University of Connecticut
DigitalCommons@UConn

Economics Working Papers | Department of Economics

11-1-2005

# The Secret Life of Mundane Transaction Costs

Richard N. Langlois
*University of Connecticut*

Follow this and additional works at: http://digitalcommons.uconn.edu/econ_wpapers

Recommended Citation
Langlois, Richard N., "The Secret Life of Mundane Transaction Costs" (2005). *Economics Working Papers.* Paper 200549.
http://digitalcommons.uconn.edu/econ_wpapers/200549

This is brought to you for free and open access by the Department of Economics at DigitalCommons@UConn. It has been accepted for inclusion in Economics Working Papers by an authorized administrator of DigitalCommons@UConn. For more information, please contact digitalcommons@uconn.edu.

University of Connecticut

*Department of Economics Working Paper Series*

# The Secret Life of Mundane Transaction Costs

Richard N. Langlois
University of Connecticut

Working Paper 2005-49

November 2005

341 Mansfield Road, Unit 1063
Storrs, CT 06269–1063
Phone: (860) 486–3022
Fax: (860) 486–4463
http://www.econ.uconn.edu/

This working paper is indexed on RePEc, http://repec.org/

**Abstract**

Transaction costs, one often hears, are "the economic equivalent of friction in physical systems." Like physicists, economists can sometimes neglect friction in formulating theories; but like engineers, they can never neglect friction in studying how the system actually does – let alone should – work. Interestingly, however, the present-day economics of organization also ignores friction. That is, almost single-mindedly, the literature analyzes transactions from the point of view of misaligned incentives and (especially) transaction-specific assets. The costs involved are certainly costs of running the economic system in some sense, but they are not obviously "frictions." Stories about frictions in trade are not nearly as intriguing as stories about guileful trading partners and expensive assets placed at risk. But I will argue that these seemingly dull categories of cost – what Baldwin and Clark (2003) call *mundane* transaction costs – actually have a secret life. They are at least as important as, and quite probably far more important than, the more glamorous costs of asset specificity in explaining the partition between firm and market. These costs also have a secret life in another sense: they have a secret *life cycle*. I will argue that these mundane transaction costs provide much better material for helping us understanding how the boundaries among firms, markets, and hybrid forms change over time.

**Journal of Economic Literature Classification:** D23, L22

**Keywords:** transaction costs, division of labor, modularity, standards, property rights.

To appear as a "Peripheral Vision" feature of *Organization Studies.* The author benefited from presenting early versions of this paper at the miniconference on "Linking Strategy, Technology and Organization: Strategic and Institutional Perspectives on Modularity," London Business School, October 2, 2003, and at the conference on "The Evolution of Designed Institutions," Max Planck Institute for Research into Economic Systems, February 19-21, 2004, Jena, Germany, as well as at seminars at the Copenhagen Business School and the Fisher School of Business at Ohio State. He would like to thank Carliss Baldwin, Giampaolo Garzarelli, Brian Kench, and Haridimos Tsoukas for helpful comments.

## *Anatomy of the mundane.*

In the preeminent book devoted (self-consciously) to transaction-cost economics, Oliver Williamson tells us this about what transaction costs are.

> Kenneth Arrow has defined transaction costs as the "costs of running the economic system" (1969, p. 48). Such costs are to be distinguished from production costs, which is the cost category with which neoclassical analysis has been preoccupied. Transaction costs are the economic equivalent of friction in physical systems (Williamson 1985, pp. 18-19).

Williamson then quite properly chastises traditional neoclassical theory for basing its policy conclusions on the assumption that all costs are production costs. Like physicists, economists can sometimes neglect friction in formulating theories; but like engineers, they can never neglect friction in studying how the system actually does – let alone should – work.

Interestingly, however, Williamson himself also neglects friction. That is, he immediately drops the characterization of transaction costs as frictions, and goes on to study transactions from a quite different perspective. Transactions, he says, have three "critical dimensions": uncertainty, frequency, and asset specificity (1979, p. 239). But the "most critical dimension for describing transactions is asset specificity" (1985, p. 30). The possibility of opportunism in the face of highly specific assets certainly does introduce a source of costs other than production costs as traditionally understood in price theory. Such

opportunism may even be a cost of running the economic system in some sense.[1] But it is not obviously a friction. The costs that opportunism introduce arise from incentive misalignment, and incentive misalignment is quite arguably the central preoccupation of the present-day economics of organization, at least as practiced within actual economics departments (Langlois and Foss 1999). But most theorists of incentive misalignment find no reason to refer to the resulting costs as transaction costs, let alone as frictions.

One reason for this odd detour is no doubt that, taken literally, friction isn't a particularly sexy source of transaction costs. As the term was originally used in monetary theory, a transaction cost occurred when, in effect, some amount of the good was used up in the transaction. In Paul Samuelson's (1954) famous iceberg model of transportation costs, a certain amount of the iceberg melts away as it is transported – or, we might add, as it waits around while being exchanged. This sounds very much like friction. Indeed, if we think of a transaction cost in this way, "in no significant way does it differ from a regular transportation cost" (Dahlman 1979, p. 144).

Many have written about the history and meaning of the term transaction cost.[2] One insightful treatment, that by Douglas Allen (2000), will serve as our

---

1 Already in 1979, Dahlman could write that the term transaction cost had "become a catch-all phrase for unspecified interferences with the price mechanism" (Dahlman 1979, p. 144).

2 For a history of the concept of transaction costs, see, for example, Klaes (2000), which, unfortunately, does not offer much critical analysis of the changing (or perhaps widening) meaning of the term.

jumping-off point. Allen argues that there are two distinct traditions using the term transaction cost, and accordingly two quite different meanings of the term. One is the *property rights tradition*, inspired by the work of Coase (1937, 1960), which (implicitly) defines transaction costs as "the costs of establishing and maintaining property rights" (Allen 2000, p. 898). These costs can be fixed (and sometimes sunk) as well as variable. They include costs not only of setting up but also of maintaining – and, importantly, of policing – the system of rights. They may be paid privately or through the government. They include the costs of locks, police, and guard dogs, as well as any deadweight loss arising out of the attempt to protect or expropriate rights (Allen 1991, p. 3). In a contractual setting, they can include monitoring costs, bonding costs (including the sunk costs of a hostage (Williamson 1985, chapter 7)), and any residual loss of having imperfectly protected one's assets.[3] This is clearly a broad set of costs, not all of which seem obviously entitled to the moniker "transaction costs." For his part, Allen is happy to identify the property rights approach with the entire field of the New Institutional Economics and with theorists ranging from Alchian and Demsetz to Klein and Williamson to Milgrom and Roberts.

The second tradition Allen identifies is the *neoclassical approach*, in which transaction costs are "the costs resulting from the transfer of property rights" (Allen 2000, p. 901). In this approach, which is the one that grew out of monetary

[3] I am here generalizing (or at least adapting) the three categories of what Jensen and Meckling (1976) call agency costs.

economics (Hicks 1935), transaction costs really are frictions: they arise only when a title to property is exchanged in a market (and only in a market, not in a firm or a centrally planned economy).[4]

An example – mine not Allen's – may help clarify. In the American Midwest before the coming of the railroad, as indeed throughout much of agricultural history, wheat was stored, shipped, and traded by the sack (Cronon 1991). Each sack of wheat was the product of a specific identifiable farmer, which meant that repeated trades could generate reputation effects that assured the quality of the grain in the market. At the same time, however, this mode of storage meant large transaction costs in the neoclassical sense: the transportation costs of shipping the sacks by wagon and river to St. Louis or Chicago; brokerage fees; insurance premia; the implicit costs of price volatility and poor information about market prices at destination; and the cost of the stevedores lugging bags from warehouse to barge to warehouse.[5] Even the burlap bags themselves cost two to four cents apiece (Cronon 1991, p. 113). Holding shipping route constant, most of these costs were arguably incurred on a per-sack basis as part of the

---

4 Lone among property rights theorists, Harold Demsetz (1988) cleaves to this definition of transaction costs, even though he is of course perfectly cognizant that there are costs of creating and defending property rights as well as costs of transferring resources in non-market settings. He would prefer to use the term *governance cost* to refer to the cost of transferring resources in either a market or a non-market setting.

5 According to one (clearly exaggerated) contemporary account, a "10,000-bushel shipment of grain arriving in St. Louis might involve 'the labor of probably two or three hundred Irishmen, negroes and mules for a couple of days'" (Cronon 1991, p. 112).

process of exchanging title to the wheat. It was as if some – perhaps most – of the grain spilled out of the sack between farmer and miller.

All of this changed with the coming of the railroad in the mid-nineteenth century. It quickly became economical to store and ship wheat in bulk, using the newly invented mechanical grain elevator. This reduced neoclassical transaction costs dramatically.[6] But, as it necessitated mixing together the grain of many different farmers, it destroyed the system of quality control that had relied on reputational effects from repeated transactions with identifiable farmers. To solve this problem, the Chicago Mercantile Exchange paid the costs of creating standardized categories for wheat and persuading farmers and buyers to adopt those standards (Cronon 1991). In addition, they needed to pay the costs of inspecting the wheat for conformance to the standards, which they did by commissioning inspectors.

The costs of this system are clearly the costs of establishing and maintaining property rights. But are they neoclassical "frictions"? The costs of establishing the standards clearly are not, since they are incurred once-and-for-all as a fixed cost and cease thereafter to enter into marginal calculations. But such fixed costs are of course related to frictional costs, in that they are *substitutes* for them: by paying the one-time cost of standards, I can avoid having to pay stevedores by the sack. The value of substituting such a fixed cost for a per-

transaction cost clearly increases with the volume of transaction. There is a more general point: costs of establishing and maintaining property rights are related to costs of exchanging property titles in that the principal (though perhaps not the only) reason for establishing rights to things is in order to trade those rights.

But what about the costs of policing the standards? That is less clear. If inspection occurred whenever a bushel changed hands – and thus if inspection were a per-transaction cost – then inspection would be both a cost of maintaining a system of property rights and a neoclassical cost of exchanging titles to property. In the event, this is very close to how the system actually worked in Chicago. What merchants traded in the meeting rooms of the Exchange were literally paper titles – elevator receipts for specific lots. By the 1860s, the Board had a city charter to appoint a "grain inspector of the city at large," who in turn hired and trained a team of assistants. For a standard fee, these inspectors would examine lots of grain and certify the grade of any elevator receipt traded on the floor (Cronon 1991, p. 119). On the other hand, if inspection were not priced by the lot exchanged – if, for example, inspectors simply strolled into elevators periodically to make sure that everything was up to snuff – then the policing costs would actually be fixed costs even though they represent ongoing outlays: they would depend on time, but not on output (number of exchanges or volume of trade), and would thus be fixed not variable in the Econ 101 sense. Much

---

[6] A typical large elevator of the era could simultaneously empty twelve railroad cars and load two ships at the rate of 24,000 bushels per hour (Cronon 1991, p. 113).

policing of property rights is presumably of this latter type. We pay police by the hour, not by the transaction or by the acre protected.

So it appears we really have three categories of cost. See Figure 1. Category (1) contains the fixed costs of establishing and maintaining a system of property rights. These include not only technological standards like those for wheat but also legal standards – including the whole of property law. It also

<table>
<tr><td>(1)</td><td>(2)</td><td>(3)</td></tr>
<tr><td>Fixed costs</td><td>Costs that are<br>a function of time</td><td>Costs that are a function of number of exchanges or volume of trade</td></tr>
<tr><td>Examples:<br><br>Legal, organizational, and technological standards; hostages and bonds; locks, closed-circuit TV.</td><td>Examples:<br><br>Salaries of police, supervisors, and other monitors; monthly protection money; maintenance of fixed investments.</td><td>Examples:<br><br>Brokerage fees, commissions; insurance premia; queuing at the bank, ATM fees; inspection and regulatory fees; per-transaction bribes.</td></tr>
<tr><td colspan="2">Costs of property rights.</td><td>Neoclassical T-costs.</td></tr>
<tr><td></td><td colspan="2">Mundane transaction costs</td></tr>
</table>

**Figure 1**: types of (transaction) costs.

includes what we might think of as "organizational" standards such as the corporate law of asset partitioning (Hansmann and Kraakman 2000). It would also include the fixed (and often sunk) costs arising from highly specific assets and the bonds and hostages sometimes used to offset those costs. And, of course, it includes ordinary fixed capital like locks. Category (2) contains costs that are paid periodically. These depend on time, but not on number or volume of transactions. This would include the salaries of police and of supervisors in firms. It would also include bribes, taxes, red tape, and protection money when the costs of those are incurred over time not per transaction. And it would include depreciation and maintenance of the fixed assets in category (1), even the "maintenance" of laws and standards. Category (3) contains all the costs that come with number of transactions or volume transacted. These include things like commissions, ATM fees, brokerage fees, insurance premia, inspection and regulatory fees, and sales taxes and transfer fees – including bribes when those are paid per transaction. They would also include technological costs of transacting, like queuing at the bank or waiting around in port to be unloaded.

Category (1) and (2) are all costs of establishing and maintaining property rights. Category (3) are the neoclassical costs of exchanging titles to property: the costs of melting icebergs and leaking burlap bags. In a sense, though, category (3) costs can also be costs of establishing and maintaining property rights: the price of setting up a system of ownership may often be that owners have to pay a per-transaction fee when they exchange title to what they own. More generally,

the costs one pays in categories (2) and (3) depend on – flow from – the costs one pays in category (1). As in the case of Midwestern wheat in the nineteenth century, the point of incurring costs of type (1) is to lower costs of type (2) and (3). This is also true, of course, in the case of production costs, where one is often happy to pay high fixed costs in order to reduce variable costs, especially when one expects to amortize those fixed costs over a large volume of output. Moreover, production-cost choices can also drive category (2) and (3 ) transaction costs. It was the coming of the railroad and the invention of the grain elevator that motivated the creation of standards and the system of inspection. One needs to think about both production costs and transaction costs, as well as about their interaction – a point often forgotten in both theory (Langlois and Foss 1999) and economic history (Langlois 2004).

In what follows I want to do two things. First, I want to think a bit more about the relationship among production costs, fixed costs of category (1), and transaction costs of categories (2) and (3). To do this, I will introduce something called the modularity theory of the firm. Second, I want to rethink the importance of the more blue-collar costs of categories (2) and (3). These two categories are what, more-or-less following Baldwin and Clark (2003), I will call *mundane* transaction costs.[7] Stories about frictions in trade are not nearly as

[7] Baldwin and Clark (2003) attribute the term "mundane transaction costs" to Williamson (1985). This phrasing is actually nowhere to be found in Williamson. But what they apparently had in mind (Carliss Baldwin, private communication) is Williamson's discussion of *mundane* versus *peripheral* activities (1985, pp. 105-106). By mundane activities

intriguing as stories about guileful trading partners and expensive assets placed at risk. But I will argue that these seemingly dull categories of costs actually have a secret life. They are at least as important as – and quite probably far more important than – the more glamorous costs of asset specificity in explaining the partition between firm and market. These costs also have a secret life in another sense: they have a secret *life cycle*. I will argue that these mundane transaction costs provide much better material for helping us understanding how the boundaries among firms, markets, and hybrid forms change over time.

### *The modularity theory of the firm.*

One important legacy of Coase, which Williamson embraces warmly, is the approach of *comparative institutional analysis*. Originally growing more out of "The Problem of Social Cost" (Coase 1960) than "The Nature of the Firm" (Coase 1937), comparative institutional analysis embodied a criticism of the Pigovian tradition of "market failure."[8] Under the doctrine of market failure, one is free to condemn real-world economic arrangements as inefficient, and free

---

Williamson means "core" activities that one would always "expect" to be vertically integrated: a classic example would be the successive stages of processing hot steel ingot. By contrast, he says, peripheral activities – what we would think of as forward or backward integration – are more interesting candidates for transaction-cost analysis, since they could potentially take place "offsite." Williamson thinks that the explanation for mundane integration is the same as that for peripheral integration – asset specificity and opportunism. The point of Baldwin and Clark is that one has first to explain why some activities are "mundane" and some "peripheral" in the first place As we will see, they look for that explanation in the underlying (modular) structure of the production process itself.

[8] This tradition is still alive and well, even if the abstract models used for comparison are different. A preeminent modern practitioner is the recent Nobel laureate Joseph Stiglitz (2002).

simultaneously to call for government intervention (often of unspecified form), based upon a comparison of the observed arrangements with a theoretical ideal that, all too often, failed to take account of all costs and constraints.[9] By contrast, comparative institutional analysis operates by comparing actual discrete institutional alternatives. As Williamson describes it, this involves "assigning" transactions to alternative organizational or institutional forms and arguing about which of the alternatives minimizes the sum of production and transaction costs[10] (Williamson 1985, p. 18).

Comparative-institutional analysis represents a great advance over Pigovian welfare economics. But our applause should not distract us from the ultimate limitations of the approach. Where exactly do the alternatives come from? More subtly: what exactly are these things called "transactions," and where do *they* come from? According to Williamson (1985, p.1), a "transaction occurs when a good or service is transferred across a technologically separable interface. One stage of activity terminates and another begins." But where do technologically separable interfaces come from? Why do activities terminate and begin where they do? Some recent literature, which we may lavishly call the *modularity theory of the firm* (Langlois 2002; Baldwin and Clark 2003), seeks to

---

9 See Demsetz (1969) for a scathing early exposition of this criticism.

10 Williamson actually uses the term *governance structure*, since, again, he is primarily concerned with how institutional alternatives deal with (govern) problems like opportunism.

answer these questions in a way consistent with the method of comparative-institutional analysis.

The approach from modularity tries to place the question of organizational and institutional choice within the roomier framework of a systems-design perspective. In the end, an institutional structure is a complex "designed" structure akin to those familiar in engineering design or biological evolution.[11] A system of social institutions reflects some elements of conscious human design; and it is certainly an evolutionary system. But, as Hayek long insisted, it is actually a third thing: it is, in Adam Ferguson's famous phrase, "the result of human action but not of human

| | $a_1$ | $a_2$ | $a_3$ | $a_4$ | $a_5$ | $a_6$ | $a_7$ |
|---|---|---|---|---|---|---|---|
| $a_1$ | x | x | x | x | x | x | x |
| $a_2$ | x | x | x | x | x | x | x |
| $a_3$ | x | x | x | x | x | x | x |
| $a_4$ | x | x | x | x | x | x | x |
| $a_5$ | x | x | x | x | x | x | x |
| $a_6$ | x | x | x | x | x | x | x |
| $a_7$ | x | x | x | x | x | x | x |

**2.1. A non-decomposable system.**

| | $a_1$ | $a_2$ | $a_3$ | $a_4$ | $a_5$ | $a_6$ | $a_7$ |
|---|---|---|---|---|---|---|---|
| $a_1$ | x | x | | | | | |
| $a_2$ | x | x | | | | | |
| $a_3$ | | | x | x | | | |
| $a_4$ | | | x | x | | | |
| $a_5$ | | | | | x | x | |
| $a_6$ | | | | | x | x | |
| $a_7$ | | | | | | | x |

**2.2 A nearly decomposable system**

| | $a_1$ | $a_2$ | $a_3$ | $a_4$ | $a_5$ | $a_6$ | $a_7$ |
|---|---|---|---|---|---|---|---|
| $a_1$ | x | x | x | x | x | x | x |
| $a_2$ | x | x | x | | | | |
| $a_3$ | x | x | x | | | | |
| $a_4$ | x | | | x | x | | |
| $a_5$ | x | | | x | x | | |
| $a_6$ | x | | | | | x | x |
| $a_7$ | x | | | | | x | x |

**2.3. A modular system with common interface.**

**Figure 2.**

[11] I will henceforth take institution to comprehend organization, though I return specifically to the question of firms. For my views on the relationship of organization and institution, see Langlois (1995).

design" (Hayek 1967), where the "not of human design" part is intended to mean not that humans never try to design institutions (or parts of them) but rather that the system as a whole does not spring entire from such efforts.

Talk of complex systems may call up a vision of Jeff Goldblum, as the supercilious mathematician in *Jurassic Park*, lecturing on the inevitability of nonlinear effects and unintended consequences whenever humans conjure with complexity. To the more sophisticated, it may bring to mind the Santa Fe Institute. In fact, however, unpredictable nonlinear effects are by no means an inevitable outcome of complexity. As writers like Herbert Simon (1962) and Christopher Alexander (1964) long ago argued, the designed systems we observe in the world are often precisely those that have found ways of evading the nonlinearities of complexity. They have done so by adopting a structure that is modular or, to put it more Simon's more precise term, *decomposable*. All systems are "modular" in the sense that they comprise parts, subsystems of parts, and relations among parts and subsystems. The real issues are (1) which parts are in which subsystems and (2) what are the relationships among the subsystems. A decomposable system is one that assigns parts to subsystems (modules) so as to minimize the total number of interactions among the modules. Complete decomposability is not an impossible ideal: all systems in which the modules have nothing whatever to do with one another are completely decomposable. But such systems are also of little interest, since the modules never work together. Of much greater interest are *nearly decomposable* systems. Figure 2

represents this in terms of a matrix of relationships. If module $a_i$ interacts with module $a_j$, then there is an X in the appropriate spot; otherwise not. If the Xs are mostly near the diagonal, as in Figure 1, then most interaction is contained within or "near" the modules themselves, and there are few far-flung interactions of the sort that so animated Goldblum. The butterfly flapping its wings in South America creates no typhoon in China.

There are two tricks to near decomposability as a design strategy. The first, for which there is no magic formula, is to figure out the best way to assign parts to modules. The second is ensure coordination among the (nearly) decomposed modules. In Figure 2.3, $a_1$ is a special module whose function is to coordinate the other modules. This module may itself be a firm – a "systems integrator" that communicates with all the other firms.[12] In other cases, we can think of $a_1$ as an abstract set of standards that govern the interactions among modules. Either way, the modules need to interact only (mostly) with $a_1$, not with each other. Baldwin and Clark (1997) generalize the idea of standards, as

---

[12] Stefano Brusoni and his coauthors (Brusoni 2005; Brusoni, Prencipe, and Pavitt 2001) like to portray the widespread need for systems integrators in distributed production networks as somehow in conflict with the theory of modular systems or perhaps even the Smithian theory of the division of labor. As Smith pointed out, however, the existence of systems integrators – agents "whose trade it is not to do any thing, but to observe every thing; and who, upon that account, are often capable of combining together the powers of the most distant and dissimilar objects" (Smith 1976, I.i.9) – is itself a manifestation of the division of labor. A firm specializes in being a systems integrator. It may well be true that, in their enthusiasm for abstract standards, proponents of the theory of modular systems have not always made clear that the coordinative knowledge possessed by the coordinating module $a_1$ can sometimes be contained within the capabilities of an organization as well as within an abstract institution. I argue that the theory of dynamic transaction costs (introduced below) offers a way to think about when coordinative knowledge will tend to be contained within

we loosely use the term in ordinary language, into a tripartite set of what they call *visible design rules*. These consist of the *architecture* – the overall assignment of parts to modules; the *interfaces*, which govern the connections among modules; and the *standards* proper, which are criteria to measure the compliance of the modules with the design rules. Consider again the case of wheat at the Chicago Mercantile Exchange. The publicly agreed-upon categories of wheat are the architecture. The criteria, enforced by inspectors, for determining the bin into which a particular farmer's crop will fall are the standards proper. What then is the interface? It is the price system.

> The whole acts as one market, not because any of its members survey the whole field, but because their limited individual fields of vision sufficiently overlap so that through many intermediaries the relevant information is communicated to all. … 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 and passed on only to those concerned. It is more than a metaphor to describe the price system as a kind of machinery for registering change, or a system of telecommunications which enables individual producers to watch merely the movement of a few pointers, as an engineer might watch the hands of a few dials, in order to adjust their activities to changes of which they may never know more than is reflected in the price movement. (Hayek 1945, pp. 526-527).

---

the capabilities of organizations and when delegated to unselfconscious mechanisms like standards.

Because of the wheat market's architecture and standards, the modules (individual farmers and wheat traders) need only (mostly) interact through prices. Relative prices are in effect module $a_1$.

The modules themselves can operate in any fashion they choose. They can have their own *hidden design parameters*. Unlike the visible design rules, which must be commonly shared, the module's own design parameters need not be communicated to others. Indeed, as the field of object-oriented programming has taught us, these hidden design parameters generally *must not* be communicated to others. This is the principle of *encapsulation* and *information hiding*. To allow another module knowledge of and access to one's inner workings is to invite those other modules to tinker – and thus to invite butterfly effects.

There is a natural fit between the theory of modular (that is, nearly decomposable) systems and the theory of property rights discussed earlier (Langlois 2002). Rights – what legal philosophers call *negative* rights – are all about exclusion. They are about encapsulation and information hiding. If rights are not defined properly, then what goes on in one module (one bundle of rights) can affect what goes on in another in ways that the interface (i. e., the legal system) doesn't take into account in coordinating the system's overall performance. We call such effects externalities. In the literature on law and economics, a well-designed system of property rights is one that minimizes the

losses from externalities (which is not, of course, necessarily the same thing as minimizing externalities). The same is true in the theory of the firm as seen from a property rights perspective: the boundaries between firm and market arise as a solution to the problem of minimizing the costs of externalities.[13]

Baldwin and Clark (2003) cast the matter in a congruent form that more explicitly accounts for production as well as transacting. The process of production, they argue, is a system of tasks assigned to various actors, both human and mechanical. The process requires the transfer of material, energy, and information among those actors. The design problem is to minimize total costs by deciding which transfers will take place within modules and which will cross module boundaries.[14] Unlike Williamson, but like the neoclassical transaction-cost tradition, Baldwin and Clark insist on calling transactions only those transfers that cross module boundaries – only those transfers that, in their terms, are *standardized, counted,* and *compensated*.

Transfers that cross boundaries incur frictional (and other mundane) transactions costs, since, among other things, it is costly to count, to value, and to

---

[13] Demsetz (1964) suggested early on that vertical integration is one way to solve paradigmatic Coasean problems of conflicting land-use. His example is the shopping mall, in which a single owner (of the mall) manages and internalizes the externalities that would otherwise be created by separate shops. From this it is a small step to Alchian and Demsetz (1972), which is really a story about how the rebundling and reassignment of property rights solves a free-rider-externality problem.

[14] Because they are writing form a product-design perspective, Baldwin and Clark tend to cast the design problem as one faced by a human designer or design team. But the same logic applies if the system actually emerges from unselfconscious design processes in a manner akin to biological evolution.

collect compensation. Typically, there are also (mostly fixed) costs of turning transfers into transactions by establishing standards and procedures for counting, valuing, and collecting compensation. When it is costly, for whatever reason, for a transfer to become a transaction, the transfer is best left to take place within a module – within the boundaries of an organization – where it need not be standardized, counted, and compensated. The overall design of the task-and-transfer system, then, will typically be one in which the most-transaction-costly transfers take place within the boundaries of modules and only the least-transaction-costly transfers actually become transactions (and only these actually incur transaction costs). Keep in mind that, since the goal is to minimize the sum of production costs and all three kinds of transaction costs, the result may not be one that minimizes transaction costs, since one may well want to increase transaction costs (of one kind or another) if that will more than proportionally reduce production costs or other kinds of transaction costs. As I have already hinted, however, there is arguably a general tendency in this design process for costs of category (1) – fixed costs of standards and the like – to substitute for mundane transaction costs (especially those of category (3)) as the extent of the market expands.

### *The secret life.*

Before there was Williamson, of course, there was Coase. As we saw, Williamson talks about transaction costs as frictions but then immediately treats

them not as frictions but as costs of maintaining property rights (costs of governing transactions). By contrast, Coase is (rightly) credited with having inspired the property-rights approach in which transaction costs are the costs of establishing and maintaining property rights. Interestingly, however, Coase *himself* viewed transaction costs as neoclassical frictions.

Coase didn't originally use the term transaction costs, and he has never seemed comfortable with the word since (Klaes 2000, p. 569). He talked instead of "the costs of using the price mechanism."[15] These seem to fall into the categories of either search costs or contracting costs. "The most obvious cost of 'organising' production through the price mechanism," he says, "is that of discovering what the relevant prices are." These are search costs, very much in the vein later mined by George Stigler (1961). Contracting costs are of two types, neither of which involves the kind of incentive misalignment that plagues contracting in the present-day economics of organization. One kind of contracting cost seems to qualify as frictional, or at least mundane: the tedious business of "concluding a separate contract for each exchange transaction which takes place on a market." Organizing within a firm, Coase thinks, is a way to reduce the friction: "A factor of production (or the owner thereof) does not have to make a series of contracts with the factors with whom he is co-operating within the firm, as would be necessary, of course, if this co-operation were as a

[15] All quotations in this paragraph are from (Coase 1937, pp. 390-391).

direct result of the working of the price mechanism. For this series of contracts is substituted one."

Ultimately, however, it is a different type of contracting cost that really attracts Coase's attention. Notice that the costs of searching and writing many contracts vanish if there is no change or uncertainty: I can search once then simply write a long-term contract of indefinite length with the contracting parties I find. But in a world of uncertainty, Coase observes, such a contract would be too inflexible to deal with change and contingency.

> Now, owing to the difficulty of forecasting, the longer the period of the contract is for the supply of the commodity or service, the less possible, and indeed, the less desirable it is for the person purchasing to specify what the other contracting party is expected to do. It may well be a matter of indifference to the person supplying the service or commodity which of several courses of action is taken, but not to the purchaser of that service or commodity. But the purchaser will not know which of these several courses he will want the supplier to take. Therefore, the service which is being provided is expressed in general terms, the exact details being left until a later date. All that is stated in the contract is the limits to what the persons supplying the commodity or service is expected to do. The details of what the supplier is expected to do is not stated in the contract but is decided later by the purchaser. When the direction of resources (within the limits of the contract) becomes dependent on the buyer in this way, that relationship which I term a "firm " may be obtained. (Coase 1937, pp. 391-392).

The firm does indeed substitute a single contract for many individual contracts, but that substitution is qualitative as well as quantitative. The single contract is not an output contract but an employment contract of the sort later spelled out by Herbert Simon: the employee agrees to let the employer choose any action **x**

from a set **X** of possible tasks in return for a specified wage[16] (Simon 1951). Rather than paying by the piece for each document filed or phone call answered, an executive pays an hourly wage and then assigns the secretary tasks on a flexible basis. The tasks thus take place within a "transaction-free zone" – the firm – and are not individually counted and compensated. "Indeed, firms can be considered as social artifacts *designed for the purpose* of encapsulating complex transfers of material, energy and information"[17] (Baldwin and Clark 2003, p. 25, emphasis original).

In effect, task uncertainty creates a coordination cost that limits the extent of the market and thus the division of labor: for if "production runs" of typing and phone answering were long and predictable enough, one could easily pay for these services by the piece, thus making them transactions rather than transfers. Perhaps we can generalize this idea. In Adam Smith's formulation, an increasing extent of the market permits a finer division of labor in part because, with larger expected output, it becomes worthwhile to pay the fixed costs of standardizing tasks and standardizing what gets transferred between tasks (Smith 1976; Leijonhufvud 1986). Division of labor does not by itself tell us when transfers will become transactions, since it is clearly possible to subdivide tasks

---

16 As Coase's disciple Steven Cheung put it, the emergence of the firm involves "the replacement of a product market by a factor market, resulting in a saving in transaction costs" (Cheung 1983, p. 3).

17 Baldwin and Clark note that other forms like families, tribes, and clans are also social artifacts for encapsulating transfers, but these were never consciously designed for the purpose. Perhaps we might add clubs as designed institutions of encapsulation.

even with in a transaction-free zone. But Smith's theorem – that the division of labor is limited by the extent of the market – does give us some clues about when transfers will be encapsulated and when they will be turned into transactions.

Notice the word "complex" in the sentence from Baldwin and Clark above. Firms encapsulate "complex" transfers, whereas less-complex transfers are more likely to become transactions. But the causality can work both ways. Transfers may be encapsulated because they are necessarily complex, involving, for example, transfers of rich and idiosyncratic information among actors. Alternatively, encapsulated transfers can be complex because they happen to be encapsulated and it simply doesn't pay to simplify and standardize them. Consider the following three possibilities.

1. *Transfers are encapsulated because transportation costs and frictional transaction costs limit the extent of the market.* In 1776, meat, bread, and ale, as well as the tasks needed to produce them, were relatively standardized; nonetheless, transfers were encapsulated within the household of an isolated Highlands farmer because the extent of the market for butchery, baking, or brewing was too small to make it worthwhile for the farmer to specialize in any one of these tasks and transact with other people through the market.[18] In

[18] "In the lone houses and very small villages which are scattered about in so desert a country as the Highlands of Scotland, every farmer must be butcher, baker and brewer for his own family" (Smith 1976, I.iii.2).

Glasgow and Edinburgh, people did specialize in these "peculiar trades"; but transportation costs prevented markets from extending into the Highlands.

2. *Transfers are encapsulated because coordination costs limit the extent of the market.* The extent of the market may be large, and tasks and transfers relatively standardized, but spatial and temporal uncertainty limit the *effective* extent of the market. In the parable of the secretary, the tasks of typing, filing, and phone answering are relatively standardized and separable; yet the unpredictability of which task will be needed at which moment limits the effective extent of the market, even when tasks in the aggregate are heavily demanded.[19] So the secretary is "encapsulated" as an employee whose transfers are not individually counted and compensated.

3. *Transfers are encapsulated because dynamic transaction costs limit the extent of the market.* The extent of the market may be (potentially) large, but something inherent in the nature of the tasks and transfers makes it prohibitive to standardize them and to create transactions.[20] Here the ultimate culprit is economic change and

---

[19] Becker and Murphy (1992) make a similar point that coordination costs limit the division of labor. But they choose to see coordination costs as something separate from the extent of the market, whereas I choose to emphasize that coordination costs limit the division of labor precisely because they effectively reduce the extent of the market.

[20] Brusoni (2005) writes that "there are cognitive limits to the extent of division of labour: what kinds of problems firms solve, and how they solve them, set limits to the extent of division

> technological innovation, especially when it is systemic in character. During the era in which Henry Ford was developing the moving assembly line, his company transformed from an assembler to a highly vertically integrated organization (Langlois and Robertson 1989). Even though the firm would have liked to have made use of outside suppliers, Ford and his team were systemically reinventing the way automobile parts were fabricated. The process of innovation made it necessary to transfer rich information quickly among actors and indeed made it undesirable to standardize the task-and-transfer system until the innovation reached maturity. As a result, it was costly for Ford to enlist outsiders through the interface of the market. These costs of informing outsiders and persuading them to cooperate in production are what I call *dynamic transaction costs* (Langlois 1992).

Now notice something interesting. In all three cases, the costs that motivate encapsulating tasks and transfers within the boundaries of the firm are not only all mundane but arguably all frictional. In the first case, it is literally

---

of labour, irrespective of the extent of the market." Again, however, I prefer to say that these cognitive limits – which are part of what lies behind the phenomenon of dynamic transaction costs – reduce the division of labor because they reduce the effective extent of the market below what it would be if cognitive ability were unbounded and free. Notice that Brusoni's argument is essentially the same as that of Becker and Murphy (1992); the principal difference is that the "coordination costs" of Becker and Murphy sound like mundane transaction costs of type 2, whereas those Brusoni discusses seem to be dynamic transaction costs.

transportation costs that matter. In the second case, coordination costs are clearly frictional: each time I wanted to transact with an office worker through the market, I would have to pay a "coordination fee" involving not only the resource costs of counting, pricing, and paying for each transaction but also the opportunity cost of slower and less effective matching of task with moment in time. Similarly in the third case: the costs of informing and persuading potential trading partners would involve a coordination fee made up not only of resource outlays but also of the opportunity costs of not having the capabilities you need when you need them (Langlois 1992, p. 113).

### *The secret life cycle.*

As we saw, when viewed from the perspective of design, the problem of explaining the boundaries of the firm is one of determining when transfers tend to remain unstandardized, uncounted, and uncompensated and when they tend to become transactions.[21] When viewed from an evolutionary perspective,

---

[21] In general, transfers tend to become transactions at "gateways" or "thin crossing points" between modules. Baldwin and Clark (2003, pp. 17ff.) use the marvelous example of a smithy and a kitchen in pre-modern times. Both units are engaged in complementary tasks that ultimately put food on the table: the smithy makes cooking implements and the kitchen prepares the food; and both employ the division of labor. At what point in this task-and-transfer system would one most expect to see a transaction? The natural place, of course, is at the handoff of finished implements between the smithy and the kitchen. This is so not merely because the transfer is relatively standardized and easy enough to count and compensate at that point but also because the smithy and the kitchen form two naturally distinct subsystems of related tasks. Compare this with George Richardson's (1972) suggestion that firms tend to take on board mostly activities that are related to one another in the sense that carrying out those activities requires similar knowledge or capabilities. In Baldwin and Clark, it is the interrelatedness of the tasks rather than similar capabilities that determines the natural boundaries of modules. But, as most students of capabilities would insist (see, notably, Nelson and Winter (1982), chapters 4 and 5), the knowledge that constitutes capabilities is often embedded in the routines and practices of the organization –

however, the problem of explaining the boundaries of the firm is also one of explaining how encapsulation boundaries and the location of transactions *change* over time. One way to think about this latter problem, I suggest, is to examine the ways in which changes in the various types of mundane transaction costs affect encapsulation boundaries.

Let's start with transportation cost and the kind of leaky-grain-sack transaction costs that are analogous to transportation costs. As these kinds of costs diminish, the extent of the market increases as local markets begin to integrate. With a larger extent of the market, and in the absence of any other kinds of transaction costs, it starts to pay not only to subdivide tasks further but also to turn more transfers into transactions.[22] Just as it is worth paying the fixed set-up costs of subdividing tasks (Smith), it is also worth paying the fixed set-up costs of standardizing transfers and creating a system to count and compensate them (Baldwin and Clark). Specialization in tasks and vertical disintegration in organization go hand in hand. The result would seem to be George Stigler's famous conclusion that "vertical disintegration is the typical development in growing industries, vertical integration in declining industries" (Stigler 1951, p. 189). I have argued, however, that, phrased this way, Stigler's conclusion isn't quite right (Langlois 1992). Stigler's conclusion is valid if we are comparing two

---

that is, capabilities are actually imbedded in the tasks themselves. In this sense, Richardson and Baldwin and Clark are making much the same point.

[22] Actually, as I'll suggest in a minute, an increase in the extent of the market can sometimes lead to the *integration* of tasks rather than to further subdivision.

states of long-run equilibrium, one with a small extent of the market and one with a larger extent of the market: in the equilibrium with the larger extent of the market, we should observe finer vertical disintegration than in the equilibrium with a smaller extent of the market. But this is not the same thing as saying that we should observe vertical disintegration in a "growing industry." Growth takes place in disequilibrium.[23] And disequilibrium can lead to dynamic transaction costs, especially when the increase in the extent of the market calls for a systemic reorganization of the task-and-transfer system (Langlois 1992). These dynamic transaction costs can lead to an even greater degree of vertical integration than existed before the extent of the market increased.[24] To put it another way, reductions in mundane transaction costs of type 1 can sometimes lead to offsetting increases in mundane transaction costs of type 3.

In the case of wheat in the American West of the mid-nineteenth century, the rapid decline in transportation costs attendant on the coming of the railroad led to systemic technological change and a major alteration in the task-and-transfer system. But it did not result in a major increase in the degree of vertical integration in the system, largely because the sectors upstream and downstream

---

[23] There is of course a notion of "equilibrium growth." But what I mean by equilibrium is that the rate of change of the degree vertical integration is either zero or close to and asymptotically approaching zero. Thus a change in the extent of the market will create a disequilibrium – a positive rate of change of the degree of vertical integration. I mean by the short run the period in real time during which that rate remains positive and by the long run the period after which it is zero or close to and asymptotically approaching zero.

[24] Phrased in terms of a *Gedanken* regression, we might say that in explaining the degree of vertical integration, the sign of the extent of the market is negative, but the sign on the rate of change of the extent of the market can be positive.

from storage and transportation – farming and local distribution – were tied to a dispersed geography and because market-supporting institutions (the standards) emerged quickly. But, as Alfred Chandler showed in *The Visible Hand* (1977), vertical integration was the rule in many other contemporary industries affected by the advent of the railroad, the telegraph, and improved inland water transportation. From petroleum to meat packing to branded goods, the effect of lower transportation and frictional transaction costs was to raise the scale of production, since it began to pay to mass produce and ship rather than to produce locally at small scale. Because these changes were systemic – involving major redefinitions of tasks and shifts of task boundaries – the necessary markets and market-supporting institutions could not be called into existence quickly, making it costly to organize many of the new transfers as transactions. Dynamic transaction costs were high, and entrepreneurs needed to encapsulate much of the system within the boundaries of large vertically integrated concerns (Langlois 2003b).

Thus the apparently anti-Smithian conclusion embodied in the title of Chandler's book: although the extent of the market was expanding, even dramatically, large swaths of the division of labor were coordinated as encapsulated transfers rather than as the market transactions Smith seemed to imply. Even when there were existing or potential "thin crossing points" between subsystems of tasks, these were regulated as uncounted and uncompensated transfers. Seen from the right perspective, however, there is

actually nothing un-Smithian going on (Langlois 2003b). A reduction in transaction costs of category 1 led to an increase in the *potential* extent of the market and a finer division of labor; but the *effective* extent of the market was small because of the systemic newness of the task-and-transfer network.

As we saw, by allowing the manager authority to adapt flexibly, the Coasean firm is one way to reduce the costs of coordination. Indeed, Chandler saw the problem in very much this way. Because the increased potential extent of the market called for mass production, and mass production called for high-fixed-cost machinery (more on which presently), it was crucial to keep the system humming in order to amortize those fixed costs over as high a volume as possible. And this required managers to adapt flexibly in order to *buffer* uncertainty and keep the effective extent of the market high (Langlois 2003a, 2003b). But flexible managers are not the only way to buffer uncertainty. Alchian and Demsetz long ago challenged the idea that authority is the fundamental difference between the contracts that constitute a firm and those that constitute a market: "Telling an employee to type this letter rather than to file that document," they famously jibed, "is like my telling a grocer to sell me this brand of tuna rather than that brand of bread" (Alchian and Demsetz 1972, p. 777). And they were right: markets *can* in fact provide flexible responses to uncertainty, though they do so in a way different from a firm. Markets and firms can be alternative ways of buffering uncertainty, alternative ways of adapting flexibly to variation in the environment (Langlois 2003b). In the case of the firm,

the manager adapts to change by using authority to direct resources flexibly as states of the world reveal themselves in real time. In the case of the Alchian-and-Demsetz grocer, the customer can also act flexibly because the wide variety of alternative commodities available in the market permits him or her to exercise authority to adapt in real time. I can choose tuna over bread as needed, just as I can choose typing over filing as needed, because multiple alternatives are available to me and I have the authority to choose on the spot without negotiation.

As time passes and change slows, markets will become thicker, and both coordination costs and dynamic transaction costs will tend to diminish, all other things equal. As this happens, it will begin to be worth paying the fixed costs of transaction-supporting institutions. More and more transfers will become transactions. And thick markets will increasingly come to supplant management as a mechanism for buffering uncertainty. Arguably this is what has happened since the era of Chandler's large vertically integrated firms. I call the phenomenon *the vanishing hand*, and I have tried to explain it in a nuanced and historically attentive way (Langlois 2003b). But the underlying theory is exactly this. In the pre-Chandlerian era, transportation and frictional transaction costs were high, and lack of scale limited the extent to which transfers could be marketized. In the Chandlerian era, transportation and frictional transaction costs fell dramatically, leading to a systemic reorganization of the task-and-transfer system – but also to high dynamic transaction costs that made it

uneconomical initially to turn transfers into transactions. In the post-Chandlerian era, dynamic transaction costs diminished as markets thickened and market-supporting institutions emerged; this led to vertical disintegration as it began to pay to turn transfers into transactions.

So Smith (and Stigler) are right. In the long run, after dynamic transaction costs have had a chance to die down and to the extent that no other disturbing changes have emerged, increasing extent of the market leads to finer subdivision of tasks *and* to the transformation of internal transfers into transactions – to the division of labor *and* to vertical disintegration. This is so because with a large and predictable extent of the market it becomes economical to pay the fixed set-up costs of task specialization and the fixed set-up costs of standardizing transfers. The two are in fact interrelated, as standardizing tasks and transfer by making operations increasingly routine is essential to the division of labor (Ames and Rosenberg 1965). Making tasks increasingly standard and routine reduces production costs, as Smith noted, because it simplifies those tasks and speeds learning by human agents. At the same time, however, standardized tasks both demand and facilitate standardized transfers between stations (Leijonhufvud 1986).

Let me now throw what may seem to be a monkey-wrench into the machinery. And that monkey-wrench in the machinery is, well, *machinery*. I have so far implied, with Smith and Stigler, that in the long run increasing extent

of the market always leads to the increasing transformation of encapsulated transfers into transactions. And I continue to hold that, in the large, this is correct. But there is in fact a tendency that would seem to work in the opposite direction. As, with an increasing predictable extent of the market, tasks become increasingly routine and transfers increasingly standardized, it sometimes begins to pay not to subdivide tasks further but rather to *integrate* tasks within a machine (Ames and Rosenberg 1965). Smith implicitly assumed that tools were already specialized to task and that the only response to increasing extent of the market is for labor to specialize to the same degree as tools. Further increases in the extent of the market would continue this process, since it would reveal the possibilities for even more specialized tools to which humans would in turn specialize themselves. But, as Ames and Rosenberg (1965) observe, *tools* can also change their level of specialization. As tasks become more and more routine and standardized, it increasingly pays not only to assign them to machine – which have comparative advantage over humans in routine tasks (Langlois 2003a) – but also to integrate several tasks into a single machine.[25]

---

[25] Smith (1976, I.1.3) famously describes the manufacture of pins in his day, in which ten men, organized according to principles of the division of labor, could make about 48,000 pins a day, or almost 5,000 per person per day. By Marx's era, making pins was already the business of machines, and a single machine could crank out 145,000 a day. But this increase in productivity came not because of a finer division of labor but because the pin-making machinery integrated many of the previously subdivided human tasks and undertook them with far greater speed and precision than humans could have mustered. In Marx's time, one woman or girl could supervise four machines, which means almost 600,000 per person per day (Marx 1961, Volume 1, Part IV, Chapter XV, Section 8, p 460). As of 1980, one person could supervise 24 machines, each making 500 pins a minute, or about 6 million pins per person per day (Patten 1980).

As we saw in the case of grain elevators, the supersession of subdivided tasks by machines often involves a dramatic redesign of the task-and-transfer network. A mechanical grain elevator is not just a faster and more efficient system of stevedores and conventional warehouses; it is something fundamentally different. Similarly, electromechanical (and later digital) telephone switches did not speed up telephone calls by subdividing the labor of human operators more finely or speeding up the way calls are connected on a switchboard; rather, they operated by internalizing transfers within a machine, which required users to master a new interface (a rotary dial and system of phone numbers instead of a nice chat with Mabel). This also changed the relationship of the worker to the machine, as the worker became a maintainer, designer, and (later) programmer rather than an "operator."[26] To put it another way, mechanization increases labor productivity not by reducing the number of workers performing the existing set of tasks but rather by crowding machines into the more routine tasks and crowding humans into tasks for which human cognition has comparative advantage (Langlois 2003a). In many cases, the role of the human worker is that of Chandler's manager in the small, namely, to buffer uncertainty: to keep the machine humming by tending it, feeding it, and maintaining it.

---

[26] As in Smith, the worker has indeed specialized to the machine. But the machine is less specialized than the Smithian tools it replaced, since it takes on several previously distinct operations (Ames and Rosenberg 1965).

That task integration into machines is also a consequence of the growth of the extent of the market does not so much contradict Smith as show how his account fits within the roomier accommodations of the modularity theory outlined above. The boundaries of modules are set by the "natural" grouping of tasks, and what is "natural" is driven by the effective extent of the market, that is, the size of production runs conditioned not only on potential demand but also on uncertainty and environmental variation. When the effective extent of the market is small, humans are unspecialized but tools are (relatively) more specialized.[27] As effective market size expands, tools become more specialized, but at the same time humans become increasingly specialized to tools. This process creates (in equilibrium) an increasing number of "technologically separable interfaces" along which transactions can flow. As the effective market expands still further, this process is supplemented by the possibility of integrating certain of the most standardized and predictable tasks into machines. This latter process has the effect of turning transactions back into transfers (even in equilibrium).

Does this mean that, as the effective extent of the market continues to grow, all tasks will be integrated into machines? The key word here is "effective." As Schumpeter (1950, p. 82) reminds us, capitalism "is by nature a form or method of economic change and not only never is but never can be

---

[27] Tools are typically more specialized than humans at low levels of market extent simply because it is relative cheap to make specialized tools. Even the most primitive hunter-

stationary." We can expect buffets of change to move the system (and especially parts of the system) back and forth, and certainly to keep us out of the realm of science fiction. Moreover, even as machines become more capable, it will still be economical to assign to humans those tasks in which humans have cognitive comparative advantage (Langlois 2003a). And, since humans are in a sense inherently technologically separable from one another, we should expect to see lots of vertical disintegration – perhaps especially in those areas of most interest to us – even in a science-fiction world in which machines take over most routine activities.

Furthermore, there is one sense in which mechanization actually increases the possibility of transactions across technologically separable interfaces. The integration of tasks into machines depends not only on the potential extent of output but also on the predictability of that output. "Predictability" here need mean only statistical predictability, not certainty. Thus machines can sometimes take over operations even in a world of type 2 mundane transaction costs. This has been true at least since the Jacquard loom, but the process has of course accelerated since the advent of cheap digital computing. When you feed your card into an ATM or surf to the Dell website, a machine not a human secretary confronts you with a menu of standardized choices over which you (the Coasean "buyer") can exercise authority flexibly. Here the combination of a standardized

---

gatherer societies possessed an amazing array of specialized implements.

architecture of choice and speedy, precise computers means that individual transactions can be counted and compensated. Of course, even digital computers are of little use in coping with dynamic transaction costs, which involve the inherently non-standard.

### *The sacred and the profane.*

The modularity theory of the firm focuses on the way in which tasks and knowledge are grouped together in the production system. By contrast, the theory of the firm in mainstream economics of organization – in which for present purposes I would include even Williamson – takes task boundaries (technologically separable interfaces) for granted and focuses, often single-mindedly, on the issue of asset ownership. In many of these theories, indeed, it is the ownership of non-human assets that defines the firm and sets its boundaries (Hart 1989). I am far from denying that issues of highly specific assets are never part of the explanation. For example, the dangers of putting specialized assets at risk may have been in the minds of some suppliers whom Chandlerian entrepreneurs found costly to persuade, and such considerations may thus have formed part of the dynamic transaction costs involved.[28] But, as I would tell the story of industrial dynamics over the last century and a half,

[28] History suggests, however, that, more often than not, it was not threats of opportunism that led to internalization but rather the opposite: suppliers – when these could even be found – more often refused to deal because they didn't have the necessary capabilities or because they actually considered the entrepreneur's scheme absurd (Silver 1984; Langlois 1992, 2004). (These latter may be two ways of saying the same thing.)

opportunism arising out of asset ownership plays at best a bit part (Langlois 2004).

The perspective from modularity I advocate here has the benefit of returning our focus to the *process* of organizational change. It also helps to connect the economics of organization to the grand theoretical landscapes of Smith and Coase. The transaction costs involved may be mundane. But it is well to remember that the older meaning of that term – still the first meaning in most dictionaries – is not "ordinary and commonplace" but "worldly." After all, like Smith and Coase, we are worldly philosophers.

***References.***


Alchian, Armen, and Harold Demsetz. 1972. "Production, Information Costs, and Economic Organization," *American Economic Review* **62**(5): 772-795.

Alexander, Christopher. 1964. *Notes on the Synthesis of Form*. Cambridge: Harvard University Press.

Allen, Douglas W. 1991. "What are Transaction Costs?" *Research in Law and Economics* **14**: 1-18.

Allen, Douglas W. 2000. "Transaction Costs," in Boudewijn Bouckaert and Gerrit De Geest, eds., *The Encyclopedia of Law and Economics*, volume 1. Chelthenham, Edward Elgar, pp. 893-926.

Ames, Edward, and Nathan Rosenberg. 1965. "The Progressive Division and Specialization of Industries," *Journal of Development Studies* **1**(4): 363-383 (July).

Arrow, Kenneth J. 1969. "The Organization of Economic Activity: Issues Pertinent to the Choice of Market versus Non-market Allocation," in *The Analysis and Evolution of Public Expenditure: The PPB system*. Vol. 1 U.S. Joint Economic Committee, 91st Congress, 1st Session. Washington D.C.: U.S. Government Printing Office, pp. 59-73.

Baldwin, Carliss Y., and Kim B. Clark. 1997. "Managing in an Age of Modularity," *Harvard Business Review* **75**(5): 84-93 (September-October).

Baldwin, Carliss Y., and Kim B. Clark. 2003. "Where Do Transactions Come from?" Working Paper, Harvard Business School, February.

Becker, Gary, and Kevin Murphy. 1992, "The Division of Labor, Coordination Costs, and Knowledge,". *Quarterly Journal of Economics* **107**: 1137–1160.

Brusoni, Stefano. 2005. "The Limits to Specialization: Problem Solving and Coordination in 'Modular Networks,'" *Organization Studies* **26**(12): 1885-1907.

Brusoni, Stefano, Andrea Prencipe, and Keith Pavitt. 2001. "Knowledge Specialisation, Organisational Coupling, and the Boundaries of the Firm: Why Do Firms Know More Than They Make?" *Administrative Science Quarterly* **46**: 597-621.

Chandler, Alfred D., Jr. 1977. *The Visible Hand: the Managerial Revolution in American Business*. Cambridge: The Belknap Press.

Cheung, Steven N. S. 1983. "The Contractual Nature of the Firm," *Journal of Law and Economics* **26**: 122 (April).

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

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

Cronon, William. 1991. *Nature's Metropolis: Chicago and the Great West*. New York: W. W. Norton.

Dahlman, Carl. 1979. "The Problem of Externality," *Journal of Law and Economics* **22**: 141-162.

Demsetz, Harold. 1964. "The Exchange and Enforcement of Property Rights," *Journal of Law and Economics* **7**: 11-26 (October).

Demsetz, Harold. 1969. "Information and Efficiency: Another Viewpoint," *Journal of Law and Economics* **12**: 1-22.

Demsetz, Harold. 1988. "The Theory of the Firm Revisited," *Journal of Law, Economics, and Organization* **4**(1): 141-161 (Spring).

Hansmann, Henry, and Reinier Kraakman. 2000. "The Essential Role of Organizational Law," *Yale Law Journal* **110:** 387.

Hart, Oliver D. 1989. "An Economist's Perspective on the Theory of the Firm," *Columbia Law Review* **89**(7): 1757-1774.

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

Hayek, F. A. 1967. *Studies in Philosophy, Politics, and Economics*. Chicago: University of Chicago Press.

Hicks, John R. 1935. "A Suggestion for Simplifying the Theory of Money," *Economica* **2:** 1-19.

Jensen, Michael C., and William H. Meckling. 1976. "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure," *Journal of Financial Economics* **3**: 305-360.

Klaes, Matthias. 2000. "The Birth of the Concept of Transaction Costs: Issues and Controversies," *Industrial and Corporate Change* **9**(4): 567-593.

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

Langlois, Richard N. 1995. "Do Firms Plan?" *Constitutional Political Economy* **6**(3): 247-261.

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

Langlois, Richard N. 2003a "Cognitive Comparative Advantage and the Organization of Work: Lessons from Herbert Simon's Vision of the Future," *Journal of Economic Psychology* **24**: 187-207.

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

Langlois, Richard N. 2004. "Chandler in a Larger Frame: Markets, Transaction Costs, and Organizational Form in History," *Enterprise And Society* **5**(3): 355-375 (September).

Langlois, Richard N., and Nicolai J. Foss. 1999. "Capabilities and Governance: The Rebirth of Production in the Theory of Economic Organization," *Kyklos* **52**: 201-18.

Langlois, Richard N., and Paul L. Robertson. 1989. "Explaining Vertical Integration: Lessons from the American Automobile Industry," *Journal of Economic History* **49**(2): 361-375.

Leijonhufvud, Axel. 1986. "Capitalism and the Factory System," in R. N. Langlois, ed., *Economic as a Process: Essays in the New Institutional Economics*. New York: Cambridge University Press, pp. 203-223.

Marx, Karl. 1961. *Capital*. Moscow: Foreign Languages Publishing House, Volume 1.

Nelson, Richard R., and Sidney G. Winter. 1982. *An Evolutionary Theory of Economic Change*. Cambridge: the Belknap Press.

Pratten, Clifford F. 1980. "The Manufacture of Pins," *Journal of Economic Literature* **18**(1): 93-96.

Richardson, G. B. 1972. "The Organisation of Industry," *Economic Journal* **82**(327): 883-896.

Samuelson, Paul A. 1954. "The Transfer Problem and Transport Costs, II: Analysis of Effects of Trade Impediments," *Economic Journal* **64**(254): 264-289 (June).

Schumpeter, Joseph A. 1950. *Capitalism, Socialism, and Democracy*. New York: Harper and Brothers, second edition

Silver, Morris. 1984. *Enterprise and the Scope of the Firm*. London: Martin Robertson.

Simon, Herbert A. 1951. "A Formal Theory of the Employment Relationship," *Econometrica* **19**(3): 293-305.

Simon, Herbert A. 1962. "The Architecture of Complexity," *Proceedings of the American Philosophical Society* **106**: 467-482, repinted in *idem*, *The Sciences of the Artificial*, 2nd ed. Cambridge: MIT Press, 1981.

Smith, Adam. 1976. *An Enquiry into the Nature and Causes of the Wealth of Nations*. Glasgow edition. Oxford: Clarendon Press.

Stigler, George J. 1951. "The Division of Labor Is Limited by the Extent of the Market," *Journal of Political Economy* **59**(3): 185-193.

Stigler, George J. 1961. "The Economics of Information," *Journal of Political Economy* **69**: 213-225.

Stiglitz, Joseph E. 2002. "Information and the Change in the Paradigm in Economics," *American Economic Review* **92**: 460-501.

Williamson, Oliver E. 1979. "Transaction Cost Economics: The Governance of Contractual Relations," *Journal of Law and Economics* **22**(2): 233-261.

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

Williamson, Oliver E. 1996. "Economic Organization: the Case for Candor," *Academy of Management Review* **21**(1): 48-57.