06-051

Where Do Transactions Come From? A Network Design Perspective on the Theory of the Firm

Carliss Y. Baldwin*
Kim B. Clark†

Copyright © 2006 Carliss Y. Baldwin and Kim B. Clark

Working papers are in draft form. This working paper is distributed for purposes of comment and discussion only. It may not be reproduced without permission of the copyright holder. Copies of working papers are available from the author.

Where Do Transactions Come From?
A Network Design Perspective on the Theory of the Firm

Carliss Y. Baldwin*

Kim B. Clark†

May 11, 2006

A very early version of this paper was presented at the Saint-Gobain Centre for Economic Research 5th Conference, Paris, FR, November 7-8, 2002. We thank Richard Langlois, Michael Jacobides, Jason Woodard, Masahiko Aoki, Ben Bensaou, Amar Bhide, Robert Boyer, Takahiro Fujimoto, Robert Gibbons, Nobuo Ikeda, Datta Kulkarni, Luigi Marengo, Sharon Novak, Mari Sako, David Scharfstein, David Sharman, Robert Solow, Robin Stevenson, and Daniel Whitney, as well as seminar participants at the MIT Design Structure Matrix Workshop, the NBER Conference on Organizational Economics, the London Business School Modularity Miniconference, Southern Methodist University, Columbia Law School, and the MIT Industrial Performance Center for thoughtful comments on previous drafts. We alone are responsible for errors, oversights and faulty reasoning.

*Harvard Business School
†Brigham Young University, Idaho

Direct correspondence to:
Carliss Y. Baldwin
cbaldwin@hbs.edu

Copyright © Carliss Y. Baldwin and Kim B. Clark, 2006

Where Do Transactions Come From?
A Network Design Perspective on the Theory of the Firm

Carliss Y. Baldwin and Kim B. Clark

Abstract

Our goal in this paper is to explain the location of transactions (and contracts) in a larger system of production. We first characterize the system as a network of tasks and transfers. While transfers between agents are necessary and ubiquitous, the mundane transaction costs of standardizing, counting, valuing and paying for what is transferred make it impossible for all transfers to be transactions. We go on to argue that the modular structure of the network determines its pattern of mundane transaction costs, and thus establishes where cost-effective transactions can be located.

Key words: transaction — transaction cost — modularity — encapsulation — information flows — division of cognitive labor — network — network design

JEL Classification: D23, L22, L23, M11

Introduction

For the last thirty years economists have used the related concepts of "transaction," "transaction cost," and "contract," to illuminate a wide range of phenomena, including vertical integration, the design of employment contracts, the relation of a corporation to its capital providers, and the economic development of societies and nations.[1] The success of this work is clear, not only from the numerous theoretical insights it has generated, but also from the fact that these concepts are now deeply embedded in the fields of economics, sociology, business and law. But although economists and management scholars have explored the design of transactions and contracts in a wide variety of settings, in most of this literature, it is assumed that a pre-existing division of knowledge and effort makes a transaction possible at a particular point in a larger productive system. The models in this literature then compare, contrast and even "choose" between different forms of transactions at the point in question. But they almost never ask why the opportunity to have a transaction occurs where it does. As a result, the forces driving the location of transactions in a system of production remain largely unexplored.[2] Simply put: where do transactions come from? Why do they arise where they do?

Our goal in this paper is to explain the location of transactions (and contracts) in a system of production. Systems of production are designed artifacts, and where to place "transactions" is

[1] Ronald Coase's two seminal articles, "The Nature of the Firm" (1937) and "The Problem of Social Cost" (1960) established the foundations of the intertwined fields of transaction-cost economics and contract theory. It is impossible to cite adequately all important contributions to this widespread literature. For a selective overview by one of the founders of the field, see Williamson (2000). For a comprehensive overview, see Furubotn and Richter (2005). On vertical integration, see especially, Williamson (1985), Grossman and Hart (1986), and Baker, Gibbons and Murphy (2002), Gibbons (2004), and Arora and Merges (2004); on the design of employment contracts, see Aoki (1988), Holmstrom (1982), Holmstrom and Milgrom (1994) and Wernerfelt (1997); on contracts between a corporation and its capital providers, see Alchian and Demsetz (1972), Jensen and Meckling (1976), and Hart and Moore (1990, 1998).

[2] Recently, a number of scholars have looked at how technology changes the structure of industries. The first to consider "dynamic transaction costs" was Langlois (1992). Recent contributions include Baldwin and Clark (1997, 2000), Langlois (2002, 2003 2006), Jacobides (2005, 2006), Cacciatori and Jacobides (2005), Jacobides and Winter (2005), and Marengo and Dosi (2005).

one of the basic problems their designers face. Sometimes the laws of physics and logic, operating through a particular technology, make the location of transactions obvious. But at other times, the designers can choose whether to have a transaction or not. The decision to have a transaction (or not) in turn affects the design of other elements in the system.

To explain the location of transactions, we will first characterize a system of production as a *network of tasks* that agents perform and *transfers* of material, energy and information between and among agents. We will then explain what a transaction is (and is not), and what having a transaction entails for the network. We will argue, in fact, that whereas *transfers* of material, energy and information in the network are necessary and ubiquitous, *transaction costs* make it impossible for all transfers to be transactions.

The particular transaction costs we are concerned with, however, are not the costs of opportunistic behavior or misaligned incentives, which are classically the focus of transaction cost economics and contract theory. Instead we will focus on the more "mundane" costs[3] of creating a transactional interface: the costs of defining, counting, valuing and paying for what is transferred. These costs, we will argue, are determined by the material, energy, and information flows in the underlying system of production. At some points, transfers are simple, and therefore easy to standardize, count, and value: mundane transaction costs are low in these places. At other points, transfers are complex, hence impossible to standardize, count, and value, and mundane transaction costs are prohibitive. We will go on to argue that the *modular structure* of the system of production determines its pattern of mundane transaction costs. In this fashion, the network design of a system of production necessarily establishes (1) where transactions can go; and (2) what types of transactions are feasible and cost-effective in a given location.

---

[3] Williamson (1985, p. 105) points to the phenomenon of "mundane vertical integration." Transaction costs, he says, explain this type of integration, but "these mundane matters go unremarked." These mundane matters lie at the heart of our analysis.

Before proceeding to the main argument, we should explain that our intentions in this paper are quite limited. We are seeking to connect the design of a system of production, specifically its modularity, to its (mundane) transaction costs. Many of the connections we make will be obvious to most readers. We think this effort is worthwhile, however, because mundane transaction costs presently lie in the background of transaction costs economics and contract theory: They are taken for granted and simply not mentioned most of the time. In this essay, we will risk saying what is obvious in order to convert implicit knowledge about mundane transaction costs into explicit knowledge. By doing this, we hope to show that there are deep and interesting connections between the currently widely separated fields of transaction cost economics and complex systems design.

The Task and Transfer Network

The basic unit of any production process is a *task*. Imagine all the tasks needed to produce all the goods in a modern economy. The tasks are linked by the logic of their underlying technologies. In particular, the outputs of some tasks are inputs to others. Tasks must be carried out by agents, including people and machines.[4] But no single agent, human or machine, is capable of carrying out all tasks, and thus it is necessary to *transfer* various things from agent to agent in the system. Effecting transfers adds a new set of tasks to the system, and thus transfers are costly. However, these additional tasks are essential given that all agents have both physical and cognitive limitations.[5]

---

[4] It may seem odd to some readers to call machines "agents" in a system of production. However, in modern economies, machines perform many tasks and make many decisions.

[5] On the design of tasks and transfers in a system of production and, especially, the level of detailed specification needed to achieve functionality and efficiency, see, Nevins and Whitney et.al. (1989) and Spear (1999, 2002). Also, the contents of tasks and the nature and location of transfers may change over time as agents learn and as new technologies introduce new agents (like computers) into the system. Thus, in

Taken as a whole, the tasks, the agents who carry out tasks, and the transfers make up a vast *network* of productive activity. The tasks are the nodes and the transfers are the links. We will call this network the "task and transfer network" or simply "the network" for short.[6] A functioning task and transfer network defines and performs tasks, including transfer tasks, and matches agents to tasks in such a way that the desired goods are obtained, and no agent has to carry out tasks that are beyond its ability. In modern economies, the totality of the task and transfer network is mind-boggling, but most of the time, we take its (relatively) smooth operation for granted.[7]

What Gets Transferred?

Material and Energy

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

addition to being finely structured, modern systems of production are inevitably dynamic. On the evolution of a manufacturing system, see, especially, Fujimoto (1999).

[6] The idea of a "task network" is fundamental to organizational design (Galbraith, 1977). However, in both theory and practice, a focus on the design of tasks often diverts attention from the equally important design of transfers—the links between tasks. We want to give equal weight to task and transfers, so have adopted a more cumbersome name.

[7] However, both Adam Smith (1776, reprinted 1999, Chapter 1) and Friedrich Hayek (1945) marveled at how well the network operates.

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*." Briefly, data is information about the world that must be received and interpreted by agents in order for the network to function properly. Designs are solutions to problems posed by data. And tags are used to identify and locate resources in the system. We expand on each of these definitions in the paragraphs below.[8]

Data includes such things as physical and biological facts, preferences, demands and prices. In general, single agents or small groups of agents cannot control data although they can (and often must) respond to it. And 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 order and its details are data, which flow "upstream" in the task network. These data must first be transferred, then absorbed and interpreted by a capable upstream agent: the data can then be used to modify the "downstream" tasks of making a particular automobile.[9]

Designs are another type of information that gets transferred in the task network. Designs are "the instructions… that turn resources into things that people use and value."[10] They are what Joel Mokyr, calls "prescriptive knowledge," and include algorithms, procedures, recipes and

[8] 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, Marschak and Radner (1972); Cremer (1980); Aoki (2001), as well as the Nobel Prize winning work of Vickrey (1971); Mirrlees (1976); Akerlof (1970); Spence (1973); and Stiglitz (1975).

[9] This stylized example has been informed by the work on "build-to-order" systems, flexible supply chains, and mass customization by Fujimoto (1999), Pine (1999), and Spear (2002).

chemical formulas.[11] The inputs to a design process are effort and knowledge; the output is a solution to some problem.[12] The solution must conform to the laws of physics and logic and the rules of society. In the task network, designs are inputs to production, hence design information flows "downstream." An assembly line that can make a green sedan with a sunroof must be designed for this purpose. Before the order was given, the design of the line must have been completed and a physical line embodying the design must have been built.

Tags are the third and final type of information transferred in the network.[13] Tags provide information about which agents can perform particular tasks or tranfers. 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. Unlike data, which generally flow upstream, and designs, which flow downstream, tag information may be broadcast every which way. Advertisements are tags, as are job descriptions and professional accreditations. 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.[14] To be effective, they must be both published and in limited supply. For example, there is an upper bound on the number of

---

[10] Baldwin and Clark (2006).

[11] Mokyr (2002).

[12] The idea of a design as the solution to a problem posed by the environment was advanced by the influential design theorist, Christopher Alexander (1964). In his essay on "The Science of Design," Herbert Simon put forward a similar idea, saying that designs are "courses of action aimed at changing existing situations into preferred ones" (Simon, 1999, p. 111).

[13] For a discussion of the role of tags in complex systems generally, see Holland (1996). The notion of tags is closely related to David Parnas' concept of "information hiding" discussed below (Parnas, 1972a,b).

[14] Kraakman and Hansmann (2002) explore the variety of verification rules by which property rights can be established under U.S. law. In our language, they explore problem of tag verification and show how the structure of rights reflects the limits of cost-effective verification procedures.

automobiles that a given assembly line can manufacture in week. Thus the salesman who took the order for the green sedan had to relay it to a production scheduler (which was probably a computer). The scheduler in turn had to convey the order to the line, taking account of other orders and the capacity of the line. Two schedulers for one line will create chaos, hence it is reasonable to give one scheduler 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;[15] second, a particular scheduler with appropriate training (if human) or programming (if a computer) is designated for a particular line at a particular time.

Note the complexity of information flows even in this stylized example. The designation of a particular scheduler is part of the detailed *design* of the task network for automobile assembly. The scheduler in turn must take account of *data* about probable orders and the capacity of the line. The right to make the decision is conveyed by property rights over the line: property rights are *tags*. Finally, the order-takers must know how to submit an order to the right scheduler (whom to contact; what to include in the order). The ordering procedure is a *design*; the order itself is *data*; the information on how to contact the scheduler is a *tag* read by the order-takers. But the scheduler also must have information about which agents are allowed to submit orders, hence the order-takers must have *tags*, too. All of these information transfers swirl around the assembly line for an automobile. The task and transfer network encompasses all of these information transfers, as well as the flows of material and energy in the line itself.

[15] On property as a "bundle" of decision rights including the right to determine use, see, for example, Demsetz (1967); Alchian and Demsetz (1973), Posner (1977), and Grossman and Hart (1986). For a historical review and critique of this concept, see Grey (1980).

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 money or credit transfers involve only information: an entry in two accounting ledgers. However, money and credit are not information in the sense of data, designs or tags. Money and credit transfers obey special rules, and play a special role in the network.[16] Thus, when the customer ordered the green sedan with a sunroof, she had to put a deposit down; when she took delivery, she either had to pay for the car immediately or enter into a binding agreement to pay over time. The car company thus obtained either cash or a financial claim, and the customer's ability to buy other goods was reduced commensurately.

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. 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. Therefore, the transfers, like the tasks between them, must be designed.

The tasks and transfers are not designed by any central planner or authority, however. They are designed by engineers and others with local knowledge, local authority, and local

[16] We thank Robert Solow (private communication, 2002) for this insight. As an example of money's special rules, consider the fact the supply of money and credit must be limited but need not be fixed. Modern macro and monetary economics addresses the optimal management of the supply of money and credit.

incentives. Because of their own physical and cognitive limitations, the engineers and other designers perforce must work on subsets of the network and on the interfaces between those subsets. Transactions, we will see, are a powerful way to create functional interfaces between subsets of tasks and transfers.

What are Transactions?

Transfers within the task network are not necessarily transactions. Indeed a transaction is more than a transfer, it is a transfer that is (1) standardized; (2) counted; and (3) compensated.[17] To see how transactions differ from transfers, consider a generic transfer, in which agent A conveys "X" to agent B, and B receives "X" from A:

"X"

A ———> B

We assume that B must have "X" in order to perform some task in the larger system of production. But (1) A and B do not have to record the fact that a transfer has occurred; (2) A and B do not have to agree on what "X" is; and (3) B does not have to pay A for "X". The transfer of "X" can be effective and productive even if *none* of these conditions is satisfied. However, we contend, the transfer cannot be a *transaction* unless *all three* conditions are satisfied.

[17] Most economists operate at a high level abstraction with respect to the system of production and thus do not distinguish between transfers and transactions. For example, Ronald Coase (1937 reprinted in 1988) implicitly conceived of production as a sequence or chain of invariant "transactions" which could take place within a firm or across firms. Oliver Williamson (1985) defined transactions as transfers: "A transaction occurs when a good or service is transferred across a technologically separable interface," (p. 1.) and went on to say that it is "easy and even natural to regard the transaction as the basic unit of analysis" (p. 88). (Notably, Williamson did not define "technologically separable interface," although he did say they were common.) If all transfers are transactions, however, the idea of a transaction as a reciprocal exchange based on common understanding is lost. Reciprocal exchanges based on common understanding play an enormously important role in the system of production, yet they are not the only way that material, energy and information move around the system. Thus, from an network design perspective, it is crucial to distinguish between transactions and transfers.

For example, suppose "X" is a piece of information. B may obtain "X" in a casual conversation with A: the transfer is unrecorded and uncounted. Also A and B do not have to agree on what B can, should or may do with the information: the transfer is not standardized. And if A does not charge for "X", B does not have to worry about its value or the means of payment ("does A take American Express?"). Hence, the transfer of information "X" can take place without being counted, standardized, valued or paid for.

This is true for material and energy as well: for example, when a host sets up a buffet table, material and energy are provided to the guests without counting the transfers ("how many shrimp did you take?"); standardizing them ("big shrimp or little shrimp?"); or demanding payment ("what will you pay for a napkin to wipe the sauce off your chin?"). Even money can be tranferred without transacting as, for example, when a parent gives lunch money to a child.

A transaction, we propose, is a transfer that *is* standardized, counted, valued and paid for.[18] Each of these steps adds tasks and transfers to the network. Thus a transaction is more than a plain transfer, it is a transfer embellished with several added, costly features. We describe the costs of making a transfer into a transaction in the paragraphs below.

Standardization places the object "X" into a defined category that is recognized by both parties and possibly others.[19] Thus standardization adds tasks of description, communication,

---

[18] 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. The definition is also consistent with the concept of "reciprocal altruism" or "social exchange" in evolutionary psychology (cf. Cosmides and Tooby, 1992, p. 177).

[19] In the language of formal contract theory, if both parties agree on the categorization ("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." Much contract theory concerns transfers that are observable, but not verifiable (Hart 1995, p. 37, note 15). Implicitly, in these models, there is an additional cost of making the transfer verifiable: this is a cost of standardization under our definition. But formal contract theory takes this cost as coming from outside the model, and does not inquire as to its causes. At the same time, formal contract theory assumes that other costs related to standardization and counting, for example the cost of definining and measuring "high" vs. "low" quality output, are zero.

and (sometimes) negotiation to the network.[20] For example, if information "X" is going to be the basis for a transaction, A and B must agree on its category (news, stock quotes, a movie, a story, a design, a recipe, a set of instructions); its form (text, pictures, code); and the mode of transmission (paper, telephone, Internet). And in setting up a buffet, if the host hires a caterer, then the two must agree on the menu, service and cleanup.

Counting gives the transferred object "X" a quantity, Q — a number, weight, volume, length of time, or flow. Hence counting requires a measurement system that is appropriate to "X"s category.[21] Counting also adds measuring tasks and transfers — taking measures, recording measures, communicating measures — to the network. For example, a caterer will generally charge by the guest, but will buy chips by the bag and shrimp by the pound; and will hire servers by the hour. Thus the host and the caterer must count guests; the grocer and the caterer must count bags of chips; the shrimp merchant and the caterer must count pounds of shrimp; and the servers and the caterer must count hours. (Only the guests don't have to count anything.)

Finally, *compensation* involves a backward transfer of "consideration," from the recipient to the provider of "X." This in turn requires systems for valuing the transferred object and paying for it. Modern market economies have highly developed and specialized systems in each of these domains. Valuations, for example, can be based on cost, on functionality, on relative performance, or on pure preference. And payments can be in cash, on credit, or in kind.[22] But

---

[20] On the tasks and competencies involved in specifying a transactable object, see Fine and Whitney (1996). On the emergence of new standardized product descriptions in the mortgage securities industry during the 1970s, see Jacobides (2005).

[21] Standardization is a pre-requisite to counting, because one can only count objects within a class or category. Economics takes the existence of standardized categories to be axiomatic: Goods are defined outside of economics; prices and quantities are determined inside of economics. Thus, as a discipline, economics takes "the existence of its very objects and the nature of their dynamical couplings as given and immutable from the very start, thus committing itself to purely quantitative and non constructive theories" (Marengo, Pasquali and Valente, 2003). Barzel (1997) is a notable exception, however, as he conceives of goods as fluctuating bundles of attributes and property rights.

[22] Modern banking and credit systems have evolved to make the per-unit cost of cash and credit payments much lower than in the past.. As a result, in-kind payments are rarer than they used to be, although they

whatever the amount or 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 or promise to pay has to be made. For example, the host must decide if he is willing to pay the price of the caterer's services. The caterer, in turn, must be sure that her revenue from the job exceeds her costs. And at the end of the party, the caterer must be paid, and she in turn must pay (or have already paid) the grocer, the shrimp merchant and the servers. These valuations and payments add still more tasks to the network.

In summary, objects that are transacted must be standardized and counted to the mutual satisfaction of the parties involved. Also, in a transaction, there must be valuations on both sides and a backward, compensatory transfer—consideration paid by the buyer to the seller. Each of these activities—standardizing, counting, valuing, compensating—adds a new set of tasks and transfers to the overall task and transfer network. Thus it is costly to convert even the simplest transfer into a transaction.

Taken as a whole, standardizing, counting, valuing and paying for transfers give rise to what we call "mundane transaction costs."[23]

Why Have Transactions?

If transactions are more costly than plain transfers, then why have transactions at all? The obvious answer for economists is: "to reduce opportunism and align incentives."[24] That is true

have not disappeared. Trading in a used car is a partially in-kind payment; swapping one company's stock for another is another.

[23] In related work, Langlois (2006) defines mundane transaction costs as the frictional costs of transacting. The cost of standardizing is an *ex ante* cost, hence is not "mundane" by his definition. Otherwise our definitions are consistent.

[24] Opportunism here includes free-riding, shirking, the consumption of perquisites or private benefits of control, holdup, defensive investments (or non-investments), misdirected effort, and excessive risk-taking.

enough. But from a network design perspective, transactions also help to coordinate the overall network. In fact, in any farflung, dynamic system of autonomous agents, standardizing interfaces and counting what flows across them is a classic way to manage complexity and regulate behavior. And local compensatory "payments" are a brilliant device for maintaining resource balance (homeostasis) and providing prompt diagnosis, triage, and repair. Thus we expect to see—and we do see—"transactions," that is, standardized, counted and compensated transfers, arising in decentralized, dynamic systems with no human actors.[25]

If transactions are desirable but expensive, then part of the job of designing the network is to *locate* the transactions among the transfers. Which transfers are the best candidates to become transactions? Here we arrive at the key point in our argument: *the mundane costs of converting a transfer into a transaction vary dramatically across different types of transfers*. It is easier to standardize a simple thing than a complex thing. Discrete objects are easier to count than flows. Transfers with predictable effects are easier to value than those with erratic effects, and routine transfers are easier to value than contingent transfers. Finally the unit cost of a single large payment is less than that of many small payments. Thus the *lowest* mundane transaction costs are associated with transfers of simple, discrete, material objects, which can be aggregated into homogeneous lots. And the *highest* mundane transaction costs are associated with contingent transfers of complex bundles of information and action that might or might not turn out to be useful.

The heterogeneity of mundane transaction costs means that cost-effective transactions will not occur randomly in the task and transfer network. This in turn raises new questions: Where should the transactions go? Specifically, how does the network structure affect mundane

[25] As one example, the TCP layer of Internet protocols works by standardizing messages into "datagrams," counting the datagrams sent, and receiving acknowledgment ("compensation") for those that have been received. For other examples, see Ashby (1960); Gerhart and Kirschner (1997), Chapter 4, "The Exploratory Behavior of Biological Systems"; and Levy (1978) on the design of computer buses.

transaction costs? To answer these questions, we need to look at the network itself in more detail, mapping the transfers of matter, energy and information. In this mapping effort, we will make use of a tool of engineering systems design, the so-called Task Structure Matrix. But first we need to introduce the twin concepts of *information hiding* and *thin crossing points*.

Information Hiding and Thin Crossing Points

An artifact must be produced before it can be used. The movement of an artifact from its producer to a user is, of course, a transfer. It is probably the most basic transfer in any economic system, and the paradigmatic "transaction" in economic theory. Thus it is worth thinking about what makes this type of transfer special. We contend that what is special about the producer-to-user transfer is *information hiding*, which in turn gives rise to *thin crossing points* in the task and transfer network.

The efficient 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 to the user, he would have had to produce the thing himself, or at least watch every step of production. The efficiency of the division of labor would then collapse. 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 the efficient division of labor is a partition and substantial hiding of information. This *information hiding* in turn supports what Williamson and Aoki call the "division of cognitive labor."[26] 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.

---

[26] Williamson (1999) as quoted by Aoki (2001) p. 96. The term "information hiding," is due to Parnas (1972a,b).

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 structure will have a *thin crossing point* at the juncture of the two subnetworks. Furthermore, *because* the transfers are relatively few and not complex, mundane transaction costs will be low at the thin crossing point. Thus, other things equal, thin crossing points are good places to locate transactions.

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

To fix the ideas of information hiding and thin crossing points, let us look at the production and use of an iron pot hook in pre-modern times. We chose this example because it is relatively simple compared to most modern task networks, and because it illustrates team effort in each domain and a team-to-team transfer: from a smithy to a kitchen.[27]

Even in primitive settings, 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. Hence the production of iron artifacts has always required multiple pairs of hands and eyes: an efficient team might range in size from two to six.[28] The same was true of cooking in pre-modern times, although kitchen teams, especially in wealthy households, often had more than six members.

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

[27] In 1750 and before, kitchens contained many wrought iron artifacts which were made by smiths. These implements included: brackets; pot-hooks; handles; spits; trivets; gridirons; toasters; conjurors; girdleplates; hand-irons; tongs; fire-shovels and dripping pans. See *Iron in the Service of Man* (South Yorkshire Industrial History Society) http://www.top-forge.fsnet.co.uk/Books/Service.htm, viewed 7/4/02. Pots and other cast iron implements were made at larger ironworking establishments.

[28] At Top Forge, a wholesale ironworks in Wortley, England, a visitor in 1640 might see "the hammerman, a boy and a man at the finery hearth, and two men at the chafery hearth. The team produced three tons of wrought iron in a good week." *Iron in the Service of Man*, op.cit. viewed 7/4/02.

material, energy, and information, the resulting graph would be bi-directional and complete.[29] 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.[30] Pot hooks and other wrought iron implements form a bridge between the two establishments: They are the products of the smithy and the tools of the kitchen.

We can represent the task and transfer network of the smithy and the kitchen using a mapping device called a "task structure matrix" or TSM.[31] First, we list the members of each team along the row 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 *j* and the row of *i*. The results of this mapping are shown in Figure 1. The dense web of transfers of material, energy and information *within* the smithy and the kitchen show up as blocks of "**x**s" in the matrix.[32] But between the two

[29] On the use of graphs to represent dependencies (transfers) in production processes, see Kusiak (1995).

[30] Our example has been deliberately selected to reflect the definition of team production put forward by Alchian and Demsetz (1972, p.779): "With team production it is difficult, solely by observing total output, to either define or determine *each* individual's contribution to this output of the cooperating units. The output is yielded by a team, by definition, and it is not a sum of the separable outputs of each of its members." Today much has changed, but some things remain the same. Ichniowski, Shaw and Gant (2002) graphed interactions amongst crews on different steelmaking lines: for the highest performing lines, the graphs were bi-directional and complete. Restaurants and fast food kitchens also have a high degree of bi-directional interaction.

[31] Herbert Simon, in his classic article "The Architecture of Complexity" (1962, reprinted in 1999), may have been the first person to represent the interdependencies of a complex system via a square "causality" matrix. Independently of Simon, Donald Steward (1981) developed techniques for mapping the actual design parameter and task interdependencies of network projects: he named these matrices "Design Structure Matrices" or DSMs. These techniques have been applied and extended by Steven Eppinger (1991) and his colleagues (cf. http://web.mit.edu/dsm/): they are known by several names, including Design Structure Matrix, Dependency Structure Matrix and Task Structure Matrix. In related work, Rivkin and Siggelkow (2001) use interdependency maps, expressed as matrices, to investigate the properties of different organizational structures.

[32] By recording only the presence or absence of transfers, the matrix abstracts from the true complexity of the actual system of production. To capture the whole process, the matrix would have to show: (1) what is transferred and how frequently; (2) under what conditions each transfer will occur (because many transfers are contingent); and (3) whether each transfer is essential to the process or optional.

establishments, there is only one point of interaction: the transfer of a completed implement, in this case a pot hook.

Figure 1
Task and Transfer Network for a Smithy, a Kitchen and an Iron Pot Hook

		Smithy					Kitchen				
		S1	S2	S3	S4	S5	K1	K2	K3	K4	K5
	S1	.	x	x	x	x					
Smithy	S2	x	.	x	x	x					
	S3	x	x	.	x	x					
	S4	x	x	x	.	x					
	S5	x	x	x	x	.					
	K1		**Pot Hook**			x	.	x	x	x	x
Kitchen	K2		**Transfer**				x	.	x	x	x
	K3						x	x	.	x	x
	K4						x	x	x	.	x
	K5						x	x	x	x	.

The TSM shows that the systems of production of 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. But as long as the smiths and the cooks agree on what a pot hook (or spit or gridiron) is, the two establishments can support one another without a lot of ongoing interaction. Hence this particular pair of subnetworks displays almost perfect information hiding. The cooks do not have to know how to make pot hooks, and the smiths do not have to know how to make stew.

It is also relatively easy to turn the completed pot hook transfer into a transaction. A smith and a cook can agree on what a pot hook is, and its salient features (size, thickness, shape), hence the transfer can be standardized. 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. Standardizing, counting, valuing and paying for the pot hook create a few more tasks in the network, but not many. Hence the mundane transaction costs at this location are relatively low.

Pushing the transaction backward into partially completed pots hooks or forward into food preparation would require much more complex definitions and systems for counting and valuing what was being transacted (a molten pot hook, a pot hook with jagged edge, a pot hook plus an onion, a pot hook plus a pot of stew). Also if pot hooks were delivered incomplete, cooks would need to learn the smiths' craft and do the smiths' work; if pots hooks were delivered with partially cooked meals, the smiths would have to learn the cooks' craft and do the cooks' work. Thus if the transaction were located at any other transfer point (and there are hundreds of transfers points in the two establishments), the mundane costs of the transaction would go up. At the same time, the two information sets, which can be almost disjoint, would have to overlap. Higher mundane transaction costs and more information overlap make for a less attractive transaction location.[33]

Visually, as we predicted, the completed-pot-hook transfer point is a *thin crossing point* in the task structure matrix: the narrow point between two densely connected subnetworks.[34] Substantively, this means there are many, non-standard and complex transfers of material,

---

[33] The so-called knowledge-based theory of the firm views firms as bundles of problem-solving routines and capabilities (Nelson and Winter, 1982; Kogut and Zander, 1992; Nickerson and Zenger, 2004). In this literature, it is argued that firms exist to economize on the production and exchange of knowledge. From this premise, it follows immediately that the boundaries of firms should be placed so as to minimize overlaps of knowledge or information between firms. However, this literature generally does not address the costs of creating and maintaining a functional interface between two firms. Brusoni and Prencipe (2001) study the degree of knowledge overlap between buyers and sellers of complex goods (e.g., airplane manufacturers and airplane engine manufacturers). They conclude that, in the best of circumstances, knowledge overlap is high, but they do not attempt to measure the relative magnitudes of the overlapping and hidden information sets.

[34] Thin crossing points are akin to "structural holes" in Ronald Burt's (1995) theory of social networks and social capital. However, in Burt's theory, structural holes are bottlenecks in the social network, whereas thin crossing points increase the efficiency of the task and transfer network. Quoting computer network designer H. T. Kung, "Striving to have the thinnest crossing point between two systems is a natural objective of design" (Kung, private communication, 2002).

energy and information that need to take place within each establishment, but only a few, standard and relatively simple transfers that need to take place between the two.

Again we must ask, why have any transaction at all? First, in a very small task network, like a manor or a plantation, transfers from the smiths to the cooks would probably not be transactions. The smiths would make iron implements, based on requests from the kitchen, and they would give the finished artifacts to the cooks.[35] But, as we said earlier, in a larger system, a transaction between the smiths and the cooks can be a useful *lateral* coordinating mechanism. Even though their physical efforts and information sets are almost disjoint, the two establishments still need to be coordinated in terms of how many and what types of iron implements the one will buy and the other will supply. Placing a transaction—a shared definition, a means of counting, and a means of payment—at the completed-pot-hook transfer point allows the decentralized magic of the price system to go to work. The cooks and the smiths can each know what a pot hook is and its price. They can compare that price to the cost of making the pot hook (the smiths' calculation) and the cost of other cooking devices (the cooks' calculation). If the price is high enough, the smiths will be motivated to make pot hooks; and if the price is low enough, the cooks will buy and use pot hooks. The price of pot hooks thus serves as a signal to producers and users alike: in our language, it is *data* in the task and transfer network. At the same time, the transfer of money (or other consideration) from the cooks to the smiths keeps each subnetwork in resource balance. Sixty years ago, Friedrich Hayek marveled at the efficient working of this decentralized, self-balancing, adaptive system:

> The most significant fact about this system, is the economy of knowledge with which it operates, or how little the individual participants need to know in order to be able to take the right action. In abbreviated form, by a kind of symbol, only the most essential information is passed on… .[36]

[35] This is an instance of Stigler's (1951) general observation that small markets call for vertically integrated production.

[36] Hayek (1945) p. 527.

Transactions, of course, are what generate this "most essential information."

The second reason to have a transaction is to adjust the incentives of the transacting parties. Humans are not only cognitively limited, but opportunistic as well. As most parents know, placing a transaction—a shared definition, a means of counting, a value and a payment—between two self-interested parties can get one to do what the other wishes, with alacrity and without coercion. Thus parents sometimes create transactions with their children, not to economize on information, but "to interest their self-love, … and show them that it is for their own advantage" to do what is required.[37] Transactions induce people to perform unwelcome tasks. Transactions also isolate groups of tasks, and thus reduce opportunities for individuals to free-ride on the efforts of others.[38] More than two hundred years ago, Adam Smith marveled at *this* aspect of transactions:

> Give me that which I want, and you shall have this which you want, … it is in this manner that we obtain from one another the far greater part of those good offices which we stand in need of.[39]

Transaction-free Zones and Encapsulated Local Systems

We have argued that the "natural" or "least-cost" places to locate transactions are at so-called thin crossing points—those places in the task network where information hiding between two subnetworks is high, and transfers commensurately simple and few in number. However, in many places, the transfers needed for production are numerous and complex: difficult to count and even more difficult to value. For example, how can we standardize, count and value the transfers that occur when a master smith watches an apprentice shape a pot hook on a forge?

---

[37] Smith (1776, reprinted 1994) p. 14.

[38] Offsetting these advantages, transactions and concomitant property rights make the system vulnerable to new forms of opportunism, notably "hold-up."

[39]Smith, *op. cit.* p. 15.

Watching creates a transfer of data from the apprentice to the master smith. The transfer is costly: watching prevents the master from doing other things. The value of this transfer, however, is contingent: it depends on what happens next. If the apprentice makes no mistakes, then the value of oversight will be relatively low. But if the apprentice makes a mistake, the master can intervene in the process, initiating new transfers of material, energy and information. In so doing, the master may save the pot hook and teach the apprentice how not to err in the future.

Practically speaking, it is impossible to construct transactions that would mirror this complex, contingent and interdependent set of transfers.[40] However, complex, contingent, interdependent transfers are extremely common in real systems of production. Simple transfers with low mundane transaction costs are the exception not the rule.

Fortunately, humans have devised ways to effect complex transfers without making each and every one a transaction. The basic strategy involves creating a *transaction-free zone*. Transaction-free zones are physical and/or social spaces where, by convention, a designated set of transfers occurs freely. In our example, the smithy and the kitchen are both transaction-free zones with respect to the tasks and transfers of iron-working and cooking. Indeed transaction-free zones are common in human affairs: every time we strike up a conversation, we are in effect creating a small, temporary transaction-free zone.

Transaction-free zones are easy to create, but not so easy to police. Property rights, by definition, are suspended in transaction-free zones. As a result, it is easy for valuable things to leave a zone that has no "walls," and rational agents will be reluctant to bring things they value into such zones. Transactions, however, can be used to count and provide compensation for

---

[40] Arrow (1964) and Debreu (1959) created a theoretical price system spanning all contingent transfers no matter how complex or microscopic. The costs of standardizing, counting, valuing and paying for microscopic transfers are what prevent the Arrow-Debreu economy from being a real economy.

transfers into and out of a transaction-free zone. In this fashion, property rights can serve as walls surrounding the zone, and transactions can function as gateways into and out of it.

If all valuable transfers into and out of a given zone are made into transactions, then the zone becomes an *encapsulated local system*. A good example of a encapsulated local system is a modern supermarket. The supermarket's space is physically bounded by walls, with limited ways to get in and out. The store owns what is in the space. However, within the store, shoppers are free to move items around, putting them in and taking them out of carts. Thus the aisles and shelves of the store are a transaction-free zone. Many complex transfers of material, energy and especially information[41] take place in this space: these transfers are not transactions, but they are necessary to the smooth functioning of the retail operation. At the same time, valuable goods are transferred in and out of the store via transactions. Products come into the store via the loading dock, where crates and boxes are counted and paid for.[42] Symmetrically when shoppers move through the checkout lane, the items in their carts will be counted and paid for. (Standardization takes place before the goods even enter the store. Valuation takes place when the shopper looks at the price of an item and puts it in the cart or back on the shelf.)

Transactions at the Boundary of an Encapsulated Local System

Some of the transactions at the boundary of an encapsulated local system are easy to locate and design. Counting boxes and cans is easy, both on the loading dock and at the checkout counter. Thus, as we said before, the task and transfer network provides some natural thin

---

[41] Consider the information embedded in the store's layout and signage, plus all the labels on all the goods. Much of this is tag information, designed to help shoppers find what they want (and perhaps attract them to things they don't want).

[42] The payment takes place by recording a debt between the store and the manufacturer. In most cases, signing for the receipt of the goods creates the debt.

crossing points where transfers are relatively few and simple, information sets non-overlapping, and mundane transaction costs low.

Other transactions are harder to design. For example, bringing labor and capital into an encapsulated local system is often quite tricky. In pre-modern economies, labor would often enter the task network by birth, adoption, or bondage: the grocer's assistant would be his son or his slave. And capital would enter through marriage or inheritance or as trade credit attached to a goods transaction.[43] But in modern economies, people are hired and capital raised via transactions. In these cases, something—effort or money—is transferred; the thing transferred is standardized in some fashion; it will be counted in some fashion; and it will be paid for in some fashion.

The key phrase here is "in some fashion." *By intent*, precise standardization, counting, and compensation are impossible (or prohibitively expensive) for most of the critical transfers within a transaction-free zone. Thus, of necessity, the transactions that bring agents and capital into the zone cannot perfectly reflect what happens inside.[44] But they don't have to. If the encapsulated local system *as a whole* can pay all of its suppliers and have something left over, then it will be *financially sufficient*. In a free-market economy,[45] financial sufficiency gives a local system—a structured subnetwork—the right to survive. In this fashion, a financially sufficient,

---

[43] For descriptions of various pre-modern arrangements for obtaining labor and capital, see, for example, Bloch (1961), Braudel (1882), North (1990), and Greif (1994, 1998).

[44] Given the inherent imperfection of employment and capital contracts, it is not surprising that much of the modern literature in organizational economics and contract theory focuses on these transactions.

[45] By "free-market" economy we mean one in which encapsulated local systems can be created *at low cost* by *local agreements* between and among individual agents. The rights of association do not have to be formal or legally constituted, but they must be effective, *de facto* rights. Many traditional and communist societies do not (or did not) give rights of association to their members (Shleifer and Vishny, 1994). Others have conveyed theoretical rights of association to their members, but have made such associations prohibitively expensive to set up (deSoto, 2000).

encapsulated local system of tasks, transfers and agents can become an autonomously governed organization in the larger economy.[46]

Transactional encapsulation is a complex social technology.[47] The technology has changed over time, becoming more efficient, and it has also diffused across cultures. Some elements have been codified in the laws of sovereign states, for example the laws governing incorporation and conferring limited liability on corporations.[48] Other procedures, such as double-entry book-keeping,[49] the functional design of an enterprise,[50] and the design of financial claims and payment systems,[51] fall under the aegis of the theory and practice of management.

In general, the social technology of encapsulation works as follow. First, an agent or agents must identify a group of tasks and transfers involving multi-directional flows of material, energy and information. Such *interdependent tasks and transfers* show up as blocks in a task structure matrix. The activities of the smiths formed one such block, the activities of the cooks formed another. The second step is to set boundaries for this group of tasks and transfers. Which tasks and transfers will be inside and which outside the local system? The bounded group is a well-defined subnetwork within the greater system. By design, many transfers within its boundaries will be complex, non-standard and difficult to value—they will take place within a *transaction-free zone*. The third and last step is to bring agents and resources into the zone, and move products out of it. This is done by means of *imperfect but cost-effective transactions and contracts*.

---

[46] Aoki (2001); Furubotn (2001).

[47] Nelson and Sampat (2001).

[48] North (1990); Hansmann and Kraakman (2000); Moss (2002); Hansmann, Kraakman and Squire (2002).

[49] Crosby (1997).

[50] Chandler (1977); Galbraith (1977).

[51] Merton and Bodie (1995, 2002).

The specific transfers within and across the boundaries of an encapsulated local system can be designed to suit its particular needs. Some transfers, such as telephone calls, may cross the boundaries without becoming transactions. Conversely, some transactions may take place within the boundaries: for example, employees may pay for lunch at a company-owned cafeteria. And there may even be subcapsules (divisions or departments), which have standardized, counted, and compensated transactions between them.

From a network design perspective, the purpose of having encapsulated local systems in a larger system of production is to facilitate complex transfers without making all of them into transactions. Hence in a well-designed network, we should see complex transfers with taking place within transaction-free zones, and simple transfers with low mundane transaction costs made into transactions at the boundaries of those zones. The transactions in turn serve as the interfaces by which encapsulated local systems deal with one another and coordinate their activities. Importantly, designers can construct this type of task network using local knowledge and lateral communication alone. A well-designed network can emerge as a self-organizing system without central control.[52]

Firms are a form of encapsulated local system.[53] Indeed, firms can be viewed as social artifacts *designed for the purpose* of encapsulating complex transfers of material, energy and

---

[52] Property rights theorists would say that there needs to be a government and an over-arching legal system to make transactions efficient and to enforce the boundaries of local systems. We agree. However, the amount of centralized coordination in economies with encapsulated local systems and transactions is very small in relation to the amount coordination that takes place locally within and between the subnetworks. On self-organization as a general phenomenon, see Kauffman (1993, 2000). On the efficacy of lateral coordination mechanisms in an abstract task and transfer network, see Woodard (2006, Chapter 3).

[53] Langlois (2002) makes a similar argument. His emphasis is on the bundling and unbundling of property rights to reduce externalities: "[T]he creation of 'new' rights and rebundling of existing rights are really manifestations of the same underlying process…. In all these cases, the driving objective is to internalize externalities subject to the costs of setting up and maintaining the rights… [One strategy is to place] all the interactions within a single module, where presumably they could be dealt with more cheaply." What Langlois labels "externalities" or "interactions," we call "transfers" that generate causal "dependencies." Thus "internalizing externalities" is equivalent to "encapsulating blocks of transfers." We have arrived at the same place by somewhat different routes.

information.[54] Families, isolated villages, and nomadic tribes are also local systems that encapsulate complex transfers, but, in contrast to firms, their encapsulation is fortuitous: they were not created for that purpose.

Why Encapsulate?

All complex systems appear to need encapsulating mechanisms. Herbert Simon was one of the first to understand this general principle. In the first chapter of *The Sciences of the Artificial*, he emphasized the need to separate the "inner" parts of a goal-directed structure from the "outer" environment:

> It is an important property of most good designs, whether biological or artifactual… [that] the designer insulates the inner system from the environment, so that an invariant relation is maintained between the inner system and goal, independent of variations …[in] the outer environment. … Quasi-independence from the outer environment may be maintained by various forms of passive insulation, by reactive negative feedback…, by predictive adaptation, or by various combinations of these.[55]

In other words, encapsulation is essential, but it can be achieved in many different ways. For example, biological systems exhibit encapsulation at various levels: The nucleus, which contains genetic material, is an encapsulated subsystem within a cell; cells are encapsulated within cell walls; within multi-celled organisms, the germ-line is encapsulated in the reproductive organs; organisms have encapsulated bodies; and social insects, such as ants and termites, build encapsulated nests.[56]

In both biological and economic systems, encapsulation appears to have two main effects. First, as Simon noted, it protects material, energy and information transfers inside the

---

[54] On the evolution of law related to corporate boundaries and on the critical importance of protecting a firm's assets from seizure by the creditors of shareholders ("asset partitioning"), see Hansmann and Kraakman (2000) and Hansmann, Kraakman and Squire (2006). Langlois (2006) points out that clubs are another social artifact designed to encapsulate complex transfers.

[55] Simon (1999, p. 8.)

[56] Maynard Smith and Szathmary (1995); Gerhart and Kirschner (1997); Turner (2000).

capsule from being disturbed by outside interference. Second, it permits the storage of material, energy and information inside the capsule over time.[57] In this fashion, encapsulation allows stable local systems, involving complex material, energy, and information transfers and temporary imbalances of material and energy flows, to exist.

In biological systems, encapsulated local systems like single and multi-celled organisms can compete with their unencapsulated surroundings *and with each other* for scarce resources. Importantly, in some cases, encapsulated local systems can also cooperate and perform specialized complementary functions. Such cooperation gives rise to the internal structure of the cell, multicellular organisms, and social species, like ants. The most successful singletons and groups will survive, either physically or, more commonly, as biochemically or genetically replicated information patterns.[58]

The surviving local systems, in turn, can sometimes (but not always) be incrementally modified, especially if their internal structure is *modular*. Thus, in addition to enabling stable local systems to form, encapsulation gives rise to two new "levels of selection:" the capsules themselves and groups of complementary capsules. The emergence of these new levels of selection in turn changes the evolutionary or adaptive trajectories that are open within the larger system. *This change in evolutionary potential is the third effect of encapsulation.*[59]

In an economy, as we said, transactionally encapsulated local systems include individuals, families, and corporations. Transactional encapsulation has many of the same effects that physical and chemical encapsulation has in biological systems. First, it can be used to isolate

---

[57] Turner (2000).

[58] Gerhart and Kirschner (1997); Haig (1997); Margulis (1970, 1981).

[59] Kirschner and Gerhart (1998).

"fragile" parts of the task and transfer network from external shocks.[60] Second, it allows not only specific materials but also general claims (money and credit) to be stored within a local system, e.g., a family or a corporation. Those stored resources in turn can be used to adjust for temporary imbalances within the local system and between the local system and the larger system. Finally, transactionally encapsulated local systems can compete *as units* in the larger economy. Those individual capsules and complementary groups that are most successful in securing and storing resources will survive, perhaps indefinitely. And those that can be modified at low cost may evolve in response to competitive pressures.

In summary, at any given time and place, the task and transfer network will have a structure, which can be made visible using task structure matrix mapping. From a network design perspective, *blocks* of complex transfers should be located in transaction-free zones, which in turn should be encapsulated via property rights and transactions. The transactions should be located at the *thin crossing points* of the network, where transfers are few and simple. At these points, transfers are relatively easy to standardize, count, and value, hence mundane transaction costs will be low.

However, real task and transfer networks do not consist entirely of dense blocks and thin crossing points. Many intermediate network structures are possible, and thus the "best place" to locate a transaction is not always as obvious as in the case of the smiths and the cooks. Complicating matters further, the network itself can be modified. Often it can be *pinched* at a particular juncture to make the crossing point thinner. At the same time, transactions can be designed to match the nature of the crossing point. For example, *relational contracts*, a special form

[60] On the disruptiveness of external shocks and the benefits of process encapsulation, see Herbert Simon's famous parable of the two watchmakers (Simon, 1999, pp. 188-190).

of transaction, are especially useful at thick crossing points. In the next section, we describe how a task and transfer network and a transaction can be codesigned to be mutually supportive.

"Pinching" the Task and Transfer Network

A particular transaction between an network plastics company and an automobile manufacturer serves to illustrate how the task and transfer network can be "pinched" to create a thin crossing point between two firms. The example also shows how relational contracts make transactions possible even when the underlying subnetworks are somewhat overlapping and interdependent.

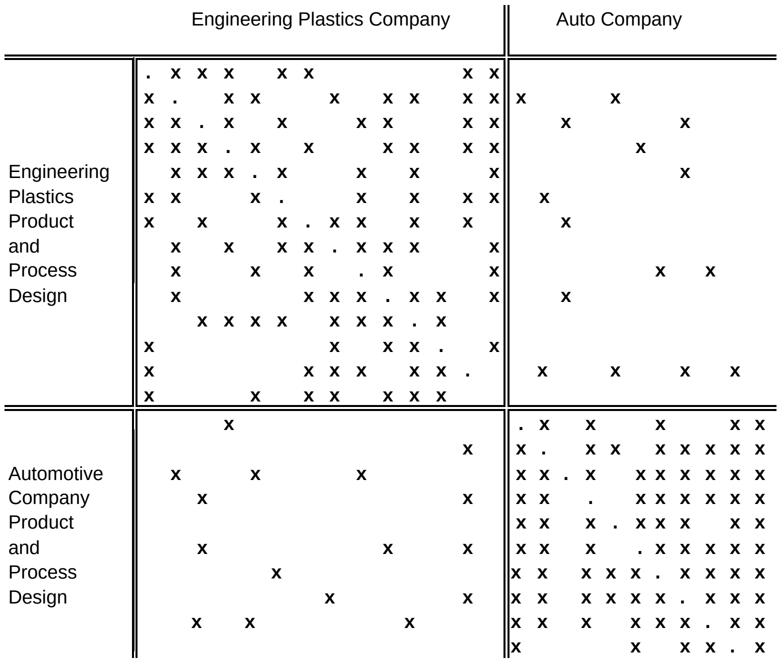
In 1994, an automobile manufacturer sought to find a new plastic with high heat resistance for automobile interiors. Kim Clark observed how the automobile company (the user) managed its subcontracting relationship with the plastics company (the producer):

> [T]he automotive customer developed "specifications" that the new material had to meet in order to qualify for and win the business. There were eight items in the specification, including heat resistance, cost, strength and so forth. Each specification was accompanied by a testing protocol and a standard that the material had to meet.[61]

Figures 2 and 3 below show how the creation of a formal contract affected the task and transfer network. Figure 2 is a schematic diagram of the "raw" or "natural" dependencies between the plastic compound and the automobile. These dependencies were numerous and flowed both ways. That is, the material properties (e.g. weight, viscosity, shapes), the energetic properties (e.g., shock absoption or brittleness) and the informational properties (e.g. color, texture) of the plastic would affect the automobile at each point where the plastic was used. But those same material, energetic and informational properties would affect the ease of finding the right chemical compound and the cost of making it.

[61] Clark (1995).

Figure 2
The "Natural" Task and Transfer Network for the Plastic Compound Design Process
Each Out-of-Block "x" Represents a Dependency to be Resolved via Consultations



The natural dependencies between the compound and the automobile needed to be resolved during the design process to achieve a satisfactory outcome. In the normal course of events, resolving these dependencies would have required many consultations, involving transfers of material, energy and especially information between the two design teams.[62] Because of these transfers, there is no thin crossing point in Figure 2— no obvious place to locate a transaction. Instead, the design process required so many transfers that it was impossible to count and arrange compensation for each one. Moreover, each point of dependency contained a potential conflict of interest. The better choices from the auto company's perspective were likely

---

[62] The mirroring of design dependencies and transfers of design information is a well-established principle in design theory. See, for example, Eppinger (1991), Baldwin and Clark (2000), and Sosa, Eppinger and Rowles (2004).

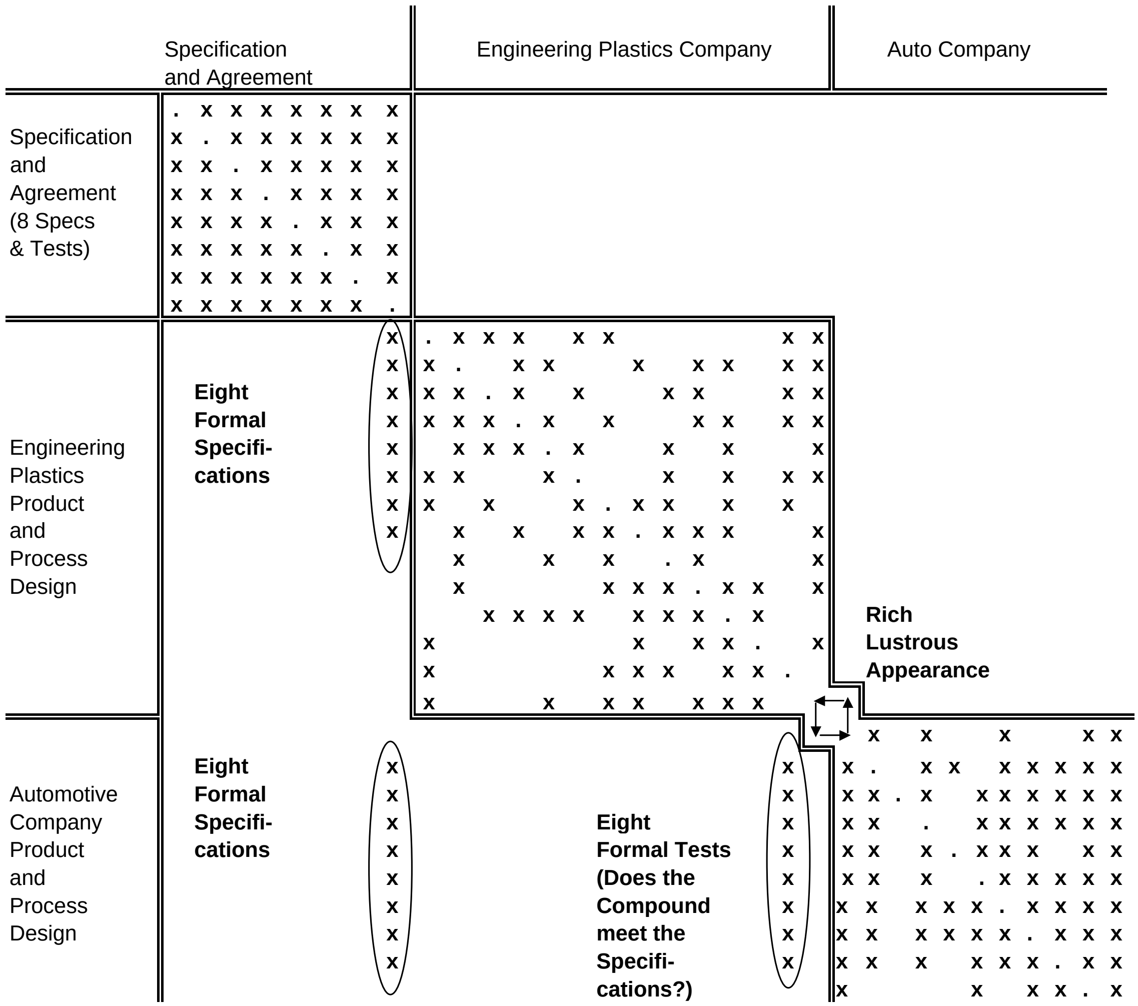
to be worse from the plastic company's point of view. These inherent conflicts meant that each consultation about a dependency had the potential to turn into a holdup or a lengthy negotiation.

There were several things the auto company could have done that would not have involved changing the task and transfer network. One possibility was to combine ("de-encapsulate") the two firms. For example, the auto company might have acquired the plastics company, or it could have developed the compound inhouse. These actions would have created a larger "transactions-free zone" spanning both subnetworks. Another possibility was to set up imprecise, but easy-to-count measures, and use those as the basis of the transaction. For example, the auto company might have paid the plastics company on a cost-plus basis. It is well-known that cost-plus contracts create perverse incentives, but the auto company might have seen this as their best alternative.

The auto company (with the consent of plastics company) chose to do something else. *They changed the structure of the original task and transfer network by standardizing the object that would be transferred.* Figure 3 shows the new network. The matrix has three big blocks, corresponding to: (1) specification and agreement; (2) the design of the compound by the plastics firm; and (3) the design of the automobile.[63]

[63] These blocks are not drawn to scale: in reality the design blocks involved many more tasks and transfers than the specification block.

Figure 3
The Task and Transfer Network Chosen for the Plastic Compound Design Process



In the specification and agreement phase of this transaction, representatives of the auto company first used their engineering knowledge to *predict and resolve* the natural dependencies shown in Figure 2.[64] As it happened, most of the dependencies could be codified in terms of eight

[64] To construct a useful set of specifications, designers at the auto company had to have knowledge about how to formulate and manufacture plastic. Thus information hiding was not as complete as in the case of the smiths and the cooks. This is a case of what Brusoni and Prencipe (2001) call "firms knowing more than they do." On the network design of specifications for purposes of outsourcing, see Fine and Whitney (1996).

measurable specifications. These specifications standardized the object of exchange (the compound). The two parties to the transaction then had a common definition of the product and an unambiguous way of determining when the design job (identifying the compound) was done to the auto company's satisfaction. In effect, the eight specifications with associated tests and standards served as an intermediate good in the production process. The specifications, tests and standards were formally agreed to by the plastics company, and communicated to designers at both firms. In Figure 3, this transfer of information is denoted by the two vertical sets of "**x**"s labeled "Eight Formal Specifications."

The agreed-upon specifications became *design rules* for the teams at both companies. In this fashion, the two companies resolved *ex ante* a large number of potential questions and disputes that might have arisen. Because these dependencies had been resolved, the need for ongoing consultations between the two design teams went down dramatically. Thus in contrast to Figure 2, Figure 3 has an obvious thin crossing point. The original task and transfer network appears to have been *pinched* in the middle.

Pinching is the act of systematically removing dependencies between two subnetworks by creating design rules binding on both of them.[65] Pinching allows two naturally interdependent subnetworks to be separated and "hidden" from each other. As we have seen, information hiding enhances the division of cognitive labor: the auto designers did not have to learn polymer chemistry, and the chemists did not have to learn mechanical engineering. At the same time, pinching limits and standardizes transfers between the subnetworks. In place of numerous consultations and debates as to whether a proposed compound was right, each proposal was subjected to eight tests—eight measurements—and if it passed all eight, was deemed satisfactory.

[65] Baldwin and Clark (2000, pp. 64-70) call this process "design rationalization."

(In reality, things were not quite this cut-and-dried: see the discussion of "rich, lustrous appearance" below.)

Like everything else, pinching a task and transfer network is costly. Specifying the plastic's performance added a whole new block of tasks to the original network. There were also costs of communicating the specification, performing the tests, and implementing progress payments. These were all *mundane transactions costs*: the specification standardized the compound, the tests measured its quality, and the progress payments provided compensation. However, both companies judged these costs to be worthwhile for two reasons. First, the new task and transfer network called for fewer consultations between the two design teams, which reduced ongoing coordination costs for both firms. Second, standardizing and testing the product along eight dimensions of performance reduced the ambiguity of the contract. As is well known, contractual ambiguity opens the door to opportunism and opportunistic transaction costs.[66] Thus for both firms, incurring higher mundane transactions costs decreased the likelihood of subsequent disagreements, holdups, and litigation, and thus reduced opportunistic transaction costs.[67]

However, the substitution of mundane transaction costs for ongoing coordination costs and opportunistic transaction costs is not the whole story of this contract. Equally interesting is the fact that some dimensions of performance for the plastic were *not* standardized. Thus in important ways, the contract was incomplete.[68] Quoting Clark again:

> [A]s development proceeded, it became clear ... that there were other characteristics of the material that were very important to important players in the auto company, which

[66] Ambiguous contracts are legally unenforceable. Most transaction-cost and contract theory models begin with the premise that some contracts are unavoidably ambiguous ("non-verifiable"), and proceed to analyze the consequences of this fact. See, among others, Williamson (1985); Grossman and Hart (1986); Hart and Moore (1990); Holmstrom and Milgrom (1994); Baker, Gibbons and Murphy (2002).

[67] A lower probability of holdup also reduced both firms' incentives to make defensive investments (or non-investments). See Hart (1995).

[68] Hart (1995).

> were not in the specs. (Example: the interior designers wanted a material with a "rich, lustrous appearance.") They were not in the specs, because the auto company had no way to make the requirement specific, no testing protocol and no standard to use in the specifications.
>
> The only way to uncover these critical but unspecified parameters was [for the network plastics firm] to … develop material for the [auto company] to test… and to move quickly to generate test quantities of material over and over again.[69]

A certain amount of indeterminacy was tolerated because the two firms enjoyed an ongoing commercial relationship, or "relational contract."[70] One particular property of the plastic, "rich, lustrous appearance," was both critical and highly ambiguous. The auto designers knew of no objective test for this property: it was a case of "I know it when I see it." The search for a compound with sufficiently rich, lustrous appearance caused many consultations and transfers of samples back and forth between the two companies. In Figure 3, these transfers, *which were neither counted nor paid for*, are shown as a circuit of arrows at the thin crossing point.

In effect, to resolve questions of "rich, lustrous appearance," the two companies had to become locally and temporarily de-encapsulated *with respect to this issue*. They had to allow a whole set of material, energy and information transfers across their boundaries, which were unstandardized, uncounted, and uncompensated. In allowing such transfers, each firm risked being overburdened and undercompensated by the other. But, because there was trust between individuals and the expectation of a continuing commercial relationship, a small transactions-free zone could be sustained between the two firms.

In summary, the task and transfer network is not cast in stone. Complex and contingent transfers between subnetworks can sometimes be eliminated by setting up prior rules binding on both sets of agents. In this fashion, a thick crossing point can be made thinner: when this

[69] Clark (1995).

happens, we say that the task network has been "pinched". Pinching is costly, but it reduces ongoing coordination costs and makes contract terms less ambiguous, thereby reducing the scope of opportunistic behavior. The costs of pinching are mundane transaction costs: they can be substituted for coordination costs and opportunistic transaction costs to achieve a more favorable overall outcome. Pinching does not have to proceed to its theoretical limit, however. Some complex and contingent transfers between a supplier and a customer may be needed to achieve the desired outcome. In such cases, a relational contract can support the creation of a limited transaction-free zone between two otherwise encapsulated enterprises.

Modularity and Mundane Transaction Costs

"We have been talking about modularity all along."

At the beginning of this essay, we promised to relate mundane transaction costs to the *modular structure* of the task and transfer network. The time has come to reveal that, just as Moliere's bourgeois gentleman had been speaking prose all his life, we have been talking about modularity all along.

A complex system is said to exhibit *modularity* if its parts operate independently, but still support the functioning of the whole. Modularity is not an absolute quality, however. Systems can exhibit different modular structures and different degrees of interdependence between their respective elements.[71]

---

[70] Baker, Gibbons and Murphy (2002).

[71] This definition is taken from Rumelhart and McClelland (1995), quoted in Baldwin and Clark (2000). Although Herbert Simon did not use the term "modularity," the attribute we now call by that name is essentially identical to the property he called "near-decomposability" (Simon, 1999; Simon and Augier, 2002.) On the economic properties of modular systems, see, among others, Langlois and Robertson (1992); Baldwin and Clark (1992, 2000); Ulrich (1995); Garud and Kumaraswamy (1995), Sanchez and Mahoney (1996); Schilling (2000); Aoki (2001); Langlois (2002); and Aoki and Takizawa (2002a, b).

In a system of production, a module is a group of tasks and transfers that are densely connected to one another, but only loosely connected to the other parts of the network. In a task structure matrix, modules appear as densely connected blocks. If there are only a few, simple out-of-block transfers, then the underlying network is highly modular. As the number and complexity of out-of-block transfers increases, the modularity of the network decreases.

Finally modules hide information.[72] In a large system, the complexity of an element can be isolated by defining an *abstraction* for the element and an *interface*. The abstraction hides the complexity of the element and the interface allows it to interact with the larger system. When most of the internal structure of the element is not visible to the rest of system and its interfaces are simple and stable, the element is said to be encapsulated and becomes a module. But, of course, there are degrees of encapsulation and degrees of modularity.

The arguments presented in this essay can now be condensed into a single statement: *The modular structure of a task and transfer network determines its pattern of mundane transaction costs, hence where cost-effective transactions can be located, and how those transactions should be structured*. Mundane transaction costs are lowest at thin crossing points, i.e., between modules. They are prohibitive within densely connected blocks, i.e., within modules. And they are high, but sometimes manageable, at thick crossing points between interdependent, but somewhat segregated groups of tasks.

Mundane transactions costs must be sufficiently low in a given location for it to be possible to have a transaction there. We believe this was what Williamson had in mind when he said transactions take place at "technologically separable interfaces." From an network design

[72] Information hiding, also called information encapsulation, is a central goal in the design of modular systems. For a discussion of what information hiding means, and why it is desirable in complex designs, see Parnas (1972a,b), Parnas, Clements and Weiss (1985), Baldwin and Clark (2000), and Sullivan et.al. (2001). In economics, Cremer (1980), Aoki (2001) and Aoki and Takizawa (2002) discuss when information encapsulation is advantageous in organizations.

perspective, technologically separable interfaces are those places in the task network where labor is divided and information hidden between two blocks. As a result, transfers between the blocks are relatively few and simple, and mundane transactions costs are not prohibitive.

Sometimes the modular structure of a particular task network is highly constrained by the laws of physics and logic. The smithy and the kitchen were natural modules, exhibiting an almost complete division of physical and cognitive labor. Between the two, there was a natural thin crossing point, which was the "obvious" place to locate a transaction.

In other cases, however, the laws of nature give designers more latitude. In the plastics example, there were three potential modular structures: (1) the "natural" structure depicted in Figure 2; (2) the "proto-modular" structure in Figure 3; and (3) a "fully modular" structure that would have eliminated the complex transfers caused by the "rich lustrous appearance" issue. The designers of the system chose the second of these three alternatives, and matched it with a relational contract. In the process, they made a number of tradeoffs and judgments. They judged that the quality of the final compound would not suffer too much from the lack of ongoing consultations between the two groups. They guessed that a compound meeting the specifications could be found by the plastics company in a reasonable amount of time. They traded off the mundane costs of setting up and enforcing the specifications against ongoing coordination costs and the opportunistic costs of contract ambiguity. They also traded off the benefits of having a limited transaction-free zone against the opportunistic costs of being overburdened or held up on the "rich, lustrous appearance" issue. Finally they judged that they would be better off working in two separate companies rather than one.

At thick crossing points in the task network, judgments and negotiations over questions like this go on all the time. In such cases, everything is up for grabs: the network's modular structure, property rights, incentives, mundane transaction costs and opportunistic transaction costs. Thus a comprehensive theory of transaction design—one with power in practice—must take all these issues into account and place them within a consistent framework. Moving beyond transactions, a comprehensive theory of organizations must address what goes on in encapsulated local systems, both in their transaction-free zones and at their boundaries. Finally, a comprehensive theory of firms in the wider economy must explain the implications of the rule "financially sufficient local systems can survive," which is a bedrock feature of all free-market economies.[73]

Such a theory is well beyond the scope of this essay. Thus we will close with just a few words about innovation.

"And what about innovation?"

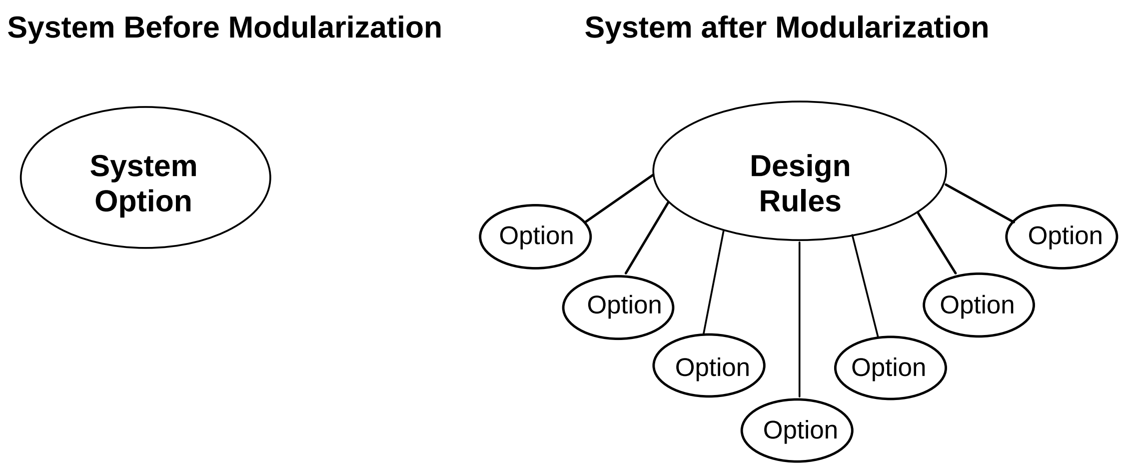
For the most part, transaction costs economics and contract theory look at static systems of production. In this paper, we have argued that the location of transactions is determined by the modular structure of an underlying task and transfer network. But this network is by no means static. Network designers, both engineers and managers, change the structure of the network all the time in search of local efficiency and competitive advantage.

In prior work we have shown that modular structure and innovation interact in a powerful way. Specifically, a modular design for a complex system increases the expected rewards and reduces the cost of search and experimentation on subsets of the system. While an "integral" system contains only one design option (to take the whole system or leave it), a modular system contains (at least) as many options as there are modules (see Figure 4). As a result, designers have strong incentives to increase modularity when the innovation potential of a

[73] Furubotn (2001) begins to work on this agenda. As he points out, the picture of the economy that emerges is vastly different than the views of either neoclassical economics or standard new institutional economics.

complex artifact or system of production is high. Indeed, the desire to realize the option value of modular innovation can provide the impetus for creating entirely new modular architectures.[74]

Figure 4
Modularity Multiplies Design Options



Modules *by definition* create thin crossing points in the task and transfer network—places where mundane transaction costs are relatively low. Thus, whether the architects realize it or not, increasing a system's modularity necessarily changes the pattern of mundane transaction costs. The new, more modular architecture will have more thin crossing points with low mundane transaction costs—places that did not exist (or were thick crossing points) in the older, more integral architecture. *Thus a modular architecture created for the sake of promoting innovation will have the secondary consequence of permitting more transactions, thus allowing more firms to participate in the overall task network.*

The converse is also true: modularity pursued for the purpose of lowering transaction costs can have the secondary consequence of fostering innovation and experimentation. For

[74] This argument is developed in great detail in Baldwin and Clark (2000).

example, suppose the engineering plastics firm insisted that the auto company specify all the dimensions of performance of the compound. The plastic design and the auto design processes would then have been perfectly modular, with each effectively hidden from the other. Both firms would have avoided the contractual ambiguity of the "rich and lustrous appearance" issue and the transfer costs involved in resolving it. But, in principle, any other company could have submitted a compound for testing. The cost to the auto company of allowing such competition would be simply the cost of conducting more tests, and the benefits would have been the chance of obtaining a better compound. Design competitions, which are common in some industries, work this way and are known to be hard on suppliers. So quite possibly, the plastics company was delighted to have an ambiguous dimension of performance to forestall innovative competitors!

"Where do transactions come from?"

We can now answer the question in the title of this paper. Transactions come from the modular structure of the underlying network of tasks and transfers of material, energy and information. *By intent*, transfers within modules in the network are numerous and complex, while transfers between modules are relatively few and simple. The between-module transfer points thus constitute a set of "thin crossing points" in the overall network. These are places where the division of labor and information hiding are high, where transfers are few and simple, and where the mundane costs of converting a transfer (or set of transfers) into a transaction are low. Cost-effective transactions can be located at these points.

The mundane costs of converting a set of transfers into a transaction are the costs of standardizing the object being transferred, counting it, valuing it, and paying for it. These costs are different from the opportunistic costs of holdup, haggling, dispute resolution and defensive investment that are the focus of transaction cost economics and contract theory.

Complex and contingent transfers are still necessary to the functioning of the productive system: indeed, they are the rule, not the exception. From a network design perspective, such complex transfers should take place in transaction-free zones within encapsulated local systems. In fact, firms are social artifacts designed for the purpose of encapsulating the complex transfers required by most production technologies. And in modern economies, the social technology of encapsulation and interfacing—of creating modules in the network and coordinating them via transactions—is highly advanced, and deserves further study.

With encapsulation and interfacing mechanisms in place, a widely dispersed network of autonomous (and opportunistic!) agents can coordinate flows of material, energy and information to design and produce a vast diversity of goods *without central control*. Thus the combination of (1) a modular network design; (2) encapsulated local systems to provide transaction-free zones; and (3) transactions to serve as signals and maintain resource balance allows a very large task and transfer network to operate in a decentralized yet coordinated way. What's more, by innovating on modules, the network itself can evolve piecemeal, with most innovations causing only local disruptions of the overall system.

It is indeed a remarkable design!

References

Akerlof, George A, (1970) "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism," *The Quarterly Journal of Economics*, 488-500 Vol. 84 (3) pp. 488-500

Alchian, Armen A. and Harold Demsetz (1972) "Production, Information Costs, and Economic Organization," *American Economic Review* 52:777-795.

Alexander, Christopher (1964) *Notes on the Synthesis of Form*, HarvardUniversity Press, Cambridge, MA.

Aoki , Masahiko (2001). *Towards a Comparative Institutional Analysis*, MIT Press, Cambridge, MA.

Aoki, Masahiko (1988) *Information, Incentives, and Bargaining in the Japanese Economy*, Cambridge University Press, Cambridge, UK.

Aoki, Masahiko and Hirokazu Takizawa (2002a) "Incentives and Option Value in the Silicon-Valley Tournament Game, RIETI Discussion Paper Series, 02-E-001, Tokyo, Japan.

Aoki, Masahiko and Hirokazu Takizawa (2002b) "Modularity: Its Relevance to Industrial Architecture," presented at the Saint-Gobain Centre for Economic Research 5th Conference, Paris, FR, November.

Arora, Ashish and Robert P. Merges (2004) "Specialized Supply Firms, Property Rights and Firm Boundaries," *Industrial and Corporate Change*, 13(3):451-475.

Arrow, Kenneth J. (1964) "The Role of Securities in the Optimal Allocation of Risk Bearing," *Review of Economic Studies* 31(April): 91-96.

Ashby, W. Ross (1960) *Design for a Brain*, 2nd Ed. Wiley, New York, NY.

Baker George, Robert Gibbons and Kevin J. Murphy (2002) "Relational Contracts and the Theory of the Firm," *The Quarterly Journal of Economics* 117:39-84.

Baldwin, Carliss Y. and Kim B. Clark (1997) "Managing in the Age of Modularity," *Harvard Business Review* Sept/Oct: 81-93.

Baldwin, Carliss Y. and Kim B. Clark (2000). *Design Rules, Volume 1, The Power of Modularity*, MIT Press, Cambridge MA.

Baldwin, Carliss Y. and Kim B. Clark (2006) "Between 'Knowledge' and the 'Economy': Notes on the Scientific Study of Designs," *Advancing Knowledge and the Knowledge Economy*, (B. Kahin and D. Foray, eds.) Cambridge, MA: MIT Press.

Barzel, Yoram (1997) *Economic Analysis of Property Rights*, 2nd Edition, Cambridge University Press, Cambridge, UK.

Bloch, Marc (1961) *Feudal Society* (tr. L.A. Manyon) University of Chicago Press, Chicago, IL.

Braudel, Fernand (1982). *The Wheels of Commerce*, (tr. Sian Reynolds) Harper & Row, New York, NY.

Brusoni, Stefano and Andrea Prencipe (2001) "Unpacking the Black Box of Modularity: Technologies, Products and Organizations," *Industrial and Corporate Change,* 10(1):179-205.

Burt, Ronald S. (1995) *Structural Holes : The Social Structure of Competition*, Reprint Edition, Cambridge, MA: Harvard University Press

Burt, Ronald S. (2005) *Brokerage and Closure: An Introduction to Social Capital*, New York, NY: Oxford University Press

Cacciatori, Eugenia and Michael G. Jacobides (2005) "The Dynamic Limits of Specialization: Vertical Integration Reconsidered," *Organization Studies*, 26(12):1851-1883.

Chandler, Alfred D. (1977) *The Visible Hand: The Managerial Revolution in American Business*, Harvard University Press, Cambridge, MA.

Clark, Kim B. (1995) "Notes on Modularity in Design and Innovation in Advanced Ceramics and Network Plastics," Paper presented at the Meeting of the American Economic Association, New York, NY.

Coase, Ronald H. (1937) "The Nature of the Firm," *Economica,* 4(4):386-405, reprinted in *The Firm, the Market, and the Law,* University of Chicago Press, Chicago, IL, 33-55.

Coase, Ronald H. (1960) "The Problem of Social Cost," *Journal of Law and Economics* 3:1-44, reprinted in *The Firm, the Market, and the Law,* University of Chicago Press, Chicago, IL, 95-156.

Cosmides, Leda, and John Tooby (1992) "Cognitive Adaptations for Social Exchange," in *The Adapted Mind: Evolutionar Psychology and the Generation of Culture* (eds. J.H. Barkow, L. Cosmides and J. Tooby) Oxford University Press, New York, NY.

Cremer, Jacques (1980) "A Partial Theory of Optimal Organization of a Bureaucracy," *Bell Journal of Economics*, 11:683-693.

Crosby, Alfred W. (1997) *The Measure of Reality: Quantification and Western Society 1250-1600*, Cambridge University Press, Cambridge, UK.

Debreu, Gerard (1959) *Theory of Value*, Wiley, New York, NY.

Demsetz, Harold (1967) "Toward a Theory of Property Rights," *American Economic Review,* 57:253-257.

deSoto, Hernando (2000) *The Mystery of Capitalism: Why Capitalism Triumphs in the West and Fails Everywhere Else,* Basic Books, New York, NY.

Eppinger, Steven D. (1991) "Model-based Approaches to Managing Concurrent Network" *Journal of Network Design*, 2: 283-290.

Fine, Charles H. and Daniel E. Whitney (1996). "Is the Make-Buy Decision a Core Competence?" MIT Center for Technology, Policy, and Industrial Development Discussion Paper, available at http://web.mit.edu/ctpid/www/Whitney/morepapers/make_ab.html, viewed January 15, 2003.

Fujimoto, Takahiro (1999) *The Evolution of a Manufacturing System at Toyota,* Oxford University Press, Oxford, UK.

Furubotn, Eirik G. (2001) "The New Institutional Economics and the Theory of the Firm," *Journal of Economic Behavior & Organization*, 45(2001):133-153.

Furubotn, Eirik G. and Rudolf Richter (2005) *Institutions and Economic Theory: The Contribution of the New Institutional Economics*, 2nd Edition, Ann Arbor, Michigan: University of Michigan Press.

Galbraith, Jay (1977) *Organization Design*, Reading, MA: Addison-Wesley.

Garud Raghu and Arun Kumaraswamy (1995) "Technological and Organizational Designs to Achieve Economies of Substitution," *Strategic Management Journal,* 17:63-76, reprinted in *Managing in the Modular Age: Architectures, Networks, and Organizations,* (G. Raghu, A. Kumaraswamy, and R.N. Langlois, eds.) Blackwell, Oxford/Malden, MA.

Gerhart, John and Marc Kirschner (1997) *Cells, Embryos, and Evolution,* Blackwell, Oxford/Malden, MA.

Gibbons, Robert (2004) "Four Formal(izable) Theories of the Firm?" manuscript. Also available at http://web.mit.edu/rgibbons/www/ , viewed 4/26/04.

Greif, Avner (1994) "Cultural Beliefs and the Organization of Society: A Historical and Theoretical Reflection on Collectivist and Individualist Societies," *Journal of Political Economy* 102: 912-50.

Greif, Avner (1998) "Historical and Comparative Institutional Analysis," *American Economic Review*, 88 (May): 80-84.

Grey, Thomas C. (1980) "The Disintegration of Property," in *Property: Nomos XXII,* (J. R. Pennock and J. W. Chapman, eds.) New York University Press, New York, NY.

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

Haig, David (1997) "The Social Gene" in *Behavioural Ecology: An Evolutionary Approach. 4th ed.* (J. R. Krebs and N.B. Davies eds.) Blackwell, Oxford/Malden, MA.

Hansmann, Henry and Reinier H. Kraakman (2000) "Organizational Law as Asset Partitioning," 44 *European Economic Review* 807.

Hansmann, Henry, Reinier H. Kraakman and Richard Squire (2006) "Law and the Rise of the Firm," 119 *Harvard Law Review* (forthcoming)

Hart, Oliver (1995) *Firms, Contracts and Financial Structure*, Oxford, UK: Oxford University Press.

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

Hart, Oliver and John Moore (1998) "Default and Renegotiation: A Dynamic Model of Debt," *Quarterly Journal of Economics,* 113:1-41.

Hayek, Friedrich A. (1945) "The Use of Knowledge in Society," *American Economic Review* 35(4):519-530.

Holland, John H. (1996) *Hidden Order: How Adaptation Builds Complexity,* Addison-Wesley, Reading, MA.

Holmstrom, Bengt (1982) "Moral Hazard in Teams," *Bell Journal of Economics,* 13(2):324-340.

Holmstrom, Bengt and Paul Milgrom (1994) "The Firm as an Incentive System," *American Economic Review,* 84(4):972-991.

Ichniowski, Casey, Kathryn Shaw and Jon Gant, (2002) "Working Smarter by Working Together: Connective Capital in the Workplace," Paper presented at the NBER Conference on Organizational Economics, Cambridge, MA, November.

Jacobides, Michael G. (2005). "Industry Change Through Vertical Dis-Integration: How and Why Markets Emerged in Mortgage Banking," *Academy of Management Journal*, 48(3):465-498.

Jacobides, Michael G. (2006) "The Architecture and Design of Organizational Capabilities," *Industrial and Corporate Change*, 15(1):151-171.

Jacobides, Michael G. and Sidney G. Winter (2005) "Co-evolution of capability and transaction costs: explaining the institutional structure of production," *Strategic Management Journal*, 26(5):395 - 413.

Jensen, Michael C. and William H. Meckling (1976) "Theory of the Firm: Managerial Behavior, Agency Costs, and Ownership Structure," *Journal of Financial Economics,* 3(4):305-360, reprinted in *Foundations of Organizational Strategy,* Harvard University Press, Cambridge, MA.

Kauffman, Stuart A. (1993) *The Origins of Order*, New York: Oxford University Press

Kauffman, Stuart A. (2000) *Investigations*, New York, Oxford University Press

Kerr, Steven (1975) "On the Folly of Rewarding A, While Hoping for B," *Academy of Management Journal,* 18:769-783.

Kirschner, Marc and John Gerhart (1998). "Evolvability," *Proceedings of the National Academy of Sciences*, USA, 95:8420-8427.

Kogut, Bruce and Udo Zander (1992), "Knowledge of the Firms, Combinative Capabilities and the Replication of Technology," *Organization Science*, 3(3):383-397

Kraakman, Reinier H. & Henry Hansmann (2002) "Property, Contract, and Verification: The Numerus Clausus Problem and the Divisibilty of Rights," 31 *Journal of Legal Studies* S373.

Kusiak, Andrew (1995) *Network Design*, Academic Press, New York, NY.

Langlois, Richard N. (1992) "Transaction-cost Economics in Real Time," *Industrial and Corporate Change*, 1(1):99-127.

Langlois, Richard N. (2002) "Modularity in Technology and Organization," *Journal of Economic Behavior & Organization* 49(1):19-37.

Langlois, Richard N. (2003) "The Vanishing Hand: the Changing Dynamics of Industrial Capitalism," *Industrial and Corporate Change* 12(2): 351-385.

Langlois, Richard N. (2005) "The Secret Life of Mundane Transaction Costs," *Organization Studies* (forthcoming). Working paper available at http://www.econ.uconn.edu/working/2005-49.pdf (viewed 1/22/06).

Langlois, Richard N. and Paul L. Robertson (1992). "Networks and Innovation in a Modular System: Lessons from the Microcomputer and Stereo Component Industries," *Research Policy,* 21: 297-313, reprinted in *Managing in the Modular Age: Architectures, Networks, and Organizations,* (G. Raghu, A. Kumaraswamy, and R.N. Langlois, eds.) Blackwell, Oxford/Malden, MA.

Levy, John V. (1978) "Buses, The Skeleton of Computer Structures," *Computer Network: A DEC View of Hardware Systems Design,* (C.G. Bell, J.C. Mudge and J.E. McNamara, eds.) Digital Press, Maynard, MA.

Marengo, Luigi and Giovanni Dosi (2005) "Division of Labor, Organizational Coordination and Market Mechanism in Collective Problem Solving," *Journal of Economic Behavior & Organization*, 58(2):303-326.

Marengo, Luigi, Corrado Pasquali and Marco Valente (2003). "Decomposability and Modularity of Economic Interactions, *Modularity: Understanding the Development of Complex Natural Systems*, (W. Callebaut and D. Rasskin-Gutman, eds.) MIT Press, Cambridge, MA.

Margulis, Lynn (1970) *Origin of Eukaryotic Cells,* Yale University Press, New Haven, CT.

Margulis, Lynn (1981) *Symbiosis in Cell Evolution,* W.H. Freeman, San Francisco, CA.

Maynard Smith, John and Eors Szathmary (1995) *The Major Transitions in Evolution,* Oxford University Press, Oxford, UK.

Merton, R. C. and Zvi Bodie (1995). "A Conceptual Framework for Analyzing the Financial Environment," in The Global Financial System: A Functional Perspective, (ed. Crane et al) Harvard Business School Press, Boston, MA.

Merton, Robert C. and Zvi Bodie, (2002) "The Design of Financial Systems: Towards a Synthesis of Function and Structure, Harvard Business School Working Paper #02-074, Boston, MA.

Mirrlees, James A. (1976) "The Optimal Structure of Incentives and Authority within an Organization," *Bell Journal of Economics,* 7(1): 105-131.

Mokyr, Joel (2002). *The Gifts of Athena: Historical Origins of the Knowledge Economy*, Princeton NJ: Princeton University Press.

Moss, David A. (2002) *When All Else Fails: Government As the Ultimate Risk Manager,* Harvard University Press, Cambridge, MA.

Nelson, Richard R. and Bhaven N. Sampat (2001) "Making Sense of Institutions as Factors Shaping Economic Performance," *Journal of Economic Behavior & Organization*, 44(2001):31-54

Nelson, Richard R. and Sidney G. Winter (1982) *An Evolutionary Theory of Economic Change,* Harvard University Press, Cambridge, MA.

Nevins, James L., Daniel E. Whitney, et.al. (1989) *Concurrent Design of Products and Processes,* McGraw-Hill, New York, NY.

Nickerson, Jack A. and Todd R. Zenger (2004) "A Knowledge-based Theory of the Firm—The Problem-Solving Perspective," *Organization Science*, 15(6): 617–632

North, Douglass C. (1990) *Institutions, Institutional Change and Economic Performance*, Cambridge University Press, Cambridge, UK.

Parnas, David L. (1972a) "A Technique for Software Module Specification with Examples," *Communications of the ACM* 15(May): 330-36.

Parnas, David L. (1972b) "On the Criteria to Be Used in Decomposing Systems into Modules," *Communications of the ACM* 15(December): 1053-58.

Parnas, David L., P.C. Clements, and D.M. Weiss (1985) "The Modular Structure of Complex Systems," *IEEE Transactions on Software Network*, SE-11(March): 259-66.

Pine, B. Joseph (1999) *Mass Customization: The New Frontier in Business Competition,* Harvard Business School Press, Boston, MA.

Posner, Richard (1998) *Economic Analysis of the Law 5th Ed.* Aspen Publishers, xxx.

Rivkin, Jan W. and Nicolaj Siggelkow (2003) "Balancing Search and Stability: Interdependencies Among Elements of Organizational Design" *Management Science* 49:290-311.

Sanchez, R. and Mahoney, J. T. (1996) "Modularity, flexibility and knowledge management in product and organizational design". *Strategic Management Journal*, 17: 63-76, reprinted in *Managing in the Modular Age: Architectures, Networks, and Organizations,* (G. Raghu, A. Kumaraswamy, and R. N. Langlois, eds.) Blackwell, Oxford/Malden, MA.

Schilling, Melissa A. (2000) "Toward a General Systems Theory and its Application to Interfirm Product Modularity," *Academy of Management Review* 25(2):312-334, reprinted in *Managing in the Modular Age: Architectures, Networks, and Organizations,* (G. Raghu, A. Kumaraswamy, and R. N. Langlois, eds.) Blackwell, Oxford/Malden, MA.

Shleifer, Andrei and Robert Vishny (1994) "Privatization in Russia: First Steps," in *The Transition in Eastern Europe, Volume 2: Restructuring* (O. J. Blanchard, K. R. Froot, and J. D. Sachs, eds.) University of Chicago Press, Chicago, IL.

Simon, Herbert A. (1999) *The Sciences of the Artificial*, *3rd ed.* MIT Press, Cambridge, MA.

Simon, Herbert A. and Mie Augier (2002) "Commentary on 'The Architecture of Complexity'," in *Managing in the Modular Age: Architectures, Networks, and Organizations,* (R. Garud, A. Kumaraswamy, and R. N. Langlois, eds.) Blackwell, Oxford/Malden, MA.

Smith, Adam (1994) *An Inquiry into the Nature and Causes of the Wealth of Nations*, (E. Cannan, ed.) New York, NY: Modern Library.

Sosa, Manuel, Steven Eppinger and Craig Rowles (2004) "The Misalignment of Product Architecture and Organizational Structure in Complex Product Development," *Management Science*, 50(December):1674-1689

Spear, Steven J. (1999) *The Toyota Production System: An Example of Managing Complex Social/Technical Systems*, unpublished DBA thesis, Harvard University, Graduate School of Business Administration, Boston, MA.

Spear, Steven J. (2002) "The Essence of Just-in-Time: Practice-in-Use at Toyota Production System Managed Organizations," Harvard Business School Working Paper 02-020, Boston, MA.

Spence, Michael A. (1973) "Job Market Signalling," *The Quarterly Journal of Economics,* 87(3):355-374.

Steward, Donald V. (1981) "The Design Structure System: A Method for Managing the Design of Complex Systems," *IEEE Transactions on Network Management* EM-28(3): 71-74 (August).

Stigler, George J. (1951) "The Division of Labor is Limited by the Extent of the Market," *Journal of Political Economy,* 54(3), reprinted in *The Organization of Industy,* University of Chicago Press, Chicago, IL.

Stiglitz, Joseph E. (1975) "The Theory of 'Screening', Education, and the Distribution of Income," *American Economic Review,* 65(3):283-300.

Sullivan, Kevin, William G. Griswold, Yuanfang Cai and Ben Hallen (2001). "The Structure and Value of Modularity in Software Design," SIGSOFT Software Engineering Notes, 26(5):99-108.

Turner, J. Scott (2000) *The Extended Organism: The Physiology of Animal-Built Structures*, Harvard University Press, Cambridge, MA.

Ulrich, Karl (1995) "The Role of Product Architecture in the Manufacturing Firm," *Research Policy,* 24:419-440, reprinted in *Managing in the Modular Age: Architectures, Networks, and Organizations,* (G. Raghu, A. Kumaraswamy, and R.N. Langlois, eds.) Blackwell, Oxford/Malden, MA.

Vickrey, William S. (1961) "Counterspeculation, Auctions, and Competitive Sealed Tenders," *Journal of Finance,* 16:8-37.

Wernerfelt, Birger (1997) "On the Nature and Scope of the Firm: An Adjustment-Cost Theory," *Journal of Business*, 70(4):489-514.

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

Williamson, Oliver E. (1999) "Human Action and Economic Organization," mimeo, University of California, Berkeley, CA, as quoted by Aoki (2001).

Woodard, C. Jason (2006) "Architectural Strategy and Design Evolution in Complex Engineered Systems," unpublished Ph.D thesis, Information, Technology & Management, Harvard University, Cambridge, MA.