer, it is a transfer embellished with several added, costly features. (See Inset Box 2-1.) I call these added costs the “mundane transactions costs” of the transaction to distinguish them from the “opportunistic transaction costs” of Williamson and property rights theorists.22

22 Mundane transaction costs include measurement costs as discussed by Barzel (1982; 1997) and Cheung (1983).
Inset Box 2-1—The Three Legs of a Transaction

Defining

*Definition* provides a description of the object(s) being transferred. It places the objects of the transaction into a defined category that is recognized by both parties. Defining adds the costs of describing, communicating and (sometimes) negotiating to the system. In contract theory, if both parties agree on the definition of what is transferred (“this is indeed a satisfactory widget”), the transfer is called “observable.” If third parties can be brought in and also agree (“anyone can see this is a satisfactory widget”), the transfer is “verifiable.” These implicit costs of observing and verifying are mundane transaction costs under my definition. Property rights theorists maintain that such costs are the underlying cause of contractual incompleteness, but treat them as axiomatic.\(^\text{23}\) It is “costly for agents to write detailed long-term contracts that precisely specify current and future actions as a function of every possible eventuality and … as a result, the contracts written are incomplete.”\(^\text{24}\)

Counting

*Counting* associates with the transferred object a quantity — a number, weight, volume, length of time, or flow. Definition is a pre-requisite to counting, because one can only count or measure objects within a class or category. Economics generally takes these pre-defined categories as givens. In other words, goods are defined outside of economics, while prices and quantities (i.e., counts) are determined inside of economics.

When I say that transacted goods must be “counted,” I do not mean to imply that transactions always involve aggregations of goods, like bushels of wheat or tons of steel. Unique goods can be transacted—their count is simply “one.” My definition of “counting” also includes all measuring processes that are used to verify the quality of the transacted object. For example, a complex good, such as a chemical plant, is a unique item, but the contract between the buyer and supplier of the plant will contain pages of detailed conditions, all of which must be met before the transaction is complete.\(^\text{25}\) These conditions define the transacted good. Verifying the conditions involves measurement, hence is a mundane transaction cost of counting.

Compensating

Finally, *compensation* involves the backward transfer of “consideration,” from the recipient to the provider of the transacted object. This in turn requires systems for valuing the object and paying for it. Modern market economies have highly efficient institutions and bodies of knowledge in each of these domains. But whatever the form of compensation, for a transaction to take place, two valuations must occur (one by the buyer and one by the seller), and a payment must be made. The costs of these valuations and payments are mundane transaction costs of compensation.

\(^{23}\) Oliver Hart gives three reasons for contractual incompleteness: (1) the inability to foresee all future contingencies; (2) the costs of negotiating over things that can be foreseen; and (3) the difficulty of writing an unambiguous contract that can be understood and enforced by a third party. Hart (1995) p. 23.

\(^{24}\) Hart and Moore (1990) p. 1122.

Where should we expect to find transactions in the larger task network? Below I will argue that mundane transaction costs are low in some places and high in others. Mindful of this fact, designers of the task network will tend to locate transactions at points where mundane transaction costs are low. They must also decide how much to spend on definition, counting and compensation. Higher levels of precision are more costly, thus designers must make judgments as to how much precision is needed in a particular setting. Finally, designers can change the task network itself to reduce mundane transactions costs in some locations.

2.5 The Determinants of Mundane Transaction Costs

Part of the job of designing a task network is to locate the transactions among the tasks. In Design Rules Volume 1, Kim Clark and I defined a module as a group of elements—in this case, tasks—that are highly interdependent on one another, but only minimally dependent on what happens in other modules. In this section I argue that mundane transaction costs are low at the boundaries of modules and high in their interiors. Thus given a choice between placing a transaction at the boundary or in the interior of a module, one should always choose the boundary. However, to explain the relationship between module boundaries and mundane transaction costs, I must introduce two additional concepts: information hiding and thin crossing points.

The economical transfer of a good from its producer to a user constrains the surrounding transfers of information quite dramatically. The user cannot know everything about how the thing was made: if that information were necessary, the user would have to produce the thing himself, or at least watch every step of production. By the same token, the producer cannot know everything about how the thing will be used, for then she would have to be the user, or watch the user’s every action.

Thus, fundamental to an efficient division of labor is substantial information hiding supporting a “division of cognitive labor.” The user and the producer need to be deeply knowledgeable in their own domains, but each needs only a little knowledge about the other’s.

If labor is divided between two domains and most task-relevant information hidden within each one, then only a few, relatively simple transfers of material, energy and information need to pass between the domains. The overall network will then have a thin crossing point at the juncture of the two subnetworks. Having few dependencies, the two domains will be modules within the larger system. In the task network, modules are separated from one another by thin crossing points and hide information.

Mundane transaction costs are the costs of defining, counting, and paying for

---

28 Parnas (1972a, b); Aoki (2001).
things transferred. At thin crossing points between modules, there are, by definition, fewer and simpler transfers than within modules. Mundane transaction costs will be thus low at thin crossing points. It follows that transactions are best located at thin crossing points, i.e., at the boundaries of modules, not in their interiors.29

2.6 An Example: The Production and Use of an Iron Pot Hook

To make this argument concrete, let us look at the production and use of an iron pot hook in medieval times. (I chose this example because it involves team production,30 but is relatively simple compared to most modern task networks.) Working with iron requires a division of labor: there are many tasks that must be carried out simultaneously in order for the metallurgical processes to work. In medieval times, the efficient production of iron artifacts required from two to six people. The same was true of cooking in a lord’s household.

Assume there are five people on the smith’s team <S1,…, S5>, and five on the cook’s team <C1,…, C5>. If we were to drop into the smith’s establishment and record all transfers of material, energy, and information, we would see that every member of the smith’s team, no matter how lowly, would at some point give material, energy, or information to every other member, and each would receive material, energy or information from every other. The same would be true of the kitchen team. Pot hooks and other iron implements form a bridge between the two establishments.

We can represent the task network of the smithy and the kitchen using a design structure matrix or DSM.31 First, we list the members of each team along the rows and columns of a square matrix. Then, if agent i transfers material, energy, or information to agent j, we place an “x” in the column of i and the row of j. The results are shown in Figure 2-2. The dense transfers of material, energy and information within the smithy and the kitchen show up as blocks of “x’s” in the task structure matrix. But between the two establishments, there is only one point of interaction: the transfer of a completed implement, the pot hook, and a backward transfer of compensation.

29 In effect, thin crossing points in the task network provide the “technologically separable interfaces” that are the starting point of Williamson’s analysis. We shall see, however, that that the task network can be changed, thus thin crossing points are not completely exogenous.

30 Team production lies at the heart of Alchian and Demsetz’s theory of the firm as a “nexus of contracts.” Alchian and Demsetz (1972); Demsetz (1988). It arises in the presence of technological interdependencies among tasks that lead to organizational complementarities among individuals. Various types of complementarity are discussed in Chapter 5 below.

The DSM shows that, in terms of tasks, the smithy and the kitchen are almost, but not quite, independent. The two establishments are *materially connected* by pot hooks and other iron implements, which are made in the smithy and used in a kitchen. And they are *informationally connected* by a set of common definitions of pot hooks and other iron implements. In terminology of Volume 1, the common definitions serve as *design rules*, and, by convention, they appear as a vertical column on the lefthand side of the DSM. The design rules are the “common ground” of the two establishments, thus I have labeled them “CG.” Given this common ground, the two establishments can support one another without a lot of ongoing interaction. Hence this particular pair of subnetworks displays almost perfect information hiding.

It is also relatively easy to turn the completed pot hook transfer into a transaction. Because of their common ground, a smith and a cook both know what a pot hook is, and can agree on its salient features (size, thickness, shape). In this fashion, the object being transferred is easily defined. Pots hooks are discrete material objects, thus easy to count. And cooks know what to do with completed pot hooks: they can easily value them and know what they are willing to pay. Defining, counting, and paying for the pot hook add a few more tasks to the network, but not many. Thus the mundane transaction costs at this location are relatively low.

As predicted, the completed-pot-hook transfer point appears as a *thin crossing point* in the task network: the narrow point between two densely connected subnetworks. Pushing the transaction backward into the smith’s establishment or forward into the

---

32 Baldwin and Clark (2000).

kitchen would require more complex methods of defining, counting and paying for what was being transacted. Thus if the transaction were located at any other transfer point in the two processes (and there are hundreds of transfers points within each), the mundane costs of the transaction and the knowledge overlap between the two establishments would go up. Higher mundane transaction costs and more knowledge overlap result in a less attractive transaction location.

The DSM shown in Figure 2-2 equates tasks with people. In fact this is an oversimplication. In reality, each of the people in the smithy or the kitchen would perform many different tasks and their tasks would change depending on the recipes being carried out at a particular time. Below in Chapter 4, I will tease apart the dependencies between tasks (the technical architecture) and the linkages between people (the organizational architecture).

Although the two forms of architecture may be distinguished conceptually, in many settings it is quite natural not to separate them. For example, in traditional societies, much technological knowledge was transmitted from one generation to the next via supervised participation in the technical process. An apprentice would experience both the technical recipe (what was done) and the organization (who did what) as an inseparable whole. It is only when technical recipes are codified as abstract instructions that technologies become separate from organizations.34

2.7 Transactions at Thick Crossing Points

Can a transaction be placed at a thick crossing point in the task network, where the divisions of knowledge and effort are not as clean as between the smiths and the cooks? Let’s think about what happens as two parts of the task network become more interdependent. To fix ideas, consider two firms, called Upstream and Downstream. Upstream designs and produces disk drives and sells them to Downstream, which designs and produces laptop computers.

The two firms’ manufacturing processes are separate, but the designs of the disk drive and the laptop depend on each other in numerous ways. Initially, much information must pass back and forth between the two firms’ designers if the final product is to work. A simplified DSM for this technical system is shown in Figure 2-3.35

---

34 Interestingly, the earliest cooking recipes consisted of only a list of ingredients. Cooking techniques and the organization of kitchens were not part of the written record, but were transmitted through apprenticeship and experience. Oliver (2018).

35 This example reflects empirical work by Clark and Fujimoto (1991); Argyres (1999); Brusoni and Prencipe (2001); Mayer and Argyres (2004); Staudenmayer, Tripsas and Tucci (2005); Hoetker (2006); and Ethiraj (2007). These studies show how firms today collaborate in the design of complex goods and services and then separately manufacture different components of the system.
In this matrix, transfers of design information are denoted by “x”s, and transfers of material by “m”s. The circled “m” near the bottom of the matrix denotes the transfer of a finished disk drive to the laptop assembly line. The cycle of arrows through “x”s depicts a process of trial-and-error problem-solving: iterations like this are the hallmark of design processes.36

In contrast to the smiths and the cooks, the disk drive and laptop firms’ task networks are highly interdependent in their design processes though nearly independent in their production processes. Transfers of design information are complex. They consist of questions whose answers are unknown and proposed solutions whose values are uncertain. The information exchanges are rich in detail, but at the same time, unstructured and poorly specified, and have uncertain and open-ended consequences.37 Yet leaving out any of them may drastically reduce the value of the final good: the disk drive may fail to work inside the laptop making the entire system worthless.

One way to design a transaction between these two firms is simply to let Downstream buy finished disk drives from Upstream the way a cook would purchase pothooks from a smith or you would purchase a candy bar from a convenience store. This minimal transaction would contain no promises of repeated business, no restrictions on future pricing, and no warranties as to the quality of the disk drive. All would be left to

---

36 Eppinger (1991); Baldwin and Clark (2000). The real task network for processes like these would have many more tasks, dependencies, and potential paths. The laptop and disk drive makers would also have common ground in the form of shared standards. To simplify the figure, I have shrunk the network, depicted only one path, and omitted the common ground.

37 Daft and Lengel (1986); Monteverde (1995); Baldwin and Clark (2000).
the uncoordinated guesses of people on each side.

Absent industry-wide technical standards (discussed below), a minimal transaction will not result in a working laptop. There is simply not enough common ground to satisfy the technological constraints imposed by the physics of electronic circuits and signals. For the disk drive to work inside the laptop, it must be designed with the specific requirements of that laptop in mind.

Design interdependency is a form of asset specificity, which in transaction cost economics is the source of opportunist transaction costs. Asset specificity arises when asset A works with asset B (and vice versa) but no other pairings are possible. Given asset specificity, once Upstream’s costs are sunk, Downstream can unilaterally set a low price on disk drive purchases, causing Upstream to lose its investment. Or if the demand for laptops is unexpectedly high, Upstream might demand a higher price for drives in return for timely shipments.

In the presence of these opportunistic hazards, the property rights theorists have shown that each party has reason to make defensive investments. For example, the drive firm might spend money to make its drives compatible with other systems and the laptop firm might look for second-source suppliers. But such ex ante defensive actions reduce the value of the entire productive system.

Finally, to produce a high-quality laptop, a great deal of design information needs to be “produced” and transferred between the two firms. But by assumption, Upstream only receives compensation for disk drives. Transfers of design information are costly to the drive maker, but unrewarded. Thus Upstream will skimp on these transfers as much as possible or shirk them altogether. Such skimping and shirking are a form of agency cost that reduces the value of the end product, hurting both firms.

Placing a minimal transaction at a thick crossing point in the task network opens the door to opportunist transaction costs, defensive investments, and agency costs. Different branches of organizational economics are concerned with each of these costly behaviors, but technological interdependence—a thick crossing point—underlies all of them.

**Formal and Relational Contracts**

Reducing opportunist behavior at a thick crossing point in the task network requires a contract, either formal or relational or both. A formal contract defines the responsibilities of each party; measures compliance; and establishes multi-dimensional compensation. Thus a formal contract reduces opportunist transaction costs by increasing mundane transaction costs.

---

38 Williamson (1985) pp. 52-56.

Relational contracts differ from formal contracts in that not all contingencies need to be foreseen and addressed at the outset. To control opportunistic behavior, a relational contract creates “a shadow of the future” and provides a means of ex post settling up to make the distribution of gains more fair. Parties to a relational contract “do the right thing” because the value of the continuing relationship exceeds the value of defecting. (The continuation value may in part lie in social relationships, not only financial compensation.)

Relational contracts economize on mundane transaction costs in that not every contingency needs to be defined, measured and valued ex ante. Only contingencies that actually arise need to be addressed, and “in the shadow of the future,” the parties will usually have incentives to settle their disputes in a constructive fashion. Nevertheless, relational contracts do incur mundane transaction costs, only in less obvious ways. Relational contracts don’t just happen: like any form of contract, they must be designed and managed. Two strangers cannot immediately arrive at a relational contract: there are numerous tasks (e.g., meetings) and transfers (e.g., conversations) involved in defining the relationship. In addition, costs of counting, valuation and payment arise in the course of adjudicating ex post settlement.

Showing how this works in practice, Kyle Mayer and Nicholas Argyres describe how, over eleven contracting rounds, a PC company and a software company learned to define, measure and provide informal compensation for more and more of their complex information transfers. Trust between these two companies grew even as their contracts became more lengthy. In this fashion, mundane transaction costs—including the costs of setting up the initial open-ended relationship—reduced opportunistic transaction costs and improved the quality of the transactional relationship.

In summary, transactions can be placed at thick crossing points in the task network, but they will be more costly than those occurring at thin crossing points. At a

40 On the economics of relational contracts, see Kreps, Milgrom, Roberts and Wilson (1982); Grief (1998); Baker, Gibbons and Murphy (1999; 2002). In the legal tradition, Macneil (1985; 1987) defines relational contracts as associations that have significant duration and involve close personal relationships, with “entangling strings of friendship, reputation, interdependence, morality, and altruistic desires” (Macneil, 1987, p. 276). In contrast, in economics and game theory, agents are assumed to be purely calculative about a continuing relationship; that is, they constantly ask the question “Is it worthwhile for me to stay in this relationship or not?” Williamson (1993a) argues that commercial relationships are fundamentally calculative, i.e., that Macneil’s form of relational contract does not really exist between commercial actors. However, see Craswell’s (1993) comment on this point and Williamson’s (1993b) reply.

41 Sako (1992, 2004); Gibbons and Henderson (2012).

42 Mayer and Argyres (2004). This finding stands in contrast to results from laboratory experiments, which show that trust diminishes in the presence of formal contracting — see, for example, Malhotra and Murnighan (2002). The resolution of this discrepancy may lie in the fact that experimental subjects interact for only short periods of time, thus do not build up their relational contract over time. Formal contracts and trust may be substitutes in the short run when the parties are signaling their respective approaches, but complements in the longer run when the parties are learning to manage their ongoing relationship.
thick crossing point, the parties to the transaction must trade off mundane transaction costs (a longer, more detailed agreement) against potential opportunistic transaction costs to arrive at a contract that works in practice. In some cases, the costs of defining, counting and valuing every interaction and contingency can be reduced by building an ongoing relationship that has continuation value for both parties.

2.8 Changing the Task Network to Support Transactions

Up to this point, I have assumed that the task network’s structure is fixed. This is the standard approach to technology in all branches of organization economics. Transaction cost economists, property rights theorists and agency theorists generally assume that technologies, modeled as payoff functions, are exogenous and unchanging. However, technical dependencies between different tasks can be created or eliminated, with the objective of making crossing points thinner or thicker.

In this section I consider the possibility of making a thick crossing point thinner through the process of modularization. We have seen that thinner crossing points have lower total transaction costs, thus firms wishing to transact can often modularize their task networks to support the transaction.

In Design Rules, Volume 1, Kim Clark and I described the process of modularization as follows:

[T]he architects… have as their goal the creation of a set of independent blocks at the core of the design process. They… then set about systematically to sever all dependencies known to exist across the proto-modules. … [I]nterdependencies can be severed by promulgating design rules early in the process.\(^{43}\)

For example, a disk drive and the laptop must pass information back and forth using the same commands, protocols, timing, and compatible physical connections. The disk drive and laptop makers could negotiate over all of these features, but often it is more efficient to adopt an industry standard. Industry standards provide design rules that settle many questions about interoperability in advance, thus eliminating many individual forward and backward dependencies between the two designs. In place of lateral interdependency, the standard specifies a set of rules that both designs must follow.

The impact of a standard on the task network of the disk drive and laptop firms is illustrated in Figure 2-4. Most of the “x”s representing design information transfers in the top-right and bottom-left quadrants of the matrix have disappeared. In their place, an interdependent block representing the choice of a standard appears in the top left corner, and a line of “s”s appears to the left of the Upstream and Downstream task blocks. Now many questions regarding interoperability of the of the disk drive and the laptop can be resolved by consulting the standard, instead of via lateral consultation and negotiation. There is a higher degree of information hiding between the two groups. Furthermore,

\(^{43}\) Baldwin and Clark (2000) p. 70.
many tests of the quality of the individual components can now be reduced to “is the component ‘standards compliant’?”

Figure 2-4  The Impact of Standards on the Disk Drive/Laptop Transaction

<table>
<thead>
<tr>
<th>Negotiate Standard and Tests</th>
<th>Upstream (Disk Drive)</th>
</tr>
</thead>
<tbody>
<tr>
<td>s</td>
<td>x x x x x</td>
</tr>
<tr>
<td>s</td>
<td>x . x x x</td>
</tr>
<tr>
<td>s</td>
<td>x x . x</td>
</tr>
<tr>
<td>s</td>
<td>x x x . x x</td>
</tr>
<tr>
<td>s</td>
<td>x . x . x x</td>
</tr>
<tr>
<td>s</td>
<td>x x x x</td>
</tr>
<tr>
<td>s</td>
<td>x x x x x x</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Design Disk Drive</th>
<th>x</th>
</tr>
</thead>
<tbody>
<tr>
<td>s</td>
<td>x</td>
</tr>
<tr>
<td>s</td>
<td>x . x x x</td>
</tr>
<tr>
<td>s</td>
<td>x x . x</td>
</tr>
<tr>
<td>s</td>
<td>x x x . x x x</td>
</tr>
<tr>
<td>s</td>
<td>x x . x</td>
</tr>
<tr>
<td>s</td>
<td>x x x . x x</td>
</tr>
<tr>
<td>s</td>
<td>x x x x</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Produce Disk Drive</th>
<th>m . x</th>
</tr>
</thead>
<tbody>
<tr>
<td>s</td>
<td>x</td>
</tr>
<tr>
<td>s</td>
<td>x . x x x</td>
</tr>
<tr>
<td>s</td>
<td>x x x x x x</td>
</tr>
<tr>
<td>s</td>
<td>x x x x x x</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Design Laptop</th>
<th>x</th>
</tr>
</thead>
<tbody>
<tr>
<td>s</td>
<td>x</td>
</tr>
<tr>
<td>s</td>
<td>x . x x x</td>
</tr>
<tr>
<td>s</td>
<td>x x . x</td>
</tr>
<tr>
<td>s</td>
<td>x x x . x x x</td>
</tr>
<tr>
<td>s</td>
<td>x x . x</td>
</tr>
<tr>
<td>s</td>
<td>x x x . x x</td>
</tr>
<tr>
<td>s</td>
<td>x x x x</td>
</tr>
<tr>
<td>s</td>
<td>x x x x x</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Produce Laptop</th>
<th>m . x</th>
</tr>
</thead>
<tbody>
<tr>
<td>s</td>
<td>x</td>
</tr>
<tr>
<td>s</td>
<td>x . x x x</td>
</tr>
<tr>
<td>s</td>
<td>x x x x x x</td>
</tr>
<tr>
<td>s</td>
<td>x x x x x x</td>
</tr>
</tbody>
</table>

Downstream (Laptop)

Using standards thus turns a thick crossing point into a thin one. A thin(ner) crossing point, as I have argued, reduces the exposure of both sides to opportunism in the form of holdup, skimping or shirking, and thus reduces the mundane transaction costs needed to control opportunism. The two firms do not have to invest as much in defining, measuring, and paying each other to share design information. They may be able to get by without a long-term relational contract.

However, as I’ve drawn the new DSM, the modularization of the task network is incomplete. This reflects two common features of the real world: (1) standards generally lag the cutting edge of technology; and (2) vendors often do not want standards to be complete. Official standards, such as the ATA interface governing disk drives, are controlled by standards setting organizations which are governed by committees and consensus. Official standards are slow to form, and the standard setting organization may be subject to opportunistic transaction costs in the form of holdup and delay. Thus there will be technical dependencies affecting interoperability, which the official

---

44 Simcoe (2012; 2014)
standards do not settle.

Component vendors will then move to fill this gap. Some—those with market power or a technical lead—will attempt to establish unilateral proprietary standards. But others will make a point of working with customers on just these issues to create customized solutions that can deliver higher levels of system performance. For example, with the so-called FC interface for hard disk drives, signal quality deteriorates at higher speeds. There are ways to compensate for this effect, but they depend on the conditions within the customer’s establishment. Fujitsu, a vendor of hard disk drives, developed a programmable device that could be customized to the needs of each site.

The literature on transaction design is full of examples of partial modularizations. In these instances, crossing points in the task network are made thinner, but still remain quite thick. For example, transacting companies often set up “single points of contact” or “liaisons” through which requests for information and change orders must pass. By requiring information to flow through a single channel, these provisions create thinner crossing points between the organizations.

2.9 Conclusion—How Technology Shapes Organizations

Technology shapes organizations by establishing artifacts (like the pothook) such that relatively few transfers of material, energy and information need to take place between agents making the artifact and those using the artifact. Such artifacts are easy to define, count and value and thus have low mundane transaction costs. Buyers and sellers can then easily agree on the terms of a transaction. It is also relatively easy for third parties to determine facts and resolve disputes. (The fact of easy verification itself will decrease the frequency of disputes.)

Transactions are cheap at these “thin crossing points.” Markets playing host to many transactions are likely to arise in these places in the task network.

Technology also shapes organizations at thicker crossing points. If a transaction is desirable at a thick crossing point, the parties must invest in organizational ties that support ongoing communication and coordination. They must take the time and expend the effort to create a robust formal and/or relational contract. They may also need to undertake unproductive investments as a defense against their counterparties’ opportunism.

Alternatively, as I shall argue in the next chapter, the parties may create a transaction-free zone where the transfers required by the technology can take place

46 Proprietary standards are discussed in detail in Chapters 15-17.
48 Sako (2004); Mayer and Argyres (2004); Staudenmayer, Tripsas and Tucci (2005).
49 Blair, O’Connor and Kirchoefer (2011).
without having to define, count or pay for each one.

However, technology itself is not exogenous or unchanging. Technical recipes may be redesigned by changing lateral dependencies (e.g., real time communication) into hierarchical dependencies (e.g., rules or standards) or vice versa. Crossing points in the task network can be made thinner to facilitate transactions or thicker to provide higher levels of information sharing and real-time coordination.

Within limits set by the laws of nature and human knowledge, the design of a technical recipe and the organization to carry it out is an interactive, two-way process. Different individuals and groups can experiment with different technical and organizational architectures. Nevertheless, to arrive at a viable combination of technology and organization, technologists and managers must be aware of how the dependencies within a given technical architecture give rise to both mundane and opportunistic transaction costs.
References


55.


Press.


