April 1996

Capabilities and Governance the Rebirth of Production in the Theory of Economic Organization

Richard N. Langlois  
University of Connecticut

Nicolai J. Foss  
Department of Industrial Economics and Strategy, Denmark

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

Recommended Citation
http://digitalcommons.uconn.edu/econ_wpapers/199602
Capabilities and Governance the Rebirth of Production in the Theory of Economic Organization

Richard N. Langlois
University of Connecticut

Nicolai J. Foss
Department of Industrial Economics and Strategy, Denmark

Working Paper 1996-02

April 1996
Abstract

Almost a decade ago, Paul Milgrom and John Roberts (1988, p. 450), two of the leaders in the formalist branch of the New Institutional Economics, made the following observation. “The incentive based transaction costs theory has been made to carry too much of the weight of explanation in the theory of organizations. We expect competing and complementary theories to emerge - theories that are founded on economizing on bounded rationality and that pay more attention to changing technology and to evolutionary considerations.” This paper argues that such theories are now emerging. We survey and synthesize a developing perspective that we label the “capabilities” view. We argue that this view complements incentive-based theory (1) by considering the problems of imperfect knowledge in production as well as in governance and (2) by considering issues not only of incentive alignment but also of qualitative coordination among holders of specialized, distributed, and often tacit knowledge. Also, focusing on capabilities brings to the fore the idea that routines and similar rule-based forms of institutionalized knowledge may be important building blocks of economic organization. As a result, the capabilities approach arguably connects more fully with the New Institutional Economics, in which rule based guides to action like norms and conventions play a fundamental role, than do approaches that take the transaction as the unit of analysis.
I. Introduction

In the last 25 year or so, the economics of organization inspired by the seminal work of Ronald Coase (1937) has emerged, perhaps belatedly, as a thriving branch of economics. In spite of some not inconsiderable variety among the contributions to this field, it is fair to say that the literature is in agreement on the fundamentals. The basic insight is this: in addition to production costs of the usual sort, one must also consider transaction costs in explaining institutions like the firm. Whether called transaction-cost economics (Williamson 1975, 1985) or the economics of organization more broadly (Milgrom and Roberts 1992), this blossoming filed has indeed focused precisely on the comparative transaction costs of alternative organizational structures, including, paradigmatically, the choice between firms and markets. Moreover, the literature has seen the “nature” of the firm — and, indeed, of other institutions — as fundamentally contractual.¹ Firms and other institutions are alternative bundles of contracts, understood as mechanisms for creating and realigning incentives. To put it another way, the economics of organization has shown a tendency (albeit an imperfect tendency) to respect an implicit dichotomy between the production aspects and the exchange aspects of the firm — that is, between production costs and transaction costs.

We do not mean to say by this that present-day theory depicts production as completely unaffected by exchange. In fact, the crucial point of some extremely influential recent research has been to demonstrate rigorously that

¹ And this is arguably so even for those who do not follow Steven Cheung (1983) and others in seeing the firm as nothing other than a “nexus of contracts.”
alternative organizational structures might be chosen because they imply different incentives to invest in specific assets (Grossman and Hart 1986; Hart 1995). In many recent models, indeed, technology and organizational structure are determined jointly (Riordan and Williamson 1985; Milgrom and Roberts 1990). What we do mean is rather that there exists an odd and unjustified allocation of responsibilities between price theory and the economics of organization. To price theory has been consigned the basic theory of production, with an implicit agreement that the production function, and its attendant assumptions, tells us what we need to know about production costs. In price theory, productive knowledge is seldom portrayed as imperfect or asymmetric, let alone tacit or “sticky” (Demsetz 1988; Winter 1988). Knowledge about alternative production possibilities is explicit, freely transmissible, and easily encapsulated in what Joan Robinson (1956) called “blueprints.”

By contrast, imperfect knowledge is arguably the raison d’être of the modern literature on the economics of organization. To an overwhelming extent, however, all such imperfections — all deviations from the assumptions of the production-function formulation — are seen as falling exclusively in the realm of transaction costs. In today’s economics of organization, transacting is fraught with hazards, and the problem of organization is one of creating governance structures to constrain the unproductive rent-seeking behavior that imperfect information permits. Indeed, it is probably not unfair to say that the heuristic driving this literature is to reduce virtually all problems of economic organization to problems of misaligned incentives attendant on imperfect information.
The result of this partition of responsibilities has been an imbalance in the economics of organization. Seldom if ever have economists of organization considered that knowledge may be imperfect in the realm of production, and that institutional forms may play the role not (only) of constraining unproductive rent-seeking behavior but (also) of creating the possibilities for productive rent-seeking behavior in the first place. To put it another way, economists have neglected the benefit side of alternative organizational structures; for reasons of history and technique, they have allocated most of their resources to the cost side.²

Our goal here is to attempt briefly to document and criticize this intellectual partition. More importantly, however, we suggest that the partition is beginning to break down. This latter point has not gone entirely unnoticed. Paul Milgrom and John Roberts (1988, p.450), two of the leaders in the formalist branch of the post-Coase literature, made the following prediction almost a decade ago. “The incentive based transaction costs theory has been made to carry too much of the weight of explanation in the theory of organizations. We expect competing and complementary theories to emerge – theories that are founded on economizing on bounded rationality and that pay more attention to changing technology and to evolutionary considerations.” We claim that these theories are now emerging.

² We are not alone in this view. In his recent review essay on Milgrom and Roberts (1992), a widely cited textbook treatment of the modern economics of organization, Brian Loasby notes that, “despite their ready acknowledgement of [Alfred] Chandler’s work, Milgrom and Roberts prefer the transaction as the unit of analysis, and do not enquire into the productive activities which a firm undertakes. The final chapter, of only ten pages, skims over technical change, team production, the creation of capabilities and organizational entrepreneurship” (Loasby 1995, p. 475).
Thus, we will document the development of a corpus of promising theories of the firm – here called generically “the capabilities view” – that are more conscious of the character and limitations of knowledge on the production side than is the mainstream post-Coase literature. These theories, we argue, have distinct implications for economic organization – implications that are not easily reached within the confines of the mainstream literature on the economics of organization.

Admittedly, the emerging capabilities view is even more heterogeneous than the post-Coase literature, partly because of its diverse backgrounds in business history and strategy, evolutionary economics, and technology studies (more on this in section IV). However, we believe that it is possible to reconstruct one of the central concerns of this body of literature in terms of a revitalized attention to the importance of production costs — now recast in a new way — for understanding the problem of economic organization. One of our important goals here is to bring the capabilities view more centrally into the ken of economists. We offer it not as a finely honed theory but as a developing area of research whose potential remains relatively untapped. Moreover, we present the capabilities view not as an alternative to the transaction-cost approach but as a complementary area of research.

---

3 We also think of the capabilities perspective as in many ways a return to common sense notions that many economists have had all along. In other words, while the capabilities perspective has not been a part of “formal” economics, it has been a part of the economist’s more “appreciative” theorizing. (On the distinction between formal and appreciative theory, see Nelson and Winter (1982, p.46).)
II. Production and Governance: the Post-Coase Literature

As we will argue in more detail below, there are in fact two principal theoretical avenues closed off by a conception of organization as the solution to a problem of incentive alignment. And both have to do with the question of production knowledge. One is the possibility that knowledge about how to produce is imperfect — or, as we would prefer to say, dispersed, bounded, sticky and idiosyncratic. The second is the possibility that knowledge about how to link together one person’s (or organization’s) productive knowledge with that of another is also imperfect. The first possibility leads us to the issue of capabilities or competences; the second leads to the issue of qualitative coordination.

Although Coase may have put aside the issue of capabilities, he did not neglect the issue of coordination. In the 1937 article, he lists several sources of those “costs of using the price mechanism” that give rise to the institution of the firm. In part, these are the costs of writing contracts. The “most obvious cost of ‘organising’ production through the price mechanism 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. These costs can be avoided by firm organization. However, a careful reading of the paper suggests that it is ultimately a quite different type of contracting cost that attracts Coase’s attention. After pointing out that the nature of the firm consists largely in substituting an employment contract for a spot contract in output, Coase

---

4 As Herbert Simon (1957) explains the employment relation, the capitalist pays a wage for the right to choose which action \( x \in \Omega \) the worker will perform at any time, where \( \Omega \) is the “job description” or set of allowable actions to which the worker agrees.
suggests that the real costs of contracts may lie in their inflexibility. “It may be desired to make a long-term contract for the supply of some article or service,” he 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.)

A close reading of this passage suggests that Coase’s explanation for the emergence of the firm is ultimately a coordination one: the firm is an institution that lowers the costs of qualitative coordination in a world of uncertainty.
Largely in a quest to make Coase’s ideas more “operational,” the literature has arguably both narrowed his explanation for the firm and moved its focus away from issues of coordination, especially qualitative coordination. More precisely, both the issue of capabilities and the issue of the coordination of production — in the sense of aligning the knowledge and expectations of the parties who need to cooperate in production — have been overshadowed by a dominant interest in issues of incentive compatibility.

Oliver Williamson, the flagbearer of the field since the 1970s, certainly cannot be accused of having a narrow conception of transaction-cost economics. But, in a manner far more explicit than Coase, he has upheld the partition between transaction costs and production costs. This he argues as a pragmatic methodological postulate: hold production costs constant and look only at transaction costs. “A useful strategy for explicating the decision to integrate,” he says, “is to hold technology constant across alternative modes of organization and to neutralize obvious sources of differential economic benefit” (Williamson 1985, p. 88). This may indeed be a sensible starting point, so long as it is not an ending point.

In Williamson’s early work (particularly Williamson 1975), issues of coordination figured prominently. For example, in an echo of the passage from Coase cited above, Williamson argued that internal organization may be a superior mode of coordination whenever boundedly rational transactors confront uncertainty.

---

5 This type of coordination was strongly emphasized by Harold Malmgren (1961) in what is arguably the first “operationalization” of Coase (1937). Langlois and Cosgel (1993) argue that this was also ultimately Frank Knight’s explanation of the firm.
If, in consideration of these [cognitive] limits, it is very costly or impossible to identify future contingencies and specify, ex ante, appropriate adaptations thereto, long-term contracts may be supplanted by internal organization. Recourse to the latter permits adaptations to uncertainty to be accomplished by administrative processes in a sequential fashion. Thus, rather than attempt to anticipate all possible contingencies from the outset, the future is permitted to unfold. Internal organization in this way economizes on the bounded rationality attributes of decision makers in circumstances in which prices are not “sufficient statistics” and uncertainty is substantial.  

(Williamson 1975, p. 9.)

But Williamson’s interest in coordination appears to have declined over time in favor of a greater preoccupation with incentive issues. Along with Klein, Crawford, and Alchian (1978), Williamson (1985) focused in on what has become perhaps the central concept in the present-day economics of organization: asset specificity. It is a concept that has virtually come to crowd out all others in the explanatory pantheon. “The main factor to which transaction-cost economics appeals to explain vertical integration,” he now believes, “is asset specificity” (Williamson 1986, p. 189).

The logic is basically simple. Assets are highly specific when they have value within the context of a particular transaction but have relatively little value

---

6 What Williamson here means by prices not being “sufficient statistics” — a reference to his interpretation of Hayek (1945) on the virtues of the price system — is that internal organization may be superior in situations requiring qualitative coordination, that is, the transmission and use of information beyond price and quantity.
outside the transaction. This opens the door to opportunism. Once the contract is signed and the assets deployed, one of the parties may threaten to pull out of the arrangement — thereby reducing the value of the specific assets — unless a greater share of the quasi-rents of joint production find their way into the threat-maker’s pockets. Fear of such “hold up” ex post will affect investment choices ex ante. In the absence of appropriate contractual safeguards, the transacting parties may choose less specific — and therefore less specialized and less productive — technology. If, by contrast, the transacting parties were to pool their capital into a single enterprise in whose profits they jointly shared, the incentives for unproductive rent-seeking would be attenuated. And, because such unified organizations would choose the more productive specialized technology, they would win out in the competitive struggle against the contractual alternative.\(^7\)

The explanation from asset specificity is at base an argument about the alignment of incentives, even if it ultimately rests on imperfect information. In a world of certainty and unrestricted cognitive ability (if one could imagine such a place), it would be easy to write and enforce long-term contracts that preempt ex ante unproductive rent-seeking behavior ex post and thus obviate internalization. This insight, indeed, has inspired one important formal strand of the literature.

The work of Oliver Hart and others (Grossman and Hart 1986; Hart 1995; Moore 1992) — called the incomplete-contracts literature — distinguishes two types of rights under contract: specific rights and residual rights. The latter are

---

\(^7\) For example, a hostage. See Williamson (1985, chapters 7 and 8).

\(^8\) This way of putting it gives an explicitly evolutionary spin to the functionalist argument more typical in transaction-cost economics. On this see Langlois (1984, 1986).
generic rights to make production decisions in circumstances not spelled out in the contract. In this literature, the choice between contract and internal organization reduces to a question of the efficient allocation of the residual rights of control when contracts are incomplete and assets highly specific. Suppose there are two parties cooperating in production, each bringing to the arrangement a bundle of assets. If none of the assets is highly specific, opportunism is impossible ceteris paribus, as either party can liquidate at no or low cost as soon as troublesome unforeseen contingencies arise. If, however, assets are specific, or if opportunism becomes possible for other reasons, it may be efficient to place the residual rights of control in the hands of only one of the parties by giving that party ownership of both sets of assets. In general, the owner ought to be the party whose possession of the residual right minimizes rent-seeking costs, which typically means the party whose contribution to the quasirents of cooperation is greater.

This is all well and good as far as it goes, which, in some respects, is not nearly as far as the mainstream economics of organization seems to think. The emphasis in the literature on misaligned incentives obscures, in our view, the fundamental role that institutions (including the firm) play in qualitative coordination, that is, in helping cooperating parties to align not their incentives

---

9 Hart and his colleagues hold that the possession of the residual rights of control necessitates ownership of the firm’s capital assets, whether tangible or intangible. This allows them to do something few in the literature have been able to do: to define the boundaries of the firm crisply and consistently. For them, a firm is defined by the bundle of assets under common ownership. (This stands in contrast to the “nexus of contracts” view, which sees the firm as a far more fuzzy notion, and to the related principal/agent theory, in which it is not possible to assign alternative contractual arrangements to specific organizational structures: a contract between employer and employee is not necessarily different from a contract between a firm and its supplier).
but their knowledge and expectations. All recognize that knowledge is imperfect and that most economically interesting contracts are, as a consequence, incomplete. But most of the literature considers seriously as coordinating devices only contracts and the incentives they embody. It thus neglects the role—the potentially far more important role—of routines and capabilities as coordinating devices. Moreover, the assumption that production costs are distinct from transaction costs and that production costs can and should always be held constant obscures the way productive knowledge is generated and transmitted in the economy.

A striking example of this incentive-oriented research strategy can be found in a recent paper by Rotemberg and Saloner (1994). They address one of the key ideas of the corporate strategy and capabilities literature, namely, that firms may be best off choosing narrow strategies. Specifically, Rotemberg and Saloner use the incomplete-contracts framework to argue that a firm may choose a narrow strategy (and thus ignore profitable opportunities) because strategic breadth leads to implementation problems ex post that distort ex ante incentives. They do note (p. 1131) that “increasing returns to specialization” (because of learning advantages from concentrating on well-defined capabilities) may be an independent reason for narrow strategies, but they do not investigate that possibility. The problem is not that such reformulations in terms of incentives are internally inconsistent. Rather, the issue is whether the mechanisms so identified are in fact plausible explanations of the phenomena under study, a question that economists do not typically feel required to pose let alone answer. In fact, it is quite likely that the mechanisms underneath narrow firm strategies
have little or nothing to do with the alignment of incentives, and have everything to do with limited knowledge and capabilities.

More generally, we are worried that conceptualizing all problems of economic organization as problems of aligning incentives not only misrepresents important phenomena but also hinders understanding other phenomena, such as the role of production costs in determining the boundaries of the firm. As we will argue, in fact, it may well pay off intellectually to pursue a research strategy that is essentially the flip-side of the coin, namely to assume that all incentive problems can be eliminated by assumption and concentrate on coordination (including communication) and production-cost issues only.

III. Production Costs Redux: Coordination and Capabilities

A. Key Ideas in the Capabilities Perspective

As we have suggested, there is now emerging a research approach that does emphasize issues of qualitative coordination and limited production knowledge.\(^{10}\) We emphasize that talking about “a” or “the” capabilities perspective in any generic sense is very much in the nature of a reconstruction, since there are a number of strands of thought involved. Section IV will try to separate out those strands; for the moment, however, we will present the reconstructed version.

\(^{10}\) There is also a recent mainstream development that pursues a research strategy that is similar to this on the overall level, namely the attempt to conceptualize, on the basis of team-theory, the firm as a communication network (Bolton and Dewatripont 1994). We discuss the relation of this work to the capabilities perspective later.
What may make it increasingly appropriate to speak of a capabilities perspective is that a small but growing list of authors has begun self-consciously referring to their work as lying within the confines of a “capabilities,” “dynamic capabilities,” or “competence” approach (Langlois 1992; Langlois and Robertson 1995; Kogut and Zander 1992; Foss 1993; Dosi and Marengo 1994; Teece and Pisano 1994). These contributions take somewhat different starting points. Thus, some begin from bounded rationality and other aspects of cognition and build up a theory of firm-specific knowledge—that is, capabilities—from this (e.g., Kogut and Zander 1992; Dosi and Marengo 1994), while others begin from the empirical generalization that productive knowledge is neither explicit nor freely transferable (e.g., Langlois 1992). Either way it boils down to the same common-sense recognition, namely that individuals—and organizations—are necessarily limited in what they know how to do well. Indeed, the main interest of the capabilities view is to understand what is distinctive about firms as unitary, historical organizations of cooperating individuals. Moreover, it is becoming an increasingly widespread recognition among contributors to the capabilities view that approaching the firm in this way has fertile implications not only for understanding the sources of firm heterogeneity, competitive advantage, and differential rents (Lippman and Rumelt 1982; Wernerfelt 1984) but also for advancing the economics of organization.

Michael Polanyi (1958) has taught us that knowledge is not all of a form that can be articulated in words or pictures for easy transmission. Much

---

11 This was clearly the position of those classical economists, particularly Smith, who wrote on specialization. For a reading of the capabilities perspective as the modern heir to the classical theory of production, see Foss (1996c).
knowledge — including, importantly, much knowledge about production — is tacit and can be acquired only through a time-consuming process of learning by doing. Moreover, knowledge about production is often essentially distributed knowledge, that is to say, knowledge that is only mobilized in the context of carrying out a multi-person productive task; is not possessed by any single agent, and normally requires some sort of qualitative coordination, for example, through direction and command, for its efficient use. Indeed, capabilities are precisely characterized by these features: they may be seen as team-embodied and partly tacit production and organization knowledge that can be employed by team-members for a strategic purpose.

In a world of tacit and distributed knowledge — that is, of differential capabilities — having the same blueprints as one’s competitors is unlikely to translate into having the same costs of production. Generally, in such a world, firms will not confront the same production costs for the same type of productive activity. Moreover, the costs that can make transacting difficult — the costs that may lead to internalization or various other business institutions — may go beyond those that arise in the course of safeguarding against opportunism or damping moral hazard through monitoring or incentive contracts. In such a world, economic activity may be afflicted with “dynamic transaction costs,” the costs that arise in real time in the process of acquiring

---

12 Of course, not all distributed knowledge requires conscious direction for its efficient utilization; in fact, it is a standard argument in favor of the market order that it better utilizes distributed knowledge than any known directed order (Hayek 1945). However, as we shall later argue, firms may derive part of their raison d’être from their (sometimes) superior abilities to coordinate (some) types of distributed knowledge. Thus, what we are after is a knowledge-based theory of the existence and boundaries of the firm.

13 Loosely, and perhaps somewhat cryptically, dynamic transaction costs are the costs of not having the capabilities you need when you need them (Langlois 1992).
and coordinating productive knowledge (Langlois 1992; Langlois and Robertson 1995) and which are different in nature from the transaction costs that are caused by problems of aligning incentives. This, in turn, implies that the capabilities may be interpreted as a distinct theory of economic organization.

B. The Capabilities Perspective as a Theory of Economic Organization

A key implication of the capabilities perspective as it relates to economic organization is that, in the terminology of G. B. Richardson (1972), the structure of complementarity and similarity among the various capabilities in the economy affects the pattern of organization (including the firm-market boundary) in ways not fully explicable in terms of the costs of transacting. Indeed, the ability to transact (and therefore the cost of transacting) is itself a capability (Winter 1988), which suggests a blurring of the boundary between production and exchange.

The idea that capabilities may be an independent causal factor behind the pattern of economic organization has recently received support from the doyen of business historians, Alfred D. Chandler.¹⁴ He traces the neglect of production in the post-Coase literature to its choice of the isolated transaction as unit of analysis. By contrast, “if the firm is the unit of analysis, instead of the transaction,” Chandler says, “then the specific nature of the firm's facilities and skills becomes the

¹⁴ This may be contrasted with Chandler’s earlier support for Williamson’s brand of transaction cost economics. Chandler (1992, p.85) says that although he has “learned much from Williamson,” there is a basic difference between them which has to do with the unit of analysis. Chandler goes on to endorse “the recently formulated evolutionary theory of the firm,” which is roughly identical to the capabilities perspective.
most significant factor in determining what will be done in the firm and what by the market” (Chandler 1992, p.86).

Even more striking, Ronald Coase (1990, p.11) has himself voiced similar views.

While transaction cost considerations undoubtedly explain why firms come into existence, once most production is carried out within firms and most transactions are firm-firm transactions and not factor-factor transactions, the level of transaction costs will be greatly reduced and the dominant factor determining the institutional structure of production will in general no longer be transaction costs but the relative costs of different firms in organizing particular activities.

However, it was G. B. Richardson who introduced the term “capabilities” to talk about the necessarily limited range of productive knowledge firms and individuals possess. Taking issue with the representation of knowledge in the production-function approach, Richardson writes:

Of course I realise that production functions presume a certain level of managerial and material technology. The point is not that production is thus dependent on the state of the arts but that it has

\[\text{15} \]

Of course, taking the firm as the unit of analysis makes it difficult to study the rationale for and the boundaries of the firm. Langlois and Robertson (1995) suggest taking capabilities and routines as the fundamental units of analysis. This has the benefit of placing the economics or organization more firmly within the structure of the New Institutional Economics more broadly (Langlois 1986), in which norms and conventions — which, like routines, are rule-based guides to action — are the fundamental concepts.
to be undertaken (as Mrs. Penrose\textsuperscript{16} has so very well explained) by organisations embodying specifically appropriate experience and skill. It is this circumstance that formal production theory tends to put out of focus, and justifiably, no doubt, given the character of the optimisation problem that it is designed to handle; nevertheless, it seems to me that we cannot hope to construct an adequate theory of industrial organization and in particular to answer our question about the division of labour between firm and market, unless the elements of organisation, knowledge, experience and skills are brought back to the foreground of our vision. (Richardson 1972, p. 888).

In Richardson’s terminology, production can be broken down into various stages or activities. Some activities are similar, in that they draw on the same general capabilities. Activities can also be complementary (in both a technical and an economic sense) in that they are connected in the chain of production and therefore need to be coordinated with one another. Juxtaposing different degrees of similarity against different degrees of complementarity produces a matrix that maps different types of economic organization. For example, closely complementary and similar activities may be best undertaken under unified governance.

Complementarity is clearly an increasingly important theme in today’s economics of organization (Milgrom and Roberts 1990); indeed, there is a widespread recognition that “strongly complementary assets should be brought

\textsuperscript{16} On whom see below.
under common ownership” (Milgrom and Roberts 1992, p. 312). But the real force of Richardson’s argument is in quite a different direction. In Richardson, the import of the concept of capabilities was their limitations. Because of what are effectively cognitive constraints, all organizations must specialize; and, since the chain of production in an advanced economy requires a diversity of very different capabilities, the costs of integrating across many links in that chain are necessarily high, and firms must rely on various kinds of market and hybrid arrangements to coordinate their activities even in the face of contractual hazards.\textsuperscript{17} Although transaction costs may outweigh the costs of dissimilarity in the case of some closely complementary activities, on the whole the limitations of capabilities outweigh transaction costs. As Brian Loasby (1991) has observed, Richardson thus stands on its head a principal, albeit tacit, presumption of transaction-cost economics, namely that, because contractual relationships among firms are fraught with hazards, integration must on the whole be relatively less costly and thus widely desirable.

Richardson’s insight is a simple but extremely profound one. For it suggests that – as a quite general matter – capabilities are determinants of the boundaries of the firm, since they determine, in Coase’s words, “the relative costs

\textsuperscript{17} A related, if not identical, position has been adopted by David Teece (1982, 1986), one of the few major scholars to have incorporated Richardson’s ideas. Unlike Richardson, who discusses the coordination of complementary activities, Teece talks about complementary assets that might be cospecialized to one another. As with Richardson’s closely complementary activities, cospecialized assets may be difficult to coordinate. But, unlike Richardson, Teece is inclined, with the broader incentive-based asset-specificity literature that has influenced him, to believe that cospecialized assets may be a cause of integration more than of cooperation, especially to the extent that integration allows an innovator to appropriate the gains from innovation in regimes in which intellectual property rights are ineffective. Thus, for Teece governance structures alternative to the market arise to prevent slippery innovative knowledge from escaping the grasp of its creators, just as, in the main current of the transaction-cost literature, alternative governance structures emerge to protect transactors from the “plasticity” of contract.
of different firms in organizing particular activities.” Problems of economic organization may crucially reflect the possibility that a firm may control production knowledge that is, in important dimensions, strongly different from what others control. Thus members of one firm may quite literally not understand what another firm wants from them (for example, in supplier contracts) or is offering them (for example, in license contracts). Because of the extreme specificity and tacitness of much productive knowledge, one firm may have difficulties understanding another firm’s capabilities; and both firms separately and together may know more than their contracts can tell (Kogut and Zander 1992; Winter 1993). In this setting, the costs of making contacts with potential partners, of educating potential licensees and franchisees, of teaching suppliers what it is one needs from them, etc., become very real factors determining where the boundaries of firms will be placed.

Note that these dynamic transaction costs are in a different category from the transaction costs usually considered in the post-Coase literature. Transacting difficulties are not a matter of incentive problems within an otherwise well-defined and well-understood exchange context. Rather, coordination problems may arise because capabilities exhibit too much “friction”: the knowledge, skills, and traditions embodied in existing governance structures (be they firms, markets, or in between) may be too inflexible, especially in the face of major “Schumpeterian” change, to seize market and technological opportunities. In such circumstances, other governance structures that can muster the necessary capabilities may arise and prosper.

Morris Silver (1984) has suggested, for example, that much vertical integration arises not when firms venture into areas of similar capabilities but
when firms are dragged, kicking and screaming, as it were, into complementary but dissimilar activities because only in that way can they bring about a profitable reconfiguration of production or distribution. Langlois and Robertson (1995) build a broad theory of industrial dynamics around this idea. The organizational question is whether new capabilities are best acquired through the market, through internal learning, or through some hybrid organizational form. And the answer will depend on (A) the already-existing structure of capabilities and (B) the nature of the economic change involved.

If a profit opportunity requires a configuration of capabilities different from what already exists in the economy, then a Schumpeterian process of creative destruction may be set in motion. If the old configuration of capabilities is decentralized into what we may loosely call markets, then a reorganization within a single organization — vertical integration — may most cheaply bring about the necessary redeployment. If, by contrast, the old configuration of capabilities lies within large vertically integrated organizations, creative destruction may well take the form of markets superseding firms. History offers many examples of both.

The organizational possibilities are tempered by the nature of the reconfiguration required. If change is systemic — if it requires simultaneous change in many parts of a complex system — internal organization may prove less costly ceteris paribus. If, however, change is autonomous — if change can take place in separate subsystems without greatly affecting the way those subsystems
are connected together — then markets, which can take advantage of specialized and decentralized knowledge, may be at a relative advantage.\textsuperscript{18}

The upshot of all this, we suggest, is that there now exists a distinct basis – a collection of ideas, concepts and mechanisms – for the capabilities view as a theory of economic organization, at least with respect to the boundaries of the firm. But what about the empirical evidence? Writers like Chandler (1962, 1977, 1990, 1992), Lazonick (1991), and Langlois and Robertson (1995) enlist economic and business history in support of a capabilities view. But more quantitative empirical studies also suggest that differential capabilities, and therefore production costs, are significant variables for explaining the boundaries of the firm. In Walker and Weber's (1984) empirical study of the make-or-buy decision, the most important explanatory variable turned out to be the indicator for differential firm capabilities, that is, for production costs. And, in a study by Monteverde and Teece (1982), which set out to find support for the standard contractual approach, the most significant variable was actually the dummy for the firm, reflecting heterogenous and unobserved firm effects (Kogut and Zander 1992, p. 394).

\textbf{IV. Capabilities and Contemporary Research Themes}

In this section, we discuss how the capabilities perspective relates to a variety of streams of thought originating from a variety of problems. Just as the post-

\textsuperscript{18} The terms “systemic” and “autonomous” are from Teece (1986). In the case of autonomous innovation, the issue of standards enters the picture: for standards are typically ways of fixing the connections among subsystems so that change is channeled in autonomous directions. Langlois and Robertson (1992, 1995) call this kind of structure a modular system.
Coase literature on the economics of organization has not been developed in an intellectual vacuum, so the capabilities perspective very much reflects a number of diverse influences. And, although the roots of the capabilities approach stretch quite far back in history, and although – until recently – the various influences have been developed independently, the emergence of the capabilities perspective seems in fact to mark a growing realization of the possibilities of convergence. Thus, we wish to provide a brief conspectus of the sources of the capabilities approach; of the many intellectual alliances that this perspective has struck or may strike; and of the work of a number of researchers who have been involved in developing capabilities insights.

What we have said so far about the post-Coase literature on the economics of organization also makes it obvious and necessary to undertake a more detailed discussion of how the capabilities perspective relates to this literature. But there are numerous other connections: to strategic management and organizational learning; to business history; to the economics of technology; to evolutionary economics; and to the economics of institutions.

A. Antecedent and Related Fields

How can firms make best use of their distinctive capabilities? How have they done this in the past? And how can they go on developing new valuable capabilities? Such questions have been central in the strategy field since its inception at the end of the 1950s, and in the related field of business history, at least since Alfred D. Chandler’s (1962) demonstration of the importance of organizational capabilities to the restructuring of the American economy that
began in the middle of the last century. This three-decades-long interest in capabilities should be contrasted with the lack of interest shown by economists, at least until recently.

The conceptualization of the firm that underlies this work was perhaps best expressed in the late Edith Penrose’s *The Theory of the Growth of the Firm* (1959), a conceptualization she explicitly differentiated from the prevailing production-function view. “The firm,” Penrose says, is “a collection of productive resources the disposal of which between different uses and over time is determined by administrative decision” (Penrose 1959, p. 24). Now, resources in Penrose’s view yield services, and it is these services – clearly a theoretical precursor to the concept of capabilities – that interest her the most. Because resources/services become specialized to firms – and mesh with each other in a team-like manner – they are worth more to the firm than to the market (meaning other firms). They therefore yield quasi-rents, some of which may be appropriated by the firm’s owners. Moreover, although resources/services are firm-specific, they are nevertheless somewhat “fungible” inside the firm, and, when in excess, provide a stepping-stone for diversifying to new markets.

Penrose’s work helped define at least three distinct areas of research. The first one partially stems from her insistence that specialized resources/services yield rents; this has helped found what is today referred to as the resource-based perspective in contemporary firm-strategy research (Lippman and Rumelt 1982; Wernerfelt 1984). The primary contribution associated with the resource-based perspective is a thorough analysis of the conditions under which resources yield rents. Thus, heterogeneous, immobile and hard-to-imitate resources that are,
moreover, acquired in imperfect factor markets (so that there may be a difference between the price of the resource and its value to an acquiring firm) are rent-yielding strategic assets to firms.

A second relevant area of research is the study of diversification (e.g., Teece 1982; Dosi, Teece, and Winter 1992), where Penrose’s notion of excess capabilities, combined with transaction-cost considerations, is perhaps the dominant mode of explanation (Montgomery 1994). Roughly speaking, the story is this. As firms carry on their normal business, they are likely to accumulate excess resources, for example, excess managerial capabilities. In principle, rents from these resources may be captured in different ways, for example, through market exchange, long-term contracts, or in-house use. Because of transaction-cost problems, which may be particularly severe when the excess resources involved are knowledge resources, in-house use is more efficient, and the firm will accordingly apply the resources that are in excess to neighboring markets.

The third area of research that Penrose’s work helped to establish is the study of organizational learning, which also owes a heavy debt to such seminal contributions to organization theory as March and Simon (1958) and Cyert and March (1963). Penrose argued that the management team holds images of the external environment and of the firm’s internal resources; that these images are produced through internal learning processes; and that they determine “the productive opportunity set” of the firm, that is, the productive possibilities that that the firm’s “entrepreneurs’ see and can take advantage of” (Penrose 1959, p.31).

The idea of the image as a shared firm-specific vision, and the implication that firms are in essence cognitive communities, is more radical than the – now
more standard – ideas of bounded rationality and tacit knowledge in action or of firms as essentially processors of (objective) information (Fransman 1994). And, with the possible exception of Arrow (1974), Loasby (1976), and Crémer (1990), few economists have shown interest in Penrose’s idea, although it is standard in contemporary organization theory (James March 1988). The image is more radical because it explicitly recognizes that agents have to make sense of their world, that agents' cognitive development is molded in social processes, and because it implies that tacitness is an aspect of virtually all acts of interpretation and meaning attribution (Marengo, 1995). In this view, the essence of decision making is not making a choice among pre-given alternatives; it is a matter of construing something resembling a decision situation by defining which variables are relevant, which in turn requires making sense of the environment, setting up procedures for solving the problem, etc.

Clearly, such a view of decision-making emphasizes the importance of coordination problems, but also points to a coordinative role for capabilities. As Langlois (1984) and Marengo (1995) argue, if agents entering the firm held the completely same habits of thought/models of the world, the only obstacle to efficient coordination of their actions would be precisely the sort of incentive problems that preoccupy modern organizational economists. However, in a world in which agents do not share exactly the same models and do not know

---

19 Langlois (1996) has argued that the set of capabilities available to an organization, and the way in which those capabilities are arranged, constitute the organization's cognitive structure, that is, its mechanism for perceiving technological and market opportunities. Thus, knowledge within the organization is perhaps even more widely distributed and variegated that the notion of an “image” suggests, and is not always confined to management.
each others' models, a collective knowledge base is required for coordination (Crémer 1990).

As simulations built on the theory of classifier systems demonstrate (Holland and Miller 1991), such a knowledge base may develop as a result of organizational learning under rather general assumptions (Marengo 1995). These attempts to construct a theory of capabilities from ideas about behavior founded in organization theory rather than in maximizing behavior currently enjoys much attention among evolutionary economists, particularly among Italian and French ones (e.g., Dosi and Marengo 1994). The basic idea, however, can be found in Nelson and Winter’s An Evolutionary Theory of Economic Change (1982), particularly chapter 4 and 5. Their widely cited analysis of “routines” builds directly on behavioralist organization theory, as well as on Michael Polanyi’s (1958) analysis of tacit knowledge. In fact, this analysis is itself an important precursor of modern work on capabilities.

In the evolutionary economics of Nelson and Winter (1982) and many (other) economists associated with the International Joseph A. Schumpeter Society, the capabilities view of the firm serves primarily as a micro-foundation for population-level analysis of industry and technology evolution. Thus, the capabilities perspective helps rationalize the variety of behaviors – including innovative behaviors – that are necessary in any evolutionary account of industry and technology evolution (Metcalfe 1989); it is an explanatory component in a broader explanation, much like the way in which the neoclassical theory of the firm is basically an explanatory component in standard price-theory.
However, there has been some important work on innovation and technological change that puts the emphasis on the firm level rather than the population/industry level. Much of this has been associated with University of Sussex Science Policy Research Unit (SPRU) and with members of the (University of) Reading school of international business, particularly John Cantwell (1994). For example, recent influential work by Keith Pavitt and Pari Patel (Patel and Pavitt 1994) use systematic information on American patenting by more than 400 of the world’s largest technologically active firms to demonstrate that the accumulation of technological capabilities is strongly path-dependent and that there are therefore severe limits on the range of exploitable technological opportunities. Moreover, they argue that technological capabilities “give a convincing empirical explanation of the boundaries (or − perhaps better − the core activities) of firms (p. 2). In other words, like many other proponents, Pavitt and Patel see the capabilities perspective as − at least potentially − an alternative theory of economic organization. We treat this issue in further detail in the next section.

B. The Capabilities Perspective and the Modern Economics of Organization

We have interpreted the capabilities perspective as reaching for a distinct theory of economic organization, one that is based on a conceptualization of the firm as a repository of productive knowledge with certain non-standard characteristics, what we have here called “capabilities.” In this story, incentive issues are suppressed in favor of a focus on problems of coordinating knowledge and expectations. We have chided the profession for its lopsided choice of the
opposite approach and for its dramatic overemphasis on transaction costs and incentive alignment, to the exclusion of production costs and issues of coordination, in explaining economic organization.

However, there has recently been some stimulating work that explicitly focuses on the coordination of knowledge and expectations in a team-theoretic framework (Arrow 1985; Crémer 1990; Radner 1992, 1996; Bolton and Dewatripont 1994). In these models, incentives move into the background. Building on earlier ideas in Marschak and Radner (1972) and Arrow (1974), these writers view the firm as a communication network that is designed to minimize both the cost of processing new information and the costs of communicating this information among agents. Communication is costly because it takes time for agents to absorb new information sent by others, but this time may be reduced by specializing in the processing of particular types of information. In Bolton and Dewatripont’s (1994) model, for example, each agent handles a particular type of information, and the different types of information are aggregated through the communication network. When the benefits to specializing outweigh the costs of communication, teams (firms) arise.

Arguably, such work captures some of the main ideas of the capabilities perspective as we have interpreted it; for example, there is an emphasis on the need for qualitative coordination, on specialization in handling knowledge, on firm-specific “codes” of communication (Arrow 1974), and on bounded rationality (Radner 1996). We conjecture that this work will become increasingly important as first steps towards the formalization of capabilities ideas.
In spite of this conjecture, one should not reject the more standard incentive-oriented work as a natural complement to the capabilities view. In fact, future work may center around modelling capabilities and incentive considerations in the same model,\textsuperscript{20} so that, for example, the role of both production costs and transaction costs in determining the boundaries of the firm becomes more visible than it is in the post-Coase literature on the economics of organization. In this respect, it is noteworthy that Williamson's primary design principle for efficient economic organization has been changed to reflect capabilities considerations: “Align transactions, which differ in their attributes, with governance structures, which differ in their costs and competencies in a discriminating (mainly, transaction cost economizing) way” (Williamson 1991, p. 79). Thus, Williamson now thinks of competencies (i.e., capabilities) as determinants of governance more-or-less on a par with transaction costs. In other words, the notion of the firm as a bundle of capabilities may harmonize with key ideas of the post-Coase literature. An excellent specific example is a model by Lewis and Sappington (1991). They analyze the firm's make-and-buy decision under the assumption that its subcontractor is known to have lower innate production costs (i.e., superior capabilities) but the firm is better able to monitor and control its own production activities. Lewis and Sappington perform various comparative-static exercises in this setting; for example, they examine how the firm's boundaries choice varies with technological change that influences production costs and monitoring,

\textsuperscript{20} Promising recent work by Aghion and Tirole (1995) incorporates both incentive considerations and information-processing considerations that are akin to the thrust of the capabilities view.
so that both incentive and capabilities considerations are allowed to enter the picture.

In the following, we briefly present a few further suggestions as to how, more specifically, key ideas from the two perspectives may be aligned.\textsuperscript{21} These suggestions keep intact the basic idea that economic organization is first and foremost a matter of efficiently aligning incentives; capabilities considerations merely serve to help extend the applicability of this basic idea. This is an interesting and legitimate research strategy, as long as we do not forget to also consider the other side of the coin: that capabilities considerations may be primary and incentive considerations secondary.

Capabilities and intra-firm agency problems. The argument here is that capabilities in firms may influence the outcomes of principal-agent-type problems: firms will often be characterized by a distinct "way of doing things" that is coded in its capabilities and is shared among input-owners. Precisely because it is shared (common), the presence of such knowledge may serve to mitigate moral hazard and adverse-selection problems.\textsuperscript{22} This is a possible interpretation of why corporate cultures may be valuable assets to firms.\textsuperscript{23}

Moreover, such an capabilities-cum-agency problems story also helps to rationalize the scope of firms: casual empiricism confirms that few firms have integrated the entire value-chain and that no firm has a stake in every product

\textsuperscript{21} For a fuller discussion of the issues involved here, see Foss (1996b).

\textsuperscript{22} There is, however, also a conflict between the agency view and the capabilities perspectives. In the first, heterogeneity of knowledge, preferences and behaviors is problematic because it causes agency problems; in the latter, it is beneficial, because it stimulates organizational learning and the development of capabilities.

\textsuperscript{23} See Kreps (1990) for a slightly different interpretation
market in the economy, the common explanation being that firms confront increasingly dissimilar capabilities as they move away from their core business (Richardson 1972). A rendering along the lines of the modern economics of organization may be: as firms move increasingly away from their core businesses, they confront increasing adverse-selection and moral-hazard problems, since management becomes increasingly unable efficiently to monitor employees or to evaluate their human capital. Agency costs rise correspondingly, producing the net profitability disadvantage associated with further integration (for a similar story, see Aghion and Tirole 1995).

Capabilities and Incomplete Contracts. In the presence of incomplete contracts and bounded rationality, something more than an allocation of rights is required to structure intra-firm interaction; firms aren’t held together solely by the thin glue of transaction-cost minimization, but rather by the thicker glue of capabilities. A key aspect of the capabilities critique of the modern economics of organization is that it too strictly dichomotizes production and organization/exchange considerations when in reality these are closely intertwined. Since the very notion of firm capability combines production and organization considerations, it is entirely likely that capabilities embodying knowledge about production at least to some extent also help solve problems of rent-seeking inside organizations.

Asset specificity and capabilities. As we have argued, the notion of specific assets is key to the modern economics of organization (but see Demsetz 1988). Not surprisingly, elaborate lists of types of specific assets have been constructed,
ranging from patents over dedicated physical equipment to site specificity (Williamson 1985; Grossman and Hart 1986).

Capabilities would certainly seem to qualify as specific assets – they are specialized to firms; they have low (or no) value in alternative uses; managers/owners can hold residual rights as to their use, etc. But the modern economics of organization does not normally view them that way (Klein (1988) and Milgrom and Roberts (1992) are exceptions).24 Part of the reason may be that capabilities are hard to treat in formal models. Another part may be that it is harder to reason about who captures rents from capabilities than from ordinary factors of production; the underlying bargaining would seem to be immensely more complicated than the bargaining game being played between the firm and the owner of an ordinary human-capital input. However, these difficulties are not insurmountable in principle, and capabilities deserve a place on the short-list of empirically important specific assets.

V. Conclusion

Our aim in this paper has been to document the importance of the capabilities perspective as an emerging perspective on economic organization. It is characterized by distinct insights, not the least the attempt to restore production and production costs to their rightful place as determinants of the boundaries of the firm, and to find a place for qualitative coordination in the theory of

---

24 This does not automatically mean, however, that capabilities are necessarily and under all circumstances best governed internally. Apart from the issue that a capability may be too dissimilar relative to an organization's other capabilities and is therefore best left to other firms (Richardson 1972), there is the issue that capabilities may exist on the industry-level as Marshallian external economies (see Langlois and Robertson 1995; Foss 1996d).
economic organization. In other words, the capabilities perspective highlights explanatory mechanisms that are different from those of the post-Coase literature on economic organization. Since the two perspectives may be read as addressing the same sort of phenomena – notably the existence, boundaries and internal organization of the firm – and employ theoretical concepts and mechanisms (incentives vs qualitative coordination, blueprint knowledge vs capabilities, etc.), they may therefore be interpreted as being competitors. Although the capabilities view is admittedly less advanced than the post-Coase literature in terms of formalization and terminological stringency, with respect to some important phenomena – notably the boundaries of firms – the capabilities perspective, we have argued, develops more plausible explanatory mechanisms.

However, rather than stressing rivalry, we emphasize the complementarity between the two perspectives and the need for more integrative efforts. Even if it is not currently fashionable among contributors to the capabilities perspective, we feel that there are strong arguments in favor of our position. Both perspectives may benefit from the ideas and insights of the other. There is important mainstream work that, if in no way identical to the capabilities view, nevertheless suggests how aspects of capabilities ideas may be formally approached and modeled (Lippman and Rumelt 1982; Arrow 1985; Bolton and Dewatripont 1994; Aghion and Tirole 1995).

Moreover, when it comes to providing convincing stories about important empirical phenomena, the relations of complementarity between the post-Coase literature and the capabilities view may appear even more striking. For
example, it is arguably hard to provide convincing stories about diversification (Teece 1982; Dosi, Teece and Winter 1992) or the organization of the innovation process (Teece 1986) without relying on both perspectives. For these reasons, the perspectives need to be integrated further. We therefore concur with one of the major scholars in today’s economics of organization when he observes that “[i]n order to fully develop its capabilities, transaction cost economics must be joined with a theory of knowledge and production (Teece 1990, p. 59). And, in fact, the reverse may also hold true, that in order to fully develop its capabilities, the capabilities view must be joined with transaction cost economics. For example, in order to understand the process of emergence and accumulation of capabilities, we need to pay attention to the incentive structure of firms, since this influences investments in human capital. In sum, whether we see it from the perspective of the capabilities perspective or from the perspective of the modern economics of organization, there is an exciting theoretical frontier ahead.
Bibliography


39


