October 2005

The Entrepreneurial Theory of the Firm and the Theory of the Entrepreneurial Firm

Richard N. Langlois
University of Connecticut

Follow this and additional works at: https://opencommons.uconn.edu/econ_wpapers

Recommended Citation
https://opencommons.uconn.edu/econ_wpapers/200527
The Entrepreneurial Theory of the Firm and the Theory of the Entrepreneurial Firm

Richard N. Langlois
University of Connecticut


June 2005, revised October 2005
Abstract

The entrepreneurial theory of the firm argues that entrepreneurship, properly understood, is a crucial but neglected element in explaining the nature and boundaries of the firm. By contrast, the theory of the entrepreneurial firm presumably seeks not to understand the nature and boundaries of "the firm" in general but rather to understand a particular type of firm: one that is entrepreneurial. This paper is an attempt to reconcile the two. After briefly delving for the concept of entrepreneurship in the work of Schumpeter, Kirzner, and (especially) Knight, the paper makes the case for the entrepreneurial theory of the firm. In such a theory, the firm exists as the solution to a coordination problem in a world of change and uncertainty, including Knightian or structural uncertainty. Taking a historical or developmental perspective, the paper then examines the changing nature of the entrepreneurial coordination problem over the life-cycle. In this formulation, "the entrepreneurial firm" is a nascent firm or proto-firm facing a problem of coordinating systemic change in economic capabilities. Lacking (by definition) adequate guidance from existing systems of rules of conduct embedded in markets or organizations, the entrepreneurial firm typically relies on a form of organization Max Weber called charismatic authority. In the end, although there is no such thing as a non-entrepreneurial firm, firms that must solve coordination problems in a world of novelty and systemic change ("entrepreneurial firms") are perhaps the purest case of the entrepreneurial theory of the firm.

Journal of Economic Literature Classification: B25, L22, M13

Keywords: entrepreneurship, transaction costs, coordination, Coase, Knight, Schumpeter, Weber.

Paper for the conference "Why Do Entrepreneurial Firms Exist?" Ohio State University, October 27-29, 2005. To be considered for a special issue of the Journal of Management Studies.
Introduction.

Why do entrepreneurial firms exist? When I first heard the question, it seemed an easy one to answer. “Entrepreneurship” and “firm” were part of the prose I had been speaking all my academic life. After a bit of thought, however, it became clear that this was not exactly the question I had been trying to answer for more than 20 years. My question has really been Coase’s question: why does the firm exist? For me, entrepreneurship is not part of the question; it’s part– or maybe even all – of the answer. The firm exists because of entrepreneurship. To put it another way, I have been working on the entrepreneurial theory of the firm, whereas the question posed seems to be calling for a theory of the entrepreneurial firm. These are not the same thing. To assert that there exists an “entrepreneurial firm” implies that there must also exist non-entrepreneurial firms and that entrepreneurial firms exist for reasons different from those that give rise to non-entrepreneurial ones.

As is so often the case, of course, some of the issues turn on how we use words. But the real issue, it seems to me, is this: firms come in different types. They are organized in a variety of ways. Some of these differences may be a matter of “life-cycle,” that is, of age or stage of development (Rathe and Witt 2001). To say that there is such a thing as an “entrepreneurial firm” – whatever that will turn out to be – is to challenge any monistic theory of the firm. One has to show that one’s theory applies, perhaps with caveats, qualifications, or
interpretations, to this special kind of firm and, more generally, to the panoply of things we call “firms” and to “the firm” throughout its life-cycle. So I welcome the question posed as an opportunity to think about that issue. Unsurprisingly, it will turn out that the entrepreneurial theory of the firm actually has a good deal to say about the entrepreneurial firm.

**Entrepreneurship.**

The firm exists because of entrepreneurship. I meant this claim as more than provocation (though I certainly meant it to be provocation). In the end, the cluster of ideas I find central in explaining the firm is the same cluster of ideas that is central to the literature on entrepreneurship. To be sure, this latter is a tangled and often somewhat amorphous literature; but there are certainly some central currents that flow out of the tradition of Knight (1921), Schumpeter (1934), and Kirzner (1973). In their recent attempt to set forth a research program in entrepreneurship, Shane and Venkataraman define the field as

the scholarly examination of how, by whom, and with what effects opportunities to create future goods and services are discovered, evaluated, and exploited. Consequently, the field involves the study of sources of opportunities; the processes of discovery, evaluation, and exploitation of opportunities; and the set of individuals who discover, evaluate, and exploit them. (Shane and Venkataraman 2000, p. 218).

Arguably, then, entrepreneurship is about the new -- new goods and services, but more generally new economic knowledge – and about how the new enters
the economic system. To put it another way, entrepreneurship is about change. It is about how the organization of economic activity extends and reshapes itself.

The theme of novelty and change is especially clear in Schumpeter, for whom entrepreneurship is the carrying out of new combinations, and in Kirzner, for whom entrepreneurship is the perception of new frameworks of means and ends. But it also present in Knight, buried inside his often-misunderstood discussion of how economic agents respond to uncertainty (Langlois and Coşgel 1993). The problem lies with the very word “uncertainty.” To be “uncertain” means to be ignorant of a well-defined piece of information within a well-defined information structure. Heads or tails? How much will Noah weigh at his next visit to the pediatrician? But what Knight meant by uncertainty is really a deeper kind of ignorance – ignorance about the information structure itself. I have long tried to rephrase Knight’s distinction between risk and uncertainty as a distinction between parametric and structural uncertainty (Langlois 1984). And structural uncertainty comes about because of the complexity of, and because of continual change in, the structure of economic activity. Thus an economic agent faced with Knightian uncertainty must confront economic change no less than must a Schumpeterian or Kirznerian entrepreneur.

---

1 For even longer, Brian Loasby (1976) has stressed the importance to economic activity of genuine ignorance rather than mere uncertainty, and his ideas influenced me from very early on.
Indeed, we might say that these three authors represent the three aspects of entrepreneurship to which Shane and Venkataraman point. Kirzner is about discovery, about alertness to new opportunities; Knight is about evaluation, about the faculty of judgment in economic organization (more on which presently); and Schumpeter is of course about exploitation, about the carrying out of new combinations and the creative destruction that often results therefrom.

*The entrepreneurial nature of the firm.*

Methodologically and by temperament, Coase is about as far from a Platonist as one can get. And that, to my mind, is all to the good. But when one writes an article searching for the nature of the firm, one is necessarily seeking something essential, some fundamental logic. Coase wasn’t looking for all the things that make real-world firms different from one another; he was looking for what is common and central to all firms. And Coase found that “essence” in the contrast, along two dimensions, between a firm and the ideal of a spontaneous market: (1) markets are about the exchange of products or outputs, and (2) that exchange is coordinated spontaneously. “Spontaneously” means that, at least in the ideal, sellers self-identify; that is to say, the sellers themselves and not the buyers determine how the sellers will allocate their effort and what they will produce. And in making their independent decisions, the sellers are guided and disciplined by relative prices rather than by administrative direction and fiat. A firm stands in contrast to both aspects of markets: it replaces contracts for
products with employment contracts, effectively substituting a factor market for a product market (Cheung 1983); at the same time, it replaces spontaneous coordination with some kind of central design or direction. Notice that this leaves two unexamined alternatives: product markets governed by central direction and factor markets coordinated spontaneously.

Figure 1 may help make this clearer. The columns distinguish whether participants in production self-identify. Do they make a product or undertake a task because someone told them what to do? Or do they choose themselves which product to make or task to undertake? The rows distinguish whether

<table>
<thead>
<tr>
<th>Participants contribute products</th>
<th>Participants self-identify (design)</th>
<th>Participants self-identify (spontaneity)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inside contracting Outsourcing “Virtual” firms</td>
<td>Classic market</td>
<td></td>
</tr>
<tr>
<td>Classic (Coasean) firm</td>
<td>Voluntary (open-source) production</td>
<td></td>
</tr>
</tbody>
</table>

Figure 1
Adapted from Langlois and Garzarelli (2005)
participants make a product or provide effort directly. If I want a widget, I can buy one from you or I can hire your effort directly. In the former case, I pay for the widget; in the latter case I pay for your time and exertion. (This is what Cheung meant when he said that the firm substitutes a factor market – buying effort – for a product market – buying widgets.) In the northeast box is the classic market of Econ 101: suppliers offer products for sale, and they do so effectively “on spec.” In the northwest box, suppliers offer products for sale, but they do so under someone else’s direction. Their products are effectively “bespoke.” Think here of Michael Dell’s highly organized use of outside suppliers or even the larger idea of a “virtual” corporation. In the southwest box lies the classic firm: participants (workers, they would be called) supply effort directly, and they do so under the direction of others. What could possibly fill the southeast box? How could participants supply effort directly without someone telling them what to do? Pursuing the question here would take us too far afield; but Langlois and Garzarelli (2005) argue that the fourth box is filled by the spontaneous collaborative mode of organization of which open-source software development is an important example. Bottom line: the classical Coasean firm-market spectrum is really the diagonal from the southwest to the northeast.

Of course, Coase asked not merely what a firm is but also the why such a modality of organization should exist. His famous answer is that there are “costs of using the price mechanism” – costs that can be avoided by organizing
activities within a firm. What are these costs? At first, they seem to be search costs and transaction costs of a purely frictional type. The “most obvious cost of ‘organising’ production through the price mechanism,” he says, “is that of discovering what the relevant prices are” (Coase 1937, p. 390). A second type of cost is that of executing separate contracts for each of the multifold market transactions that would be necessary to coordinate some complex production activity. Notice, however, that costs of these two sorts really exist only under circumstances of novelty and change. If nothing changes in my pattern of transacting, I won’t need to keep searching, and I can continue to trade with the same partners over time. Moreover, if I’m sure nothing unexpected will happen, I can further reduce the frictional costs of contact-writing by arranging a single long-term contract with each partner.

Thus, as Coase recognizes implicitly, transaction costs are always costs of novelty and change, and it is here that the firm possesses a cost advantage. “It may be desired to make a long-term contract for the supply of some article or service,” Coase writes.

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 commodity or service. 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. ... 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 essence of the firm, and its source of advantage over spontaneous product markets, lies in its flexibility in circumstances of change and uncertainty.²

Let’s unpack this further. Why is a firm more flexible? By substituting an employment contract for a spot contract in output, “the buyer” can manage economic activity in real time. As Herbert Simon (1951) explains, under an employment relation, “the buyer” pays a wage for the right to choose which action $x \in \Omega$ the worker will perform, where $\Omega$ is the “job description” or set of allowable actions for which the worker contracts. The worker thus agrees ahead of time to the abstract contours of what he or she may be asked to do; the worker also agrees that, within those limits, the wage-payer has authority – the right to dictate a decision in any circumstances not spelled out explicitly in the original contract (Tirole 1988, p. 464).

Okay. But why are markets not flexible in the face of uncertainty? Alchian and Demsetz challenged the idea that authority is the fundamental

---

² The reader who approaches from the mainstream economics of organization may be astounded that I follow Coase (and Knight) in seeing the firm as solving a coordination problem rather than an incentive problem in the face of highly specific assets. To a follower of Williamson (1985), explaining the firm without asset specificity would seem like performing Hamlet without the Danish Prince. I won't rehash the arguments here, but I have suggested elsewhere (Langlois and Foss 1999) that, in effect, it's actually more like
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). The answer is that 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 2003a, 2003b). In the case of the firm, the buyer (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 buyer (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. (As when my wife calls with a changed menu after I’ve already begun shopping.) Effectively, a thick (spot) market for commodities plays a buffering role similar to Simon’s set \( \Omega \): 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.

There are important implications to all this. Firms are not lower-cost mechanisms of adaptation under all circumstances and in all times and places.

---

performing Hamlet without Gildersleeve. (My coauthor on that paper is a Dane who lives only a few miles from Elsinore, so he should know.)
Whether a firm is superior to markets will depend (1) on the state of relevant actually existing markets and (2) on the nature of the adaptation problem involved.

The firm, we saw, involves some degree of design and direction. In a firm, workers do not self-assign to tasks, and the overall pattern of the organization of tasks bears the stamp of design. Thus, when firms are superior to (have advantages over) markets in flexibility, it must be when design and direction of tasks has advantage over “spontaneity” and self-identification to tasks. To put the matter in business school terms: for firms to have advantages over markets, using design and direction to buffer uncertainty must add value. (Indeed it must add enough value to also overcome the costs of the moral hazard created by a wage contract.)

It is easy to add value over markets if markets simply don’t exist. When Alchian and Demsetz walked into an Albertson’s in Los Angeles in the 1970s, the shelves were full of stuff, and there was a Safeway just down the street. So it was extremely easy for them to exercise their “authority” over the grocer – that is, to design whatever pattern of final goods they wanted without having to manage

---

3 As in the Coasean tradition, I have been portraying the line between firm and market as a sharp one. In fact, however, these are both ideal types. Actually existing firms arguably develop in ways that are “spontaneous” in many respects; and actually existing markets arguably exhibit important elements of design. As Alfred Marshall understood, both firms and markets “are structures for promoting the growth of knowledge, and both require conscious organization” Loasby (1990, p. 120). As I will suggest below, what distinguishes the ideal type of the market from that of the firm (along this dimension) is the degree of abstractness of the design involved (Langlois 1995b).
directly the effort that went into producing those goods. Consider, by contrast, Charles Ingalls and family in the woods of Wisconsin in the 1860s. With no grocers to boss around, they found it much cheaper to operate as a kind of firm – to decide how to allocate their own effort to baking bread and catching fish (though not tuna, presumably). There were perfectly good markets for bread and fish in Milwaukee, but transaction costs – transportation costs, really – made it uneconomical to extend those markets very far. To put it another way, the extent of the local market was too small to make trade in products economical, necessitating the direct management of effort.

But this is exactly why we pay secretaries by the hour rather than by the piece. The extent of the market for any particular service (like typing) is too small to make it worth paying the fixed set-up costs and variable measurement costs of a per-piece system.\(^4\) In this case, however, it is uncertainty rather than the friction of wagon wheels on rutted dirt roads that makes markets costly. What an office actually produces is coordination, and the costs are high of failing to match the proper coordinating service (typing, filing, phone answering) with the proper moment in time. This raises the costs of a per-piece system or, to put

---

\(^4\) In the days before computers, there was a cottage industry in typists around universities who would type student papers at a per-page rate. In large organizations of the period, there was a specialization of roles, like receptionists, telephone operators, and the “typing pool,” though employees in these roles were typically paid by the hour not the piece. (Notice that those functions least costly to specialize and pay by the piece are precisely the ones most likely to be taken over by computers (Langlois 2003a).) Nowadays there is a good deal of “outsourcing” of copying and “document processing” functions, but this often takes the form of substituting workers from the document-processing company for one’s
it another way, raises the extent of the market for coordinating services that would be necessary for such a market to buffer the uncertainty effectively. Moreover, as Coase points out, the costs fall more heavily on the buyer than the seller. This gives the buyer of coordinating services an incentive to specialize (and arguably a greater competence) in directly designing the work and directing the effort of others.

The coordination problem of the office as I have described it is really a problem of parametric uncertainty. All tasks fall into well-defined categories – typing, filing, answering the phone – on which all participants agree. The problem of coordination is one of knowing which task will be required at which moment. This gives the advantage to direction. I argue that this is only one of many possible ways in which change and uncertainty (including the Knightian kind) lead to a specialization by a buyer in directly managing effort – that is, lead to a firm in the essential Coasean sense. In this respect, Coase’s theory and that of Knight are fundamentally similar, stressing “the need for flexibility in an uncertain and hard-to-predict world” (Foss 1996, p. 78). In another sense, I will argue, the two theories are at odds, though the conflict will largely turn on different definitions of what constitutes a firm.

Knight’s theory is essentially Smithian, with a cognitive twist. The firm arises as a form of specialization. The source of the underlying advantage of this

---

own workers; the contractor’s workers then take direction from both their employer and
specialization, and thus the basis on which individuals will be sorted into this role, is the human faculty of judgment (Langlois and Cosgel 1993). Judgment is the (largely tacit) ability to make, under conditions of structural uncertainty, decisions that turn out to be reasonable or successful ex post. Knight located his ideas about behavior under uncertainty within the cognitive philosophy of his day, notably William James and Henri Bergson, and much of it is arguably armchair psychology (Foss 1996, p. 79). From my own armchair, I have suggested that modern currents in psychology, including evolutionary psychology (broadly understood) and the lessons of the debate over artificial intelligence, lead to a similar place. Humans have a cognitive comparative advantage over mechanistic rule-based system (like computers, bureaucracies, or even some kinds of formal markets) in dealing with situations of structural uncertainty, qualitative change, and informational ambiguity (Langlois 2003a). Interestingly, Knight himself put an evolutionary spin on the argument.

When uncertainty is present and the task of deciding what to do and how to do it takes the ascendancy over that of execution, the internal organization of the productive group is no longer a matter of indifference or a mechanical detail. Centralization of this deciding and controlling function is imperative, a process of

their employer’s customer.

Schumpeter described a similar faculty when he argued that entrepreneurship “depends on intuition, the capacity of seeing things in a way which afterwards proves to be true, even though it cannot be established at the moment, and of grasping the essential fact, discarding the unessential, even though one can give no account of the principles by which this is done. Thorough preparatory work, and special knowledge, breadth of intellectual understanding, talent for logical analysis, may under certain circumstances be sources of failure” (Schumpeter, 1934, p. 85). The sort of cognition that Knight and Schumpeter describe is normally discussed today in terms of Michael Polanyi’s (1958) notion of tacit knowledge.
“cephalization,” such as has taken place in the evolution of organic life, is inevitable, and for the same reasons as in the case of biological evolution. (Knight 1921, III.ix.8.)

In humans, cephalization involved seizing what the evolutionary psychologists call the *cognitive niche* (Tooby and DeVore 1987): an ability to learn, adapt, and solve problems in a creative way that enabled those humans to colonize and master virtually every kind of environment and every part of the planet. In effect, humans have always attempted to adapt to a complex and changing environment by design and direction.

Brian Loasby points out that, for Knight, “intelligence and entrepreneurship are both responses to uncertainty, to situations in which there is no correct procedure for deciding what to do. In such situations people must create their own structures for interpretation and decision, or find some ready-made structure that they are prepared to adapt; and these activities, rather than the rational choices that can be produced by automata, are the distinguishing characteristics of human intelligence.” (Loasby 2004). Judgment is Knight’s term for the process of creating frameworks of interpretation and decision. Those who specialize in the exercise of judgment we call entrepreneurs (Casson 1982, 2005). As in Coase, the firm arises because of an incentive for some to specialize in directing the effort of others. Those who lack competence or confidence in their judgment have a corresponding incentive to specialize in being directed.
Famously, of course, Coase (1937, p. 400-401) objected to Knight’s argument that specialization in judgment is the essence of the firm. Part of his objection is clearly wrong. When people have distinctive abilities, said Coase in effect, they specialize and trade with one another in markets; so why can’t one simply buy the (specialized) services of judgment on a market? The answer is: one can’t for a lot of reasons. Most prosaically, the sale of judgment in a market would lead to problems of moral hazard, since the seller of judgment could always provide less service than contracted for by disguising poor judgment as bad luck. So it makes sense for the provider of judgment to hire the services of those who are less costly to monitor rather than the other way around – that is, it normally makes better sense for the wielder of judgment to become residual claimant and to pay the other factors contractually (Barzel 1987). Moreover, there are well-known problems of selling any kind of information: if you don’t already have the information, you don’t know how much it’s worth; but if I tell you the information, then you don’t need to buy it (Arrow 1962). For all of these reasons, entrepreneurial judgment is in the end non-contractible.

Contrary to what I and others may have said or implied in the past (Langlois and Coşgel 1993; Langlois 1995a; Foss 1993, 1996; Foss and Klein 2005), however, the non-contractibility of judgment is not a sufficient explanation for the firm in Coase’s sense. The second part of Coase’s objection is right on target. To say that the entrepreneur cannot sell judgment is not to say that the entrepreneur cannot make use of markets to profit from that judgment. Suppose
I have an idea for a new kind of hand-held computer, and my powers of judgment tell me that it will be a hit. I can hire a design firm like Ideo to design it for me, a contract assembler like Solectron to build it, a New York ad agency to market it, and Amazon.com to sell it. And I can do this using contracts in output and without using employment contracts. I retain not only the right to residual income but also the right to residual control in my one-man virtual corporation. I do not have to create a firm in Coase’s sense. But I do end up creating a firm in Knight’s sense. For Knight, the essence of the firm lies in design and direction, not in contractual form.\(^6\) Coase’s firm is only the southwest quadrant of Figure 1; Knight’s is the entire western half.

Coase is in the end quite right that Knight’s theory cannot by itself distinguish when authority and control will be exercised through contracts in output (bossing around the grocer) and when exercised through employment contracts (bossing around the secretary). That requires comparing the relative costs and benefits of contracts in product and of employment contracts. Yet Knight’s analysis of behavior under uncertainty, or something very like it, actually provides some insight into that question. It may indeed provide a way to generalize to the case of structural uncertainty the coordination explanation Coase offers for (what I have argued is) the case of parametric uncertainty.

---

\(^6\) Most of the time, of course, Knightian “cephalization” does indeed imply a Coasean firm. As Knight insists, the exercise of judgment in business is almost always a judgment about other people’s judgment. And those other people often, and perhaps mostly, include
When the entrepreneur seeks to profit from judgment using solely market contracts for product, the result is a one-person firm or “virtual” organization. More typically, however, all the capabilities the entrepreneur needs are not available at low (transaction) cost like bread and tuna at Albertson’s. As we saw, this may be so for reasons of simple frictional transaction or transportation costs. Far more interestingly, and perhaps far more typically, this is so because of uncertainty. If parametric uncertainty can lead to asymmetrical costs in a problem of timing coordination, think what structural uncertainty can do. In a world in which market participants may not even share the same conceptual categories or interpret “information” in the same way, problems of coordination are even more fundamental. In such a world the impediments to market transaction are largely cognitive and informational. As Morris Silver points out, the problem of selling a genuinely new idea is not that someone else is likely to steal it but that no one else is likely to believe it. The entrepreneur’s problem “is that he cannot, at reasonable cost, convey his implausible ‘secret’ to those with the technical capabilities needed to produce the required operations at the lowest cost” (Silver 1984, p. 17). The transaction costs involved -- what I like to refer to as dynamic transaction costs (Langlois 1992) – are effectively costs of teaching and persuasion not of moral hazard and opportunism. These costs can be avoided by employees. Of course, if both parties to the transaction are exercising judgment, this raises more complex issues of the assignment of decision rights, to which I return.

---

7 The notion that the diffusion of knowledge is not instantaneous and requires persuasion is less foreign to those who actually study innovation than it is to economists. For example, in
(or at least reduced) by using an employment relation, since “[r]eliance on the
direction of some by others allows ... knowledge to be put to use guiding ...
activities without requiring those who actually execute them to master the
knowledge themselves” (Demsetz 1995, p. 32).

Notice that if economic change is slow or nonexistent, it may in fact be
cheaper in the long run to pay the set-up costs of transferring the necessary
knowledge to the workers themselves and to dispense with (what would then be)
the superfluous role of the manager. It is only under circumstances of
uncertainty, change, and information complexity that direct authority over the
effort of others retains advantage. Just as in the case of the timing problem of
secretarial services, entrepreneurs in a world of structural uncertainty may face
the costs of not having the capabilities they need when they need them, and they
may respond by directly designing and directing the effort of others.

The evolutionary nature of the firm.

So far I have tried to make the case that the cluster of factors allied to the idea of
entrepreneurship – novelty, change, structural uncertainty – are also at the heart
of the nature of the firm. But what of the entrepreneurial firm? To the extent
that an entrepreneurial firm is a type, or perhaps a stage in the life, of firms in
general, explaining the nature of such a firm involves moving from the abstract

his famous book Diffusion of Innovations, Everett Rogers (1995, pp. 167 ff.) emphasizes the
“persuasion stage” in the innovation-decision process. For a recent attempt to model the
phenomenon of dynamic transaction costs formally, see Hellmann (2005).
considerations adduced so far toward a more concrete and historical (if still in the end theoretical) account of the firm. As I have long contended (Langlois 1984), the way to do this is to place the Coasean question within a dynamic framework.\textsuperscript{8}

In a sense, of course, an entrepreneurial account of the nature of the firm is inherently dynamic, in that it encourages the future to intrude on the explanation of the firm. The explanation for the firm lies not, or at least not solely, in forces that are visible today; it lies instead in how people (sometimes) organize to deal with an uncertain future. (This is why we cannot understand comparative-institutional analysis as merely a comparison of snapshots capturing a moment in time.) But the past also intrudes on our explanation of actually existing firms. Especially when the environment provides a weak selection mechanism, we should expect organizational forms to be path-dependent and sometimes even ill adapted.\textsuperscript{9} Moreover, boundary conditions and facts about the world – what a historian would call history – also matter. As I’ve already suggested, the Coasean nature of the firm (as amended) does not accord firms an advantage over markets unconditionally in all times and places;

\textsuperscript{8} By “dynamic” I actually mean evolutionary (Langlois 1984, 1986), but I don’t want to get bogged down in the methodological details here.

\textsuperscript{9} For example, if the selection environment changes over time, an organizational form may be selected out in one period even though it would have proven highly adapted in a later period (Langlois 1984). If the selected-out form cannot be “reinvented” quickly or cheaply, we may find ourselves observing a survivor that is not “optimal” in the sense that it is not the best adapted of all historically known forms (let alone all possible forms, whatever that would mean).
instead, any such advantage depends on the state and extent of existing markets and on the structure of the coordination problem that presents itself.

With all this in mind, Paul Robertson and I (Langlois and Robertson 1995) have proposed a way to think about organizational change and development. Three factors are important:

1. The pattern of existing capabilities in firm and market. Are existing capabilities distributed widely among many distinct organizations or are they contained importantly within the boundaries of large firms?

2. The extent of the market and the level of development of market-supporting institutions. To what extent can the needed capabilities be tapped through existing arrangements and to what extent must they be created from scratch? To what extent are there relevant standards and other market-supporting institutions?

3. The nature the economic change called for. When technological change or changes in relative prices generate a profit opportunity, does seizing that opportunity require a systemic reorganization of capabilities (including the learning of new capabilities) or can change proceed in autonomous fashion along the lines of an existing division of labor?

One scenario in which the firm has an advantage over markets is when (1) existing capabilities are dispersed into decentralized markets; (2) markets in general are thin and market-supporting institutions are weak; and (3) an
entrepreneurial opportunity arises that demands a systemic rearrangement of capabilities, possibly including the development of wholly new capabilities. In this context, the dynamic transaction costs of informing, teaching, and persuading suppliers – if these can even be found – are high. It is cheaper for the entrepreneur to integrate the necessary capabilities vertically in order to directly direct and design the effort of others.

Robertson and I pay attention to the case of Henry Ford, who integrated vertically to an extremely high degree at both Highland Park and the Rouge. He did so, we argue, because the moving assembly line was a systemic reorganization of the way automotive parts were produced. Although Ford would have liked to have enlisted outside suppliers, he and his engineers were in fact still inventing the process; they were engaged in and benefiting from the transfer of rich systemic information within the organization, information whose complexity made it costly to transfer to outsiders with any alacrity (Langlois and Robertson 1989). Indeed, this scenario arguably provides a coherent theoretical account of the general rise of the large vertically integrated organization that Alfred Chandler chronicled in The Visible Hand (1977) and other works (Langlois 2003b). As population and per-capita income rose in the second half of the nineteenth century, and as the railroad, telegraph, and inland waterways lowered transportation costs and frictional transaction costs, there arose entrepreneurial opportunities for those who could systemically rearrange
industries away from small-scale local production toward centralized mass production with efficient transport.

This is very clearly an entrepreneurial account of the formation and *raison d'être* of firms. But two rather glaring questions quickly emerge. First, this is an account of why firms should have advantages over markets in the entrepreneurial short run. What of the longer term? The Langlois-and-Robertson framework would predict that, as economic change slowed and time passed, the new pattern of capabilities would filter out to the market and dynamic transaction costs would decline. How then can the framework explain the *persistence* of the large vertically integrated Chandlerian firm? Second, did the *raison d'être* of these firms remain an entrepreneurial one as the Rockefellers, Swifts, and Fords gave way to Chandler’s managerial class and to public stock ownership? Aren’t Chandlerian firms in any case a far cry from the kind of firm Coase seemed to be describing? Isn’t General Motors a far cry from bossing around the secretary? Can such firms really have the same “nature” as the abstract firms Coase discussed? Let me take these questions in reverse order.

Although the entrepreneurial account I have given of the rise of the large vertically integrated corporation is fully consistent with the story Chandler tells, the *raison d'être* Chandler himself assigns to these firms is actually something rather closer in spirit to Coase’s problem of temporal coordination. In both cases, timing is asymmetrically important to the entrepreneur. In Chandler’s case, this
is because of the kinds of production processes made profitable by lower transportation and transaction costs typically involved high fixed costs. In order to keep unit costs low, firms had to ensure high throughput so they could amortize the fixed costs over as large an output as possible. This meant using management – authority to design work and direct effort – to buffer uncertainty and keep the system humming (Langlois 2003a, 2003b). Especially since, as I argue, the right kinds of markets were frequently scarce and the right kinds of market-supporting institutions weak, direct management was initially the only way to solve the buffering problem. So firms integrated vertically, bringing many stages of the process within the buffering ambit of management.

It is a crucial element of Chandler’s story that, as these firms became larger and more successful, they found it necessary to subdivide managerial labor and to organize management in hierarchical fashion, ultimately leading to what Chandler calls the M-Form (multidivisional) structure. Coase (1937) never deals with such issues of structural elaboration, but Knight does. For Knight, the exercise of judgment almost always involves the judgment of other people’s judgment.  

In a hierarchical organization, then, when a manager assign a worker to task $x \in \Omega$, that $x$ is frequently a task of judgment, which may in turn require the worker to select a task from some other worker’s set $\Omega$. And so on down the line. In Knight’s schema, it is really ultimate judgment that is not
contractible, where ultimate judgment is the act of judgment at the very top of the hierarchy, the act that sets the others in motion.

In the simple theory of the firm, the person who exercises (noncontractible) judgment becomes the owner and possesses the claim to residual income and the residual rights of control. But who is the “owner” of the corporation? This is actually quite controversial. Some, including Demsetz (1995), think that managers are the real owners of the corporation, since they are the ones who exercise effective day-to-day control. Knight would have argued otherwise (Langlois and Cosgel 1993). For Knight, “[w]hat we call ‘control’ consists mainly in selecting someone else to do the ‘controlling’” (Knight 1921, III.x.2). As is it is ultimately the holders of common stock who decide who shall be the managers, it is the stockholders who are the owners, possessing both the residual claim and the residual control rights. This may sound absurd. Surely my one share of Microsoft doesn’t give me any effective ability to boss Bill Gates around. But Henry Hansmann (1988) has explained cogently why it may not be absurd at all. In the case of the corporation, it is desirable to place ownership in the hands of a class of patrons – the stockholders – who do not have an effective day-to-day ability to exercise that control. The reason is that giving stockholders the right is the least evil of all alternative evils, since any other allocation of rights

---

10 “The ability to judge men in relation to the problems they are to deal with, and the power to ‘inspire’ them to efficiency in judging other men and things, are the essential characteristics of the executive” (Knight 1921, III.x.34).
(notably one in favor of the managers) would result in a worse allocation of resources.

Bottom line: judgment is at the heart of even an articulated hierarchical firm. Managers at all levels are acting “entrepreneurially” whenever they employ the faculty of judgment. The nature of the corporation is the same as the nature of the Coasean firm.

So why do firms, including Chandlerian firms, often continue to exist long after the original entrepreneurial circumstances of their birth have faded. One simple answer is that, although the nature of the entrepreneurial problem that call forth a firm may change, the firm stays in business so long as their remain some kinds of entrepreneurial problems calling for judgment. Such an answer would not be wrong; but it would also not really be very helpful, since it doesn’t say much about the ways in which entrepreneurial problems differ among firms or change over the life of a firm.

A more complicated answer would look something like this. Chandler’s Visible Hand is not the only scenario that flows from the Langlois-and-Robertson framework. One would expect that, as time passed, the capabilities that were once managed more cheaply within the corporation would start to leak out to others. At the same time, markets would reorient themselves to the new structure of production and would become thicker with economic growth. And market-supporting institutions (like bodies of law and technical standards)
would emerge to lower the costs of market transaction. As a result, we would expect that, absent major new exogenous sources of systemic change, the vertical structure of production would slowly come unglued, leading to widespread “deverticalization” and a return to something that looks a lot more like market coordination. And, by the end of the twentieth century, this is exactly what did happen – a phenomenon I call the Vanishing Hand (Langlois 2003b, 2004; Lamoreaux, Raff, and Temin 2003). There are certainly still plenty of large firms around. But the likes of Microsoft, Intel, and Cisco are far less vertically integrated than the IBM of 1960 or the Standard Oil of 1910. And firms today can buy on markets a plethora of services – design, contract manufacture, payroll management, shipping logistics – that Chandlerian firms once had to provide for themselves.

But what happened in the century between the railroad and the Internet? Why did high levels of verticalization persist until the late twentieth century, long after the passing of the original entrepreneurial design problem that gave rise to most of these firms? The answer has to do with path dependency and the nature of the selection environment (Langlois 2003b). Once the Chandlerian vertical structure of production took hold in an industry, that structure tended to be self-reinforcing. For one thing, the buffering problems of high-throughput production demanded tight temporal coordination; and if all firms in a particular industry are vertically integrated, it becomes difficult for specialized supplier firms to get a foothold let alone to grow abundant enough to be able to take over
the role of buffering that uncertainty. At the same time, a highly integrated structure tended to select for and adopt innovations consistent with, and therefore amplifying of, that structure. Moreover, the golden age of corporate America – the decades just after World War II – were arguably a period in which the selection environment was rather lenient. The major competitor economies of Japan and Europe took decades to rebuild after the war, giving American firms a period in which they could enjoy the rents of reduced international competition. One of the ways in which they took those rents was organizational inertia. With the rise in oil prices in the 1970s and the simultaneous resurgence of Japan and Europe, however, the chickens of international competition came home to roost, and the leveraged-buyout wave of the 1980s began to unmake Chandlerian firm after Chandlerian firm. By the 1990s, new digital technologies and globalization had begun completely to redesign the landscape.

**The nature of the entrepreneurial firm.**

The pipe-smoking, thin-lapelled executive of a 1950s Chandlerian firm was still – I claim – a Knightian entrepreneur. He (probably not she) exercised judgment in an uncertain world and coped with problems of coordination by directly designing and directing the effort (including the judgment) of others. This is technically true. At a fundamental philosophical level, the nature of even a
Chandlerian firm in the black-and-white era is entrepreneurial. But let’s get real. This is not what most people have in mind by an entrepreneurial firm.

We have learned from Edith Penrose, George Richardson, and others that firms are both enabled and constrained by the resources and capabilities they already possess. Those capabilities are really an accretion of the practical knowledge and rules of conduct built up through successful action over time, what Nelson and Winter (1982) call routines. The idea of an organization as a system of rules is also familiar from Max Weber’s (1947) famous account of bureaucracy. The upshot of this is not that firms are doomed to inertia and cannot cope with let alone generate novelty; rather, the implication is that what novelty firms can generate will be limited by the capabilities they already possess. As Penrose (1959) explains, firms actively seek new entrepreneurial opportunities in order to take advantage of unused resources, importantly including knowledge resources. But the other side of the coin is that the firm is limited to those bits of novelty to which its existing resources are relevant. As Richardson (1972) puts it, firms are only good at taking on new activities when

---

11 Which is not, of course, a trivial assertion, since it contradicts the widely held view of John Kenneth Galbraith (1967) that the modern corporation had entirely eliminated uncertainty through rational “planning,” a course of action he recommended for the economy as a whole.

12 Notice that the development of rules and routines within organizations is a (path-dependent) substitute, or partial substitute, for the development of external markets and market-supporting institutions. (These last, of course, are also systems of rules of conduct.) In both cases, the passage of time leads to the handing-off of conscious direction to an unselfconscious process (Langlois 1995c) – effectively moving to the right in Figure 1. I return to this point below.
they require knowledge similar to what the firm already possesses. Cohen and Levinthal (1990) have argued that, even in the realm of research and development, firms are very often engaged not so much in a search for novelty as in a quest for sufficient “absorptive capacity” to be able merely to recognize and understand relevant novelty when they see it.

It is certainly entrepreneurial to seize profit opportunities that fit in with a highly articulated structure of existing knowledge. But much of the sense of the term “entrepreneurial” carries with it the implication of novel recombination that is somehow more radical, or at least less constrained. Like Adam Smith’s philosopher, an entrepreneurial firm combines together “the powers of the most distant and dissimilar objects” (Smith 1976, I.i.9). In order to effect a radical recombination of elements, especially one that is systemic, a firm must be free of the past. It must be free of memory in order to imagine the future in a new way. This implies that an entrepreneurial firm must be either a new firm or a firm somehow willing and able to creatively destroy its own memory. As Schumpeter noted, the former is far more likely than the latter: “new combinations are, as a rule, embodied, as it were, in new firms which generally do not arise out of the old ones but start producing beside them; ... in general it is not the owner of stage-coaches who builds railways” (1934, p. 66). The issue is fundamentally a cognitive one (Langlois 1995b; Noteboom 2003), and it extends beyond even firms to networks and communities of practice. “Old communities and traditions virtually never give birth to radically new technologies,” writes the
historian of technology Edward Constant. “No manufacturer of piston aircraft engines invented or independently developed a turbo-jet. No designer of conventional reciprocating steam engines invented a steam turbine, no manufacturer of steam locomotives independently developed diesel engines. In the case of both firms and individuals, community practice defines a cognitive universe that inhibits recognition of radical alternatives to conventional practice” (Constant 1980, p.).

All of this means that “the entrepreneurial firm” as I understand it is a firm that operates in the kind environment Silver (1984) described, a world in which dynamic transaction costs are high. In such a world, “cephalization,” the deliberate arrangement and design of capabilities, is necessary to take advantage of a profit opportunity.

Under the right circumstances, of course, an entrepreneur can effect systemic recombination using markets. Sometimes markets for components are thick enough and market-supporting institutions (notably, in this case, technical standardization) are strong enough that recombination can proceed in autonomous fashion – the personal computer is a case in point (Langlois and Robertson 1992). In other instances, there may be a life-cycle effect at work. In the American automobile industry, for example, the earliest car makers were assemblers who bought or commissioned parts from existing small machines shops. It was only with the invention of mass production at Ford that systemic
change required integration\textsuperscript{13} (Langlois and Robertson 1989). In other cases, like many of those Chandler chronicled, systemic change was essential from the start and relevant markets were thin or nonexistent. In still other cases, existing organizational capabilities had to be destroyed and systemically reconstituted, as happened in the Swiss watch industry in the 1980s (Langlois 1998b).

What is common to the entrepreneurial firm in all these cases is that it involved self-conscious design. Entrepreneurial firms are always on the far west of Figure 1. This is so because, rather by definition, they do not draw on existing unselfconscious repositories of knowledge and capability, whether these be existing market patterns or existing systems of rules of conduct within organizations.\textsuperscript{14} This is why entrepreneurial firms are sources of systemic novelty.

If we mean by an entrepreneurial firm a firm in Coase’s sense not just in Knight’s, then design also involves direction of the effort of others. The

\textsuperscript{13} Clayton Christensen and his coauthors have turned this scenario into a general cyclical theory of vertical integration. In their account, vertical integration has advantages over market procurement because it allows better “fine-tuning” and therefore higher performance. When performance ceases to be an issue – as when technological change has increased levels of performance beyond what the market demands – costs become a consideration, at which point markets begin to excel. Performance and costs issues can alternate in importance, leading to a cycle of vertical integration and disintegration. Noteboom (2005) also proposes a cyclical theory of entrepreneurial organization, though his is based on the cognitive issues discussed earlier. Rather than seeing these as the general theory, however, I tend to see them as interesting special cases of the theory of integration and disintegration in Langlois and Robertson (1995).

\textsuperscript{14} More correctly: entrepreneurial firms rely on systems of rules of conduct that are more abstract, less concrete, than those built up over time in firms and markets (Langlois 1995c). Even classical markets rely on institutions, notably a system of property rights, that provide the abstract design rules of the system as a whole (Langlois 2002).
transaction-cost problem here is one of reducing the costs of informing and persuading those others with whom the entrepreneur collaborates. As we saw, designing and directing the effort of others with the help of an employment relation reduces these dynamic transaction costs because it reduces the amount of knowledge that has to be transferred: the entrepreneur retains the knowledge and simply chooses activities for the workers that will make use of the knowledge. Costs of persuasion are also low, as Coase pointed out, because the worker is indifferent about which \( x \in \Omega \) the entrepreneur chooses so long as he or she gets an hourly wage.

Knight reminds us, however, that the “effort” the entrepreneur must direct actually involves the exercise of judgment. This complicates the problem somewhat, since, even in a small entrepreneurial firm, there may be a good deal of delegation of judgment, and that judgment may cover a wide ground. In a large, older, more-articulated firm, judgment is coordinated – that is, workers are kept on the same page – in large part by the systems of rules of conduct that are also the firm’s capabilities and its organizational memory. This was Max Weber’s point about bureaucracy. By (my) definition, however, an entrepreneurial firm lacks such systems of rules of conduct. How then is judgment coordinated in the entrepreneurial firm? The answer is what Ulrich Witt (1998) calls \textit{cognitive leadership} and what, following Weber, I call \textit{charismatic authority} (Langlois 1998a, 1998b).
Essentially, charismatic authority is a mechanism for coordinating innovation, for coordinating organization once it steps outside the cognitive bounds of existing firms and networks of practice. Weber describes it this way.

There is no system of formal rules, of abstract legal principles, and hence no process of judicial decision oriented to them. But equally there is no legal wisdom oriented to judicial precedent. Formally concrete judgments are newly created from case to case and are originally regarded as divine judgments and revelations. ... The genuine prophet, like the genuine military leader and every true leader in this sense, preaches, creates, or demands new obligations.

In the pure type of charisma, these are imposed on the authority of revolution [sic] by oracles, or of the leader’s own will, and are recognized by the members of the religious, military, or party group because they come from such a source. (Weber 1947, pp. 360-361.)

“Within the sphere of its claims,” he adds, “charismatic authority repudiates the past, and is in this sense a specifically revolutionary force” (Weber 1947, pp. 361-362).

We normally think of charismatic authority – whether in the form of a military commander, cult leader, or visionary entrepreneur – as “irrational.” In Weber’s meaning, it is exactly that, since its logic derives from neither traditional nor rationally designed rules. But that doesn’t make it irrational in the modern economist’s sense (that is, inefficient). As the sociologist James Coleman (1990) argued, charismatic authority is rational in the sense that it solves a problem of coordination. In a world without rules, charismatic authority provides a structure to which entrepreneurial collaborators can orient themselves. Quite apart from any motivational effects it may have, such authority serves to keep
everyone on the same page. This is not in the end far removed from Coase’s conception of the coordination benefits of authority, though it speaks to issues of cognition as well as flexibility. Charismatic authority is a way of reducing dynamic transaction costs by packaging a bundle of complex knowledge and information in a form that others can cheaply absorb. As Witt puts it, the entrepreneur implements and defends his or her own business conception as “a tacit cognitive frame collectively shared within the firm” (Witt 1998, p. 161).
References.


