On the Foundations of the Strategic Theory of the Firm:
Should We Rely on Governance, Capabilities, or Both?

Nicolai J. Foss
Department of Industrial Economics and Strategy
Copenhagen Business School
Nansensgade 19,6
DK-1366 Copenhagen K, Denmark
+45 3815 2562 (phone)
+45 3815 2540 (fax)
esnjf@cbs.dk

August 14, 1997

Abstract
This paper is taken up with discussing the extent to which existing economic theories of the firm are helpful for constructing what Rumelt called a “strategic theory of the firm”. Such a theory explains the existence, boundaries, organization and competitive advantage of the firm. The modern economics of organization (e.g., agency theory, transaction cost economics, incomplete contracts theory) is characterized by the heuristic of reducing literally all aspects of economic organization to the alignment of incentives. It thus arguably misrepresents many management and strategy issues. On the other hand, the dominant contender, the capabilities perspective, is lacking in several respects, for example, with respect to its lack of crisp microfoundations, and its inability to predict, explain ownership and explain the existence of the firm. The best way forward in the construction of a strategic theory of the firm is probably to draw eclectically on both approaches.

Acknowledgments
Conversations with, or comments from, Tom Elfring, Kirsten Foss, Richard Langlois, Frans van den Bosch, Charles Baden-Fuller, Henk Volberda, and Mark Baaij are gratefully acknowledged. This paper is a slightly revised
version of “Toward a Strategic Theory of the Firm”, Management Report No. 16 (13), Rotterdam School of Management.
I. Introduction

There can be little doubt that the modern economics of organization\(^1\) – or what is often referred to as “organizational economics” (e.g., Paul Milgrom and John Roberts 1992) – has made a substantial impact in the field of strategic management; indeed, that it “... has become one of the most influential theoretical perspectives in strategy research (Edward Zajac 1992: 76). As a result, many issues of strategic management are now framed as problems of efficient governance. It would also seem to be a reasonable conjecture that this impact is unlikely to diminish in importance in the field for some time to come. However, the modern economics of organization has recently been challenged by an approach – here generically called “the capabilities perspective” – that is seen by (a number of) its proponents as an approach that is competitive to the economics of organization.\(^2\)

The present paper is in the first instance a comparative discussion and evaluation of these two approaches as possible bases for building what Richard

---

1 By “the modern economics of organization” or “organizational economics”, I here have in mind a rather broad menu of theories, such as transaction cost economics (Oliver Williamson 1996), the incomplete contract approach (Oliver Hart, 1995), principal/agent theory (Bengt Holmström 1982), and the nexus of contracts approach (Armen Alchian and Harold Demsetz 1972; Steven Cheung 1983). Unfortunately, terminology is somewhat obscure here, not only with respect to the constituent element of organizational economics, but also with respect to the generic category itself. For example, Barney and Ouchi (1986) call their reader, which also includes the evolutionary theory of the firm, Organizational Economics. Thus, their understanding of the term is broader than the one employed here. Moreover, organizational economics is sometimes also referred to as “the theory of the firm”. However, the latter term is more narrow than the former.

2 In this case, too, a broad menu of theories is involved, including “the evolutionary theory of the firm” (Richard Nelson and Sidney Winter 1982; Luigi Marengo 1995; Ulrich Witt 1996; Nicolai Foss 1997b&c), “the competence perspective” (e.g., Nicolai Foss 1993; Geoff Hodgson 1996), “the capabilities perspective” (Brian Loasby 1991; Richard Langlois and Paul Robertson 1995), “the dynamic capabilities perspective” (David Teece and Gary Pisano 1994), “the resource-based approach” (Birger Wernerfelt 1984), and “the knowledge-based theory of the firm” (J.-C. Spender 1996; Robert Grant 1996). Foss (1997b&c) addresses on methodological level the issue of whether the capabilities/evolutionary perspective is competitive or complementary to the modern economics of organization.
Rumelt (1984) in a classic paper called “a strategic theory of the firm”. By that term, I will here understand a theory of the existence, organization, boundaries and competitive advantage of the firm. Rumelt argued that “… it appears obvious that the study of business strategy must rest on the bedrock foundations of the economists model of the firm”, and explicitly referred to the work of Coase and Williamson as demonstrating that “… economic concepts can model and describe strategic phenomena” (p. 557). He also noted, however, the absence of contact between strategy and the economic theory of the firm.

Much has happened since Rumelt wrote this, and it is perhaps increasingly pertinent to perform a stock-taking. What has happened is not only that the modern economics of organization has met with an impressive rate of acceptance among strategy scholars, and scholars in management studies more broadly (Section II, “Organizational Economics and the Strategy Field: Relevance and Impact”). It is also that the last five years have witnessed the emergence of the capabilities perspective which is claimed by its proponents to be able to successfully address the set of issues that once was the exclusive domain of the modern economics of organization, that is, the existence, boundaries and organization of firms.

At the same time, the capabilities perspective is more close to traditional strategic management concerns. For example, it directly addresses the issue of the sustainability of competitive, arguably the key issue in strategy. Moreover, the capabilities perspective link up with more behavioral perspectives (e.g., Kogut and Zander 1992; Grant 1996). Thus, it is perhaps not difficult to understand the present enthusiasm for the capabilities perspective: it is an approach to economic

---

3 As Charles Baden-Fuller pointed out to me, such an exercise requires that the two approaches are really talking about the same beast, and this may not at all be the case, since a capabilities theorist and an organizational economist may be in disagreement about whether a certain entity is really “a firm”. For example, in terms of knowledge flows, it may be a firm, but not in terms of ownership structure. I will disregard this difficulty, partly because I doubt whether this is a real difficulty, and partly because so far the participants in the debate has not been troubled by it.
organization that avoids (what many would regard as) the worst excesses of the modern economics of organization (e.g., it is opportunism-independent) and is explicitly strategic, while staying in close contact with behavioral perspectives and also more content-oriented strategy perspectives. It would seem to be a more natural candidate to the title of a strategic theory of the firm than the modern economics of organization.

However, as I will argue in this paper, this enthusiasm may be somewhat excessive, because of the heterogeneous and comparatively under-developed nature of the capabilities perspective, and because it is unclear whether the perspective really is able to do what its proponents claim it is (e.g., explaining the existence of firms). Thus, while one may agree with much of the critique of the modern economics of organization from the camp of the capabilities perspective – and I, too, will present such a critique – a balanced perspective requires that we also recognize the limitations of the capabilities perspective. It is this recognition that has arguably been missing so far.

To begin with organizational economics, the critique here is that it is in reality still committed to a “production function view” – in spite of the critique that has (rightly) been launched against this view from prominent organizational economists (such as Williamson, 1996). Specifically, the production function view of the nature of productive knowledge has been kept. One consequence of this is that differential capabilities do not enter the picture as determinants of economic organization. Another consequence is that productive knowledge is assumed to be already coordinated so that the scope for managerial action is narrowly circumscribed. At most, managerial action is a matter of pushing individual effort levels towards their contracted-for levels (given, of course, the costs of doing so) by means of incentives and monitoring, so that the firm will operate on or near the frontier of its production possibility set (Section III, “The Modern Economics of Organization: Contributions and Problems”).
While the emerging capabilities perspective pays much more attention to differential capabilities and the coordination of knowledge as determinants of economic organization, it is also lacking in a number of respects. Compared to the modern economics of organization, it sacrifices generality (e.g., it does not explain the existence of the firm), ability to predict, and explanatory elegance (e.g., its microfoundations are unclear), in favor of greater realism in some dimensions (e.g., differential capabilities) (Section IV, “The Capabilities Perspective”).

Thus, I will criticize both the two dominant approaches to the theory of the firm: the modern economics of organization view for operating with an impoverished conception of production and knowledge; the capabilities perspective for lacking in terms of generality, microfoundations, etc. While I add to the capabilities critique of the modern economics of organization, in contrast to the tenor of the capabilities literature, I also recognize the weaknesses of this literature. Thus, we are dealing here with imperfect alternatives; imperfect, that is, for furthering the construction of a strategic theory of the firm. In fact, I will end by arguing that strategy researchers need ideas from both perspectives; hence, the title of this paper (Section V, “Toward a Strategic Theory of the Firm: Research Strategies”).

II. Organizational Economics and the Strategy Field: 
Relevance and Impact

A.
The recent (or, relatively recent) infatuation with organizational economics in the strategy field may be seen as a fad inspired by the broader movement (if that is the
hard to deny the relevance of a more rational explanation: organizational economics deals with themes that are admittedly crucial to strategic management.

To illustrate, it has recently been argued that the so-called “boundary school” has the potential of playing the role of a synthesizing school in strategic management: it bears promise of being able to organize crucial strategic issues and to draw on complementary insights (Tom Elfring and Henk Volberda 1997). As indicated by its name, the boundary school may be understood as a compilation of theories that are all concerned with analyzing what economists call “the boundaries of the firm”.

In economics, the boundaries of the firm are normally defined in terms of ownership (Hart 1995): if firm A has ownership rights over asset a and firm B does not, asset a is inside the boundaries of firm A and outside the boundaries of firm B. More generally, the boundaries of firm A are defined by those assets that firm A owns. This has implications for the way that the boundaries issue is normally understood in the management studies, namely as the organization of transactions, for assets and transactions are usually bundled.4

Therefore, the boundaries of the firm issue relates to which activities or transactions should be undertaken in firms (hierarchies), which should take place in various intermediate forms (such as franchising, licensing arrangements, long-term supplier contracts, joint-ventures, etc.), and which should be handled in “anonymous” markets. Evidently, these issues are of paramount strategic significance. For example, the issues of diversification, outsourcing, partnering, strategic alliances, virtual corporations, how foreign markets should be serviced, etc. all directly involve the issue of the boundaries of the firm.

4 For example, internalizing a transaction means obtaining ownership rights to the equipment that support the transaction.
From a mainstream strategy perspective (e.g., the resource-based perspective), the boundaries of the firm derive their strategic importance from the at which resources may be acquired, influence the extent to which rents may be appropriated from, for example, valuable knowledge, etc. Thus, in this sustained competitive advantage. This is because knowing something about a firm’s boundaries (and therefore also its contracts), also tells us something about for a firm’s boundaries, including in a wider reading, its relations to, for example, outside suppliers, may in themselves constitute strategic resources, as in the case corporate strategy, and many of business strategy, involve the boundaries of the firm.

The Impact on the Strategy Field
It is hard to disagree with the proposition that since the mid-1980s, the modern field, and indeed on management studies in general. Impressionistic evidence for this is the increasing number of readers and textbooks, aimed at an consisting of management students and practicing managers, and based on and presenting the main texts or the main principles of the modern economics of

5 The journals certainly aren’t a different story: rudimentary of Strategic Management Journal, Science, Academy of Management Journal reveal that organizational economics is indeed a strong voice in the conversation. How did all this begin?

The perhaps first detailed argument that the – then relatively recent – economic theory of organization may be helpful in a strategy context was made by Richard Caves (1980). In “Industrial Organization, Corporate Strategy, and Strategy”, Caves argued that corporate strategy issues should be framed as constrained optimization problems, and that recent work of Alchian and Demsetz (1972), Mirrlees, Williamson and other economists of organization would be helpful for understanding the nature of this problem.

A later and highly influential contribution was David Teece’s 1984 paper, published in The California Management Review, on “Economic Analysis and Strategic Management”. Its main point was that while economics until recently had been of little use for the strategy scholar because of its highly stylized assumptions, this situation was now rapidly changing, not the least because of the emergence of organizational economics, “... an important set of new ideas and theories ... with strong normative implications for several strategic management issues” (p.98). In the paper, Teece further indicated how capabilities considerations may be integrated with transaction cost arguments in order to enrichen strategic management research, an approach that he had pioneered earlier in his work on diversification (e.g., Teece 1982).

A number of later contributions have made similar points (Joseph Mahoney and J. Rajendran Pandian 1992; Edward Zajac 1992; Anju Seth and Howard Thomas 1994; Christian Knudsen 1995). These papers all survey the economics of organization, and some (e.g., Knudsen 1995) point out that while the theory marks an important breakthrough of considerable significance for the further development of the strategy field, it is also limited because it fails to sufficiently incorporate bounded rationality, limited cognition and differential capabilities. However, although these contributions recognize that the economics of organization can not alone constitute a full-blown strategic theory of the firm, the main message of organizational economics, that economic organization is above all
a matter of harmonizing misaligned incentives attendant on incomplete information, is left unchallenged.

management studies in general from, primarily, organization scholars (Charles Perrow 1986; James Robins 1987 Lex Donaldson 1990). The critique has reflected much of the traditional presumed under-socialized conception of agents has been criticized, as have the related ideas of methodological individualism and self-interest as the sole

On the whole, strategic management scholars appear to have been less critical in their reception of organization economics than organization influence on strategic management than on organization theory, and, second, because strategic management has been (even) more fragmented than organization outside.

Conner (1991) and Kogut and Zander (1992). Starting from a capabilities perspective, they concentrated on the assumption of opportunism/morally – a crucial (and necessary) assumption in organizational – and argued that it is possible to derive a rationale for the existence capabilities considerations.7

In economics, Harold Demsetz (1988) is probably the first sustained critique of the modern capabilities perspective on economic organization (e.g., Conner and Prahalad 1996; Madhok 1996; Grant and Baden-Fuller 1996) have drawn heavily on Demsetz' reasoning. Conner and Prahalad (1996) and Kogut and Zander (1996) are powerful restatements, as
rationale from their superior ability (relative to “the market”) to cultivate dynamic (learning) capabilities. What they perhaps did not explain convincingly was why markets cannot develop these capabilities in the absence of opportunism. However, in order to understand the significance and context of their arguments, it is necessary to make a somewhat more detailed look at the modern economics of organization.

III. The Modern Economics of Organization: Contributions and Problems

A. Coase and the Post-Coase Literature

As the story is normally told, the modern economics of organization traces its existence back to Coase’s landmark 1937 essay on “The Nature of the Firm”. What Coase observed was that, in the world of neoclassical price theory, firms have no reason to exist. According to the textbook, the decentralized price system is the ideal structure for carrying out economic coordination. Why then do we observe some transactions to be removed from the price system to the interior of organizations called firms? The answer, Coase reasoned, must be that there is a “cost to using the price mechanism” (Coase 1937: 390). Thus was born the idea of transaction costs: costs that stand separate from and in addition to ordinary production costs.

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: 390). A second type of cost is that of executing separate contracts for each of

---

8 This section builds on Langlois and Foss (1997).
the multifold market transactions that would be necessary to coordinate some complex production activity. Firm organization may avoid these costs, and exists for this reason. Including also various costs of internal organization helps explaining, by means of standard marginal reasoning, where the boundaries of the firm is located.

Much later, Coase (1972: 63) observed tartly that his 1937 essay had been “much cited and little used.” The landscape of economic thought changed significantly in the years that followed, and a large body of literature quickly emerged that not only “used” but in many ways sprang from Coase’s paper. It is in fact ironic that precisely at the time of Coase’s lamentation, serious work on economic organization that rested on distinctly Coasean foundations had actually begun.

More specifically, two landmark contributions were published in 1971 and 1972, namely Oliver Williamson’s “The Vertical Integration of Production: Market Failure Considerations,” and Armen Alchian and Harold Demsetz’ “Production, Information Costs, and Economic Organization,” respectively. These two contributions are seminal not the least because they helped founding different approaches within the modern economics of organization. Thus, Williamson’s work pointed the way to not only his own (as well as various associates’) work on transaction cost economics, but also to the more formal recent work of, for example, Grossman and Hart (1986). Alchian and Demsetz’ work, in turn, pointed the way to later work on the principal-agent relation (e.g., Bengt Holmström 1982). All this work followed Coase in conceptualizing the firm as a contractual entity whose existence, boundaries and internal organization could be rendered intelligible in terms of economizing with transaction costs. Where it differed from Coase, however, was in the emphasis that was placed on morally hazardous/ opportunistic behavior. Coase, in contrast, had not assumed this sort of behavior (see Coase 1988).
The present flagbearer of the field, Oliver Williamson has 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 come to crowd out all others in the explanatory pantheon. 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 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.

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

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 (e.g., Alchian and Demsetz 1972), 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.

B. Incentives and the Neglect of Knowledge

Whatever their differences may be, one central heuristic characterizes all the different streams that together constitute the modern economics of organization: an overriding emphasis on conceptualizing virtually all problems of economic organization as problems of aligning incentives. In the view taken here, this not only misrepresents important phenomena but also hinders understanding other phenomena.
A specific example of the heuristic mentioned is a recent paper by Julio Rotemberg and Garth 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 contract 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 – because this would mean breaking with the heuristic of reducing all problems of economic organization to problems of aligning incentives.

A broader example is supplied by Paul Milgrom and John Roberts’ successful textbook, Economics, Organization and Management (1992). As Brian Loasby (1995) points out in a review essay, the book contains much interesting material about economics, less material about organization and relatively little that a more traditional scholar in the field would call management, including strategic management, proper. To put in a pointed way, the book certainly does not tell us “... how Jones should decide what to do at nine o’clock on Monday morning” (Loasby 1995: 474). But this is the crucial issue: how is the set of possible choices discovered? How are the consequences known? These are issues of discovery and imagination, and they are not easily translated into the language of optimization and incentives. However, they are what strategic managers do.

The problem is certainly not that reformulations of traditional management and strategy issues in terms of optimization and incentives are internally inconsistent. Rather, the issue is whether the mechanisms so identified are in fact plausible explanations of the phenomena under study. In fact, it is quite likely that the mechanisms underneath, for example, the narrow firm strategies that Rotemberg
and Saloner (1994) talk about have little or nothing to do with the alignment of incentives, and have everything to do with limited knowledge, understanding and perception, in short, with firm capabilities.

The point here is not that one should never make these assumptions or that one should never model problems of organization or management as largely problems of incentive alignment. The point is rather that to translate these assumptions into an exclusive and near-universal research strategy arguably closes off a range of plausible alternative explanations of what firms are and what strategic managers do.

As I 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, if you prefer, dispersed, bounded, sticky and idiosyncratic. The second is the possibility that knowledge about how to link together one person’s (or firm’s) productive knowledge with that of another is also imperfect. The first possibility leads us to the issue of capabilities; the second leads to the issue of qualitative coordination in a world of bounded rationality and cognition. These issues are treated seriatim in the following.

**IV. The Capabilities Perspective**

A. Path-Dependence in the Economics of Organization.

The recent blossom of interest in the economics of organization has clearly been driven by a dynamic within present-day economic theory, one fueled mostly by advances in the economics of information and by applications of game-theory and recent mathematical methods. At the same time, however, the modern literature also owes much to the way Coase originally sought to explain the existence, the
boundaries, and the internal organization of the firm. Coase was not writing in a vacuum: he was working within the context of pre-war economic theory. As a result, today's economics of organization bears the imprint of the economics of the 1930s.  

One legacy of this “path-dependent” history has been a tendency (albeit an imperfect tendency) to respect an implicit dichotomy between the production aspects and the exchange aspects of the firm or, to put it another way, between production costs and transaction costs, or, 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” (Harold Demsetz 1988). Knowledge about alternative production possibilities is explicit, freely transmissible, and easily encapsulated in blueprints – in price theory as in the economics of organization.

The result of this partition of responsibilities has admittedly 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.  

B. The Capabilities Perspective

---

9 See Langlois and Foss (1997) for more on this.

10 Williamson argues this 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: 88). This may indeed be a sensible starting point, so long as it is not an ending point.
This is changing because of the development of a corpus of 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 economics of organization. Indeed, in its incarnation as a theory of economic organization, the capabilities perspective begins from the characteristics of productive activities rather than from the characteristics of the exchange process.

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: 24). In her work, Penrose is careful to point out that allocation by means of "administrative decision" is fundamentally different from allocation by means of the price mechanism (1959: 20), thus establishing a link to Coase. But whereas Coase (1937) was eager to emphasize that he was simply extending existing economic analysis, Penrose makes an explicit break with mainstream economics. Hers is a subjectivist and disequilibrium theory of the firm. Moreover, it is a theory that stresses entrepreneurship. In Penrose's story, the management team holds images of the external environment and of the firm's internal resources (this is the subjectivist part of her analysis). She further argues that these images are produced through internal learning processes, and that they determine the constantly changing "productive opportunity set" (this is the disequilibrium part of her story) of the firm, that is, the

---

11 Arguably, this may underlie Williamson's (1994) attempt to construct a false dichotomy between "strategizing" and "economizing". See Foss (1997a) for a critique of Williamson.

12 Although her book was apparently written in ignorance of Coase's contribution.

13 This may be contrasted with the attempts in some quarters to align the resource-based perspective with mainstream price-theory (e.g., Jay Barney 1991; Margaret Peteraf 1993).
productive possibilities that that the firm’s “entrepreneurs’ see and can take
advantage of” (Penrose 1959: 31).

Thus, Penrose explicitly begins from cognition. In contrast, modern
contributions to the capabilities perspectives simply begin from the empirical
generalization that productive knowledge is neither explicit nor freely
transferable.14 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
co-operating individuals.

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 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 operated by team-members for a strategic purpose.

It is well-known that this basic analysis forms the backdrop to an analysis of
firm heterogeneity, competitive advantage, and differential rents (Steven
Lippman and Richard Rumelt 1982; Birger Wernerfelt 1984; Barney 1991; Peteraf
1993). What is important in the present context, however, is that it is becoming an
increasingly widespread recognition among contributors to the capabilities view
that conceptualizing the firm in the above way has fertile implications not only for understanding the sources of competitive advantage etc. but also for advancing the economics of organization (Conner 1991; Langlois 1992; Kogut and Zander 1992; Foss 1993; Langlois and Robertson 1995; Conner and Prahalad 1996; Kogut and Zander 1996; Madhok 1996; Grant 1996). I discuss this in the remainder of the present section.

C. The Capabilities Perspective as a Theory of Economic Organization

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 what Richard Langlois has called “dynamic transaction costs”, the costs that arise in real time in the process of acquiring 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, suggests that the capabilities perspective may be interpreted as an alternative theory of economic organization.

That a capabilities perspective may lead to a distinct theory of economic organization is not a new recognition. For example, Penrose (1959: 146-148) clearly hinted at this. However, the first full-blown argument was presented by George B.

---

14 The work of Kogut and Zander (1992, 1996) and Spender (1996) are partial exceptions to this, however.
Richardson (1972) (who coined the capabilities terminology). In his 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 are best undertaken under unified governance.

Richardson’s basic point clearly is that capabilities may be determinants of the boundaries of the firm. A possible starting point of the argument is that a firm may control production knowledge that is, in important dimensions, strongly different (“dissimilar”) from what others control. And an implication may be that 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. 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.

As observed already, such dynamic transaction costs are in a different category from the transaction costs usually considered in the modern economics of organization: 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, cognitive frames, and skills 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 (Alfred Chandler 1992). 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.

The upshot of this section is that the capabilities perspective indeed is a distinct emerging perspective on economic organization, one that would appear to be particularly well suited to explaining the boundaries of the firm in dynamic environments. It is characterized by highlighting explanatory mechanisms that are different from those of modern economics of organization, not the least with respect to 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 economic organization.

D. Weaknesses of the Capabilities Approach to Economic Organization

There is, on the whole, an optimistic tone to the new capabilities oriented contributions to economic organization. Thus, a claim made is that it is possible to construct a theory of the firm that does not rely on the seemingly cynical assumptions as to nature of human action that is characteristic of, for example, Williamson’s brand of organizational economics. This is likely to appeal to the many critics of these assumptions in management studies (e.g., Ghoshal and Moran 1996).

Second, an accompanying claim is that it is possible to say something substantial about the explananda of the modern theory of the firm – namely the existence, boundaries, and internal organization of the firm – from a completely
different starting point. Thus, instead of starting from an analysis of the exchange nexus, and derive propositions about economic organization from an analysis of individual transactions, what is highlighted instead are differential capabilities and the nature of knowledge embedded in capabilities. In other words, the analysis begins from production rather than exchange. This is likely to appeal to strategy scholars, who already possess a relatively well-developed analysis of how differential capabilities are linked to competitive advantage.

The purpose of this section is to identify a number of weaknesses of the capabilities perspective as a distinct perspective on economic organization; weaknesses that may have been neglected under the impact of the present enthusiasm. The critique should not be taken as decisive in the sense that I argue that the capabilities is, for example, inherently unable to predict or explain the existence of firms. What I am saying is rather that there are weak areas that deserve to get more attention, particularly considering that many of the weak points have been dealt with rather convincingly by the modern economics of organization.

The critique falls in five points, some of which are related: The capabilities perspective 1) does not explain the existence of the firm; 2) does not convincingly explain asset ownership; 3) is not predictive; 4) does not explain when capabilities matter for economic organization and when they don’t; and 5) lacks clear microfoundations and modeling heuristics. I discuss these seriatim and in a telegraphic form.

The Existence of the Firm
The arguably strongest arguments that the capabilities perspective can successfully

(1992) and Conner (1991). To quote Kogut and Zander:

15

... differ radically from that of the firm as a bundle of

entily property rights” (ibid.). Specifically, it does

moral hazard.

and Zander say, “

that the market supposedly cannot

forward a quite related argument:

relations joint activity

[and this explains why they exist, NJF]. Further, the advantage of firms in

opportunism-control advantage” (Conner 1991).

and the full realization of the values of assets/ resource can only take place within the

arguments, placing the emphasis on the role of firms as integrators of specialized knowledge.

to include “shared coding schemes”, “values”, “a shared language”, as well as “mechanisms by

context of a set of higher order organizing principles of shared cultures, languages, codes, etc. characteristic of firm organization.

The basic problem with this reasoning is that such embeddedness does not conceptually presuppose common ownership/firm organization. In a moral utopia, characterized by the absence of opportunistic proclivities (the setting that implicitly underlies Conner, Kogut and Zander’s analyses), the gains from resources/assets being embedded in higher order organizing principles could be realized over the market. Agents (human resources) could simply meet under the same factory roof, own their own pieces of physical capital equipment or rent it to each other, and develop value-enhancing higher order organizing principles among themselves, or in other ways integrate their specialized knowledge (as a team). In the absence of opportunism/moral hazard, the degree of “integration”/co-specialization among the various resources/capabilities/assets or their embeddedness in “higher-order organizing principles” would not seem to have any implications for the issue of who should own what: firm organization would be precisely as efficient as market organization.

This is not to deny that their reasoning (including Conner and Prahalad (1996) and Kogut Zander (1996)) may be useful for understanding the employment contract. However, they seem to think that the employment sufficiently characterizes the firm, and defining the firm on the basis of the employment contract alone is futile (cf. Coase 1988). As Stephen Cheung (1983: 17) points out, “... according to one’s view, a ‘firm’ may be as small as a contractual relationship between two input owners or, if the chain of contracts is allowed to spread, as big as the whole economy ... [because] the delegation of use rights and the transmission of price information are matters of degree”.17

---

17 As examples of the difficulties of defining the firm on the basis of the employment relation, consider the following: If I hire a gardener to tend my garden under my supervision and instructions during the next week, does this mean that the gardener and I now constitute a
Explaining Asset Ownership

As argued by Williamson (1985, 1996), Grossman and Hart (1986) and others, asset-ownership is the key to defining the firm. In this tradition, indeed, the firm is defined as the collection of assets that the firm’s owners/managers control. One reason why asset ownership matters is that it allows us to understand bargaining power and hence the derived concept of authority: the owner/manager, who controls assets, can threaten the employee of depriving him of the assets with which he is currently working. We may disagree with the idea that the pattern of asset ownership is indeed the essence of the firm, but it is hard to disagree with the proposition that the pattern of asset ownership is a crucial aspect of the firm, and that any comprehensive theory of the firm must address this issue. Capabilities theories of the firm do not address it.

Lack of Predictive Capability

As it stands presently, the capabilities perspective is more of an explanatory than a predictive approach. That is to say, it allows the analyst to tell some ex post story about the causes of the success of a given firm, or why the boundaries of that firm are what they are. But it does not, in its present version, allow us to predict future success...
or future boundary decisions. There are two obvious and related problems with this. First, the normative value of the capabilities perspective is unclear. Two, the perspective cannot be tested and risks being tautological.

Again, this is certainly not to argue that the capabilities perspective is inherently non-predictive. For example, it may be possible to operationalize Richardson's (1972) categories of “similar-dissimilar” and “complementary-closely complementary” activities, and use these as independent variables in regressions, the dependent variables being the boundaries of firms.

**Explaining When Capabilities Matter for Economic Organization**

In many contributions to the capabilities perspective (e.g., Kogut and Zander 1992), it is presented as an alternative and general perspective on economic organization. However, it is easy to imagine cases in which capabilities do not influence economic organization, such as the boundaries of the firm. For example, when technological knowledge is essentially shared in a population of firms and firms therefore confront the same direct production costs for a given productive activity and do not risk running into communication problems with suppliers or licensees, we would not expect capabilities to be determinative of observed boundary choices. Incentive considerations would dominate.

The point here is that we need to have more fine-grained reasoning that allows us to identify under which circumstances capabilities matter and when they don’t.\(^{20}\) Work here may begin on the foundations laid by Langlois and Robertson (1995), one of only two contributions so far to present a detailed analysis of the issue under consideration here (the other one is Argyres 1996). In their theory, the organizational question is whether new capabilities are best acquired through the market, through internal learning, or through some hybrid organizational form. And the

---

\(^{20}\) Of course, this point is closely related to the above point about predictive ability.
second, the nature of the economic change involved.

what already exists in the economy, then a Schumpeterian process of creative
decentralized into what we may loosely call markets, then a reorganization within
necessary redeployment. If, by contrast, the old configuration of capabilities lies
the form of markets superseding firms. History offers many examples of both.

reconfiguration required. If change is — if it requires simultaneous
costly . If, however, change is — if change can take place
connected together — then markets, which can take advantage of specialized and
standards enters the picture: for standards are typically ways of fixing the
directions. Langlois and Robertson (1995) call this kind of structure a
system

Lack of Clear Microfoundations and Modeling Heuristics

organization is characterized by having rather unambiguous microfoundations and
organization (patterns of property rights, incentives and monitoring) as the
incomplete information setting.
In contrast, there is no agreement on the microfoundations of the capabilities perspective. What precisely is assumed about the motivation, cognition and knowledge of the agents that populate the firms treated in the perspective? Some implicitly decide to side-step this issue, in favor of the empirical generalization that firms are simply constrained in what they know to do well (Langlois 1992). Others think that it is necessary to draw on classic contributions to epistemology (Loasby 1991; Spender 1996), while others begin from mathematical computation theory (Dosi and Marengo 1994). A possible result of the lack of agreement about (and interest in?) microfoundations is the plethora of names that are now placed on strategic knowledge assets, such as routines, competencies, capabilities, dynamic capabilities, etc., and the lack of agreement about how these concepts relate.

Moreover, how knowledge considerations relate to (non-incentive) economic considerations are seldom made clear. To which extent may knowledge-based arguments about economic organization be interpreted in terms of standard information costs? What exactly is the relation between, for example, tacitness and complexity on the one hand and information and communication costs on the other hand? Is the perspective consistent with optimization or is another behavioral logic required? Is it consistent with equilibrium? Where is the model that identifies the threshold level where the firm is indifferent between make or buy? In other words, clear modeling heuristics are lacking. In contrast, organizational economics is characterized by rather clear modelling heuristics (e.g., Bengt Holmström and Jean Tirole 1989).
V. Toward a Strategic Theory of the Firm:
Research Strategies

A. Taking Stock
The argument so far can be summarized thus: in searching for a strategic theory of the firm, we are confronting two imperfect contenders that are characterized by explaining economic organization in widely different ways. I have discussed their imperfections and explained why they both at their present state of development are unlikely to serve as bases for a strategic theory of the firm. To put it briefly, while the economics of organization is of considerable relevance to the strategy field, and while it is characterized by (relative) explanatory elegance and simplicity, it is also likely to misrepresent many strategy issues. On the other hand, while the capabilities perspective is much truer to the traditional interests and concerns of the strategy field (because this is partly where it was originally developed), it is at present much too dispersed and explanatorily unclear to serve as a strategic theory of the firm.

However, both approaches are likely to further develop. For example, we may imagine that the economics of organization will become more attentive to the type of non-incentive coordination problems that are highlighted in the capabilities literature, that is, to say the problems of integrating dispersed and sticky knowledge (e.g., Grant 1996; Langlois and Foss 1997). And we may imagine that the capabilities perspective will gradually accumulate greater explanatory clarity and power as unambiguous terminology becomes terminology, the occasional math model emerges, etc. One possibility is that the two approaches will more or less continue to develop in isolation.

21 Roy Radner (1992, 1996) has long been interested in the non-incentive aspects of organizational coordination, and his approach seems to have met with some interest lately.
Another possible – and in my view desirable – outcome is that the two approaches may increasingly make contact and be developed jointly. In fact, some integrative work already exists (e.g., Teece 1982; Dosi, Winter and Teece 1992; Teece and Pisano 1994), and a number of writers have suggested that a fuller integration of capabilities and modern economics of organization perspectives is something to be striven for (e.g., Mahoney 1992; Seth and Thomas 1994). However, they have not fundamentally explained, first, why such a fuller integration is desirable nor, secondly, how it should be accomplished.

With respect to the first issue, the answer may be relatively simple. We need ideas from both approaches to explain a number of real world phenomena. For example, a comprehensive understanding of diversification or the economic organization of the innovation process would seem to require that capabilities considerations be combined with transaction costs considerations. More generally, opportunistic or morally hazardous behavior are, if we like it or not, facts of the real world. To the extent that the capabilities approach insist on abstracting from opportunism (e.g., Kogut and Zander 1992), it simply does away with an arguably important determinant of economic organization, in the same way that organizational economics may be criticized for doing away with differential capabilities as determinant of economic organization. The second issue is much more difficult to approach, and I can only offer some loose speculation.

B. Some Ingredients of a Strategic Theory of the Firm
The development of a strategic theory of the firm is a process that should be based on a realistic understanding of what strategic managers do. If this sounds like a platitude, recall that I criticized the modern economics of organization for constructing a whole theory of the firm and the managerial task (e.g., Milgrom and Roberts 1992) on a basis that is arguably much too narrow in its understanding of

among formal economists of organization (e.g., Patrick Bolton and Matthias Dewatripont 1994).
directly useful for understanding what strategic managers do.

The essential ingredients in a strategic decision-making sequence are imagination, cognition, coordination, and capabilities, along with incentives and governance. My aim is not really to elaborate on these ingredients (in fact, much of the view taken here, strategic managers exercise judgment (Frank Knight 1921) rooted in (imperfect) cognition and knowledge (Luigi Marengo 1995; Ulrich Witt 1993)). Something imagined deemed possible on the basis of the available evidence possible; the essence of strategic decision making is not choice among given (including “his” firm’s resources), define which variables are relevant, attaches knowledge of this sort of problem-solving behavior in complex environments (including its trial-and-error component), and, secondly, is of considerably broader scope.

---

22 Knight (1921), Foss (1993), and Casson (1997) argue that the existence of the firm is
If (strategic) managers are judged by their ability to make the right choices among a set which they themselves partly generate, they are also judged by their ability to break with existing practice, to mediate and to engage in sense-making (Karl Weick 1997). Thus, leadership and the provision of cognitive frames enter the picture.\textsuperscript{23}

These are phenomena that are also hard to accommodate in the context of organizational economics, which – like the rest of mainstream economics – assumes that agents come endowed with essentially the same cognitive frames. Thus, the manager of division A has essentially the same understanding of what goes on “out there” as the manager of division B, and as top-management do. Therefore, his failure to take the right action (or provide the right level of effort) can only be ascribed to morally hazardous behavior. In reality, however, an important part of what managers do, and a major organizational design problem, is to get everybody on the same wavelength, as it were.

Clearly, this has implications for understanding the coordination of local and dispersed knowledge, arguably a major design problem for any social system (Friedrich von Hayek 1945).\textsuperscript{24} But it has also implications for understanding economic organization. Different cognitives frames – that may clearly be partly tacit (Polanyi 1958; Marengo 1995; Witt 1996) – are what underlie concepts such as “similar” and “dissimilar” capabilities (Richardson 1972). Thus, an analysis of organizational cognition may help strengthen the capabilities theory of the boundaries of the firm: it helps understanding why contracts, say, between two firms, may not be efficient means of coordinating activities and the size and nature of the information costs that cause this. Moreover, such an analysis helps understanding patterns of diversification, for example, to the extent that

\textsuperscript{23} See Witt (1996) for an extended analysis of leadership as the provision of cognitive frames.

\textsuperscript{24} For an analysis of this issue and its connection to the theory of the firm, as well as other issues connected to Austrian economics, see Foss (1997d).
“relatedness” a crucial concept in the study of diversification is not something Stimpert and Irene Duhaime 1997).

As argued earlier, the emphasis in the modern economics of organization on seriously considered as coordinating devices. The role of cognitive frames, J.C. Spender (1994: 354) points out, there may ironically be a similar tendency

“When we overlook the resource application processes we miss what overlook the core of our theory of the firm, the process of has paid little attention to the construction and management of the

Thus, hitherto capabilities have been examined because they may be key strategic because they tell us something about the firm’s internal “resource application

For obvious reasons, however, a strategic theory of the firm need to pay capabilities perspective. We also need to better understand capabilities in order to the whole issue of organizational learning. Without a solid theory of how organizations learn (Kogut and Zander 1992). Luckily, this has recently received
much attention in the organizational knowledge literature (e.g., Nonaka and Takeuchi 1995; Grant 1996) and in evolutionary economics (e.g., Dosi and Marengo 1994).

**Incentives and Governance**

The position taken here is that transaction and information costs complement the capabilities perspective, and do so on several levels.

First, in the absence of these costs, there would not be any differential capabilities at all; hence no competitive advantage (Demsetz 1973). And economic organization would not make a difference; for example, whether a transaction was organized on an arms-length basis or in-house would have no implications for profitability, since both alternatives would be equally efficient. Thus, information and transaction costs must be the ultimate foundation of a strategic theory of the firm.

Secondly, it is simply naïve to deny the real-world importance of opportunism and the need for monitoring and incentives to curb opportunistic behavior (even bearing the lessons of Ghoshal and Moral (1996) in mind). Indeed, we need standard organizational economics insights to help us explain that part of economic organization that is not best understood in terms of capabilities (e.g., the choice between employees or sales agents, the choice between franchising and in-house distribution, etc.).

Third, although capabilities are emergent phenomena, they are partly composed of individual agent’s stock of knowledge, even if not fully reducible to these. And agents do not just accumulate valuable knowledge that may fit into existing capabilities or help creating new capabilities: they normally need to be given incentives (though not necessarily monetary ones) to do this, and opportunistic inclinations – even if these have been strongly exaggerated in organizational economics – that may harm these processes need to be curbed.
Thus, to sum up, there is in prospect an exciting and ambitious research program, aimed at building a general strategic theory of the firm and using and combining the best ideas of the capabilities perspective and the modern economics of organization.

VI. Conclusion

I hope to have demonstrated in this paper that in searching for the foundations of a strategic theory of the firm, we are dealing with highly imperfect alternatives. Admittedly, both alternatives, organizational economics and the capabilities perspectives, are on their way to becoming strategic theories of the firm in the sense of the term employed here. Thus, both have things to say about the four issues that enter into the construction of such a theory, that is, the existence, organization, boundaries, and competitive advantage of the firm. However, both the capabilities perspective and organizational economics may be criticized on several scores.

Thus, I have added to the capabilities critique of the modern economics of organization by identifying the underlying heuristic of this work – namely the attempt to reduce literally all aspects of economic organization to matters of aligning incentives in a setting of asymmetric information – and by pointing out that this procedure arguably misrepresents a host of strategic management issues and issues of economic organization. However, I have also argued that the capabilities perspective as a perspective on economic organization suffers from a number of severe handicaps, some of which have to do with the fact that it still an emerging approach.

Given this situation, two possible research strategies immediately present themselves: to develop further the modern economics of organization or the
capabilities perspective in isolation, and hope that as a result of intense research efforts, one of the perspectives will eventually produce a satisfactory strategic theory of the firm. However, I have argued that a more sensible research strategy is to combine elements from both approaches. Thus, the message conveyed is essentially similar to that of contributions such as Mahoney (1992), Zajac (1992) or Knudsen (1995) which also pointed to the need for more integrative efforts.

I think that reflection will show that this is the sensible position. After all, it will probably be agreed that it is meaningless to identify a single aspect of the firm, declare this to be the essence of the firm, and build a theory of economic organization from this presumed essence. However, there is a tendency to this unfortunate essentialism in both the modern economics of organization and in the capabilities perspective.

Thus, firms are input combiners, bundles of contracts, cognitive entities, knowledge-accumulating entities, strategizing entities, political entities, information-processing entities, etc. – all of which may have implications for understanding economic organization. To focus on only one aspect may pay off as a short-term research strategy; in the longer run, we surely want to consider all those aspects that have a bearing on economic organization.


Organizational Economics


Information and Coordination


R. Fuchs, ed., Organization


and Five Schools of Thought Within Industrial Organization Economics: Do Journal of Management 17


