Laboratory of Economics and Management
Sant'Anna School of Advanced Studies

Piazza Martiri della Libertà, 33 - 56127 PISA (Italy)
Tel. +39-050-883-341 Fax +39-050-883-344
Email: lem@sssup.it Web Page: http://www.sssup.it/~LEM/

LEM
Working Paper Series

Toward an Evolutionary Theory of Production

Sidney G. Winter *

* *The Wharton School*

2002/27 **December 2002**

Toward an Evolutionary Theory of Production

Sidney G. Winter

The Wharton School

September 2001

Department of Management
The Wharton School
Philadelphia, PA 19104-6370

Phone: (215) 898-4140   Fax: (215) 898-0401
e-mail: winter@wharton.upenn.edu

Toward an Evolutionary Theory of Production

Sidney G. Winter*

Since the time of Adam Smith, Francois Quesnay and David Ricardo, economists have sought to ground their theoretical analyses of economic organization in an appreciation of the nature of real production activity. Any such an effort must balance two competing concerns. On the one hand, the obviously fundamental role of productive activity in economic life seems to demand a highly accurate appraisal, presumably based on detailed scrutiny. On the other hand, the objectives of economic science often seem best served by a broad-brush characterization done from a considerable distance. Those scientific objectives are, after all, quite different from those of engineering or operations management.

In mainstream neoclassical theory, the second of these considerations seems clearly dominant. Production theory as it has developed in that tradition is strong on abstract generality and treats production in a way that is convenient for the neoclassical analyst. Mainstream production theory is partly for answering questions about production and its place in economic organization, but it is at least equally concerned with sealing off questions that are not considered fruitful for economists. It places a boundary marker that serves to identify the limits of the specifically economic concern with production, beyond which lie areas of concern to engineers, managers and technologists.

It should be obvious that evolutionary economics needs to strike quite a different balance. Evolutionary thinking sees questions of production as tightly and reciprocally connected with questions of coordination, organization and incentives. Also, production activity is embedded – now more than ever – in a variety of processes of knowledge creation; theory needs to make room for those links. A major deficiency of the mainstream theory is its isolation from the realities of organizational knowledge. Above all, the evolutionary economist needs theory to address questions of economic change, not the principles of resource allocation in a hypothetical static world.

As is discussed below, a review of the historical development of production theory shows conclusively the shaping role of the analytical objectives held at each stage by the economists who made the contributions. On the contemporary scene, production function formulations now dominate because they provide a convenient basis for applied econometrics, and because the sorts of questions that require more general apparatus are no longer as salient for the discipline as they were a few decades back.

The dominance of the production function apparatus in contemporary mainstream treatments of technological change is also a "Panda's thumb" phenomenon; it reflects the logic of path-dependent evolution (Gould 1980). The apparatus was created and developed for various other reasons, and when questions of technological change came to be confronted it was conveniently available. The inherited apparatus was then extended and supplemented by a variety of formal treatments of technological change, the simplest being the introduction of the multiplicative factor A in the relation $Q = A f(x)$. Negligible attention was paid to the question of whether plausible real-world mechanisms might actually produce knowledge changes with effects that correspond to these formalisms; the formalisms are convenient and hence chosen for reasons other than micro-level verisimilitude.[1] The major investment in building a truly knowledge-based production theory, well-suited to close analysis of the problems of change, was never made. Recently, however, some beginnings have at least been made.

This paper sketches some of those beginnings, and attempts to establish their links both to parts of mainstream production theory and to the actual phenomena. The following section reviews the historical development of production theory and substantiates the claims made above. Section 2 argues that the sort of "knowledge" that is applied in productive activity – henceforth, *productive knowledge* -- has specific attributes that make it quite different from the things termed "knowledge" in other contexts. Section 3 examines some deep issues surrounding the seemingly straightforward notion of "spatial replication," i.e., the idea that the same knowledge can be used in more than one location. The concluding section discusses the place of production theory in the evolutionary economics framework and identifies some of the research tasks on the agenda.

1. Origins and Varieties of Production Theory

The topics that we now count as part of the theory of production were, over much of the history of the subject, addressed primarily in the context of the problem of distribution, which was itself viewed as the problem of explaining the division of the social product among the classic "factors of production"—land, labor, and capital. The major purposes now served by theory of production—in analysis of the role of production possibilities in the determination of relative prices, and in the efficient allocation of resources—only began to acquire something like their present importance with the advent of neoclassical economics late in the nineteenth century. In classical economics, it was the marginal productivity schedule, and not the production function or the cost function, that was the focus of attention. The question was not, "how much output can be obtained, at a maximum, from this list of inputs?" but rather, "by how much will output increase if the amount of this particular input is increased some, with all other inputs held constant?"

From our present-day theoretical standpoint, we recognize that the latter question is not well posed unless there is some understanding about how the increment of the variable input is to be used. Our standard resolution of this problem is to understand both the original input list and the augmented list to be used in such a way as to produce a maximum output given the "state of the arts"—i.e., to refer back to the first of the two questions just stated. In the classical treatments, however, the discussion of the "laws of returns" is not so clearly founded on the concept of technical efficiency. Rather, the propositions are advanced more as observational laws characterizing the results of experiments—actual or imagined, natural or planned. These experiments involve simple incremental modifications of familiar methods for producing output. The increments of input are apparently conceived as being applied in a reasonable way, but there is little suggestion that the matter requires careful consideration. The important point is that the quasi-empirical marginal productivity schedule involved in the classical conception does not raise any troubling questions of how it is known that the methods involved actually yield maximum output for given inputs. To put it another way, the conception does not challenge the theorist to develop the idea of a "state of the arts" with the care appropriate to the importance bestowed on that concept by the modern interpretation of the production function.

The primacy of the distribution problem and of the quasi-empirical conception of the marginal productivity schedule persist in early neoclassical discussion of production. Particularly significant here is Wicksteed's 1894 work, An *Essay on the Coordination of the Laws of Distribution* (Wicksteed 1894). His title names his task, which was to show that separate laws of distribution, involving a classical application of marginal productivity principles to each factor in turn, are mutually consistent in the sense that the resulting factor shares exhaust the product. It was in introducing this problem that he made the notion of the production function explicit in economic analysis for the first time in the following terms:

> The Product being a function of the factors of production we have $P = f(a, b, c \ldots)$ [2]

Neither in this statement nor in Wicksteed's subsequent analysis is there any hint that there is anything conceptually problematic about the idea of such a function; it is merely a mathematically explicit expression of the long-familiar idea that if the input quantities vary, the output quantity will vary as well, and in certain characteristic ways.

Even today, introductory treatments of the production function and of factor substitution in textbooks and lectures often follow much the same Ricardian path, with the same agricultural examples. The strength of the approach lies in the plausibility of variable proportions production in the agricultural context, and in the simple, commonsense arguments that establish the general character of the response of output to variation in a single input. A very casual acquaintance with a relatively simple production technology is the only prerequisite for understanding. But this strength is also a source of weakness of the sort suggested above; the loose way in which the technology is discussed tends to leave obscure the conceptual connections among productive knowledge, the production function, and technical efficiency.

No sooner had the production function made its explicit appearance in economic analysis than it was immediately—in the lines immediately following the above-quoted line from Wicksteed—specialized to being homogeneous of the first degree. This was the basis of Wicksteed's coordination of the laws of distribution, i.e., his demonstration that the product is precisely exhausted when inputs are paid according to their marginal products. In beginning his examination of the validity of the constant returns to scale assumption, Wicksteed stated:

> Now it must of course be admitted that if the physical conditions under which a certain amount of wheat, or anything else, is produced were exactly repeated the result would be exactly repeated also, and a proportional increase of the one would yield a proportional increase of the other.[3]

Elaborating this statement, Wicksteed made clear that the replication had to be understood to involve a replication of the inputs in exact detail; otherwise one would not be exactly repeating the original condition. He then went on to deal, in a somewhat confused way, with the fact that the economic laws of distribution must involve something more than the physical conditions of producing the "mere material product." What he did not pause to explain—but presumably had in mind—was that the "physical conditions" that are "exactly repeated" include not merely an identity of context but an identity of production method. This supposition does make the result obvious in the sense that a replication of a given physical experiment, hypothetically perfectly controlled, necessarily yields the same result. What is noteworthy in the present context is (a) that this interpretation is inappropriate given the modern understanding of the production function, which presumes that there is a choice of method, and (b) that Wicksteed's discussion manages to slide past the problem of characterizing the set of available methods, or "state of the arts."

At some point, the connection between the production function concept and technical efficiency began to be emphasized. A clear statement may be found in Sune Carlson's 1939 book, A *Study on the Pure Theory of Production* (Carlson 1956).[4]

> If we want the production function to give only one value for the output from a given service combination, the function must be so defined that it expresses the *maximum product* obtainable from the combination at the existing state of technical knowledge. Therefore, the purely *technical* maximization problem may be said to be solved by the very definition of our production function.[5]

To whatever historical depth some awareness of this point might be traced, it seems clear that its salience was vastly enhanced by the advent of new approaches to production theory that gave explicit consideration to production methods *not* "on the production function." These new approaches comprised the members of the family of linear models of production, including linear activity analysis, linear programming and input-output analysis, and also such descendants and relatives of this family as process analysis, nonlinear programming, and game theory. They made their appearance in the work of von Neumann, Leontief, Koopmans, Kantorovich, Dantzig, and others over the period 1936 to 1951.[6]

The linear activity analysis framework, as developed by Koopmans, is most relevant here. This contribution introduced into economics a workable abstract representation of "the existing state of technical knowledge." Productive knowledge was described first of all by "basic activities," formally represented by vectors of technical coefficients, but conceived as corresponding to identifiable, concrete "ways of doing things." Further, the theory adduced a set of principles that described how the productive knowledge represented by the basic activities could be extended in scope, combined, and modified. Central to these principles were the assumptions that activities could be scaled up or down at will while maintaining the same proportional relations among inputs and outputs, and that the results of activities performed simultaneously would be additive. If these assumptions held true, then the whole scope of technological possibilities could be characterized in relation to the basic activities involved. This would mean, in particular, that in the case of a single output production process, the numerical specification of all basic activities would make it possible, in principle, to determine the maximum output that could be produced from a particular input combination. If the data were available and the problem not too large relative to the computation budget, linear programming solution algorithms would make such a determination possible not merely in principle, but in practice.

Still another mode of abstract representation of technological possibilities became common in economic theory with the development of modern general equilibrium theory by Arrow, Debreu, and others(Arrow and Debreu 1954; Debreu 1959). This approach generalizes the earlier ones by going simply and directly to the abstract heart of the matter. Commodity outputs in amounts represented by the list $\boldsymbol{q} = (q_1, \ldots, q_M)$ may or may not be producible from input commodities in amounts represented by the list $\boldsymbol{x} = (x_1, \ldots, x_N)$. If $\boldsymbol{q}$ is producible from $\mathbf{x}$, then the input-output pair (x, q) is "in the production set" (or "production possibilities set"). Whatever is known or considered plausible as a property of the structure of technological knowledge is, in this representation, treated as a postulate about the properties of the production set. For example, the linear activity analysis model is recovered as a special case if it is postulated that the production set comprises a finite set of basic activities, plus the combinations and modifications permitted under the activity analysis assumptions.

It is useful to note what is gained and what is lost in going from linear activity analysis to a general production set. What is gained is generality—there is a simple abstract representation for states of knowledge whose structure may not conform to that postulated by activity analysis. What is lost, naturally enough, is specificity—and potential empirical content. No longer is there the suggestion that it might be possible to fully characterize an actual state of technological knowledge by looking in the world for "basic activities." In particular, there is no guidance as to how one would test the claim that a specific input-output pair not actually observed in the world is "not possible given the existing state of technical knowledge." Thus, to refer back to earlier discussion, the concept of a production function that expresses the "maximum product obtainable" from each input combination is once again left without any direct empirical interpretation. This is not to say that its status as a purely logical construct is impaired; given additional mathematical restrictions that are commonly imposed, it is logically possible to define such a function on the basis of an underlying production-set representation of the state of technical knowledge. The construct thus defined may be employed in further theoretical development and perhaps in that way be related, ultimately, to empirical data.

The situation may be summed up in the following terms. It is the production set concept that stands, in contemporary formal theory, for the classical idea of a "state of the arts" or for an "existing state of technical knowledge." Arrow and Hahn concisely say

> Thus the production possibility set is a description of the state of the firm's knowledge about the possibilities of transforming commodities.[7]

To assume that the production set has certain properties—for example, those that correspond to the linear activity analysis model—is thus an indirect way of imputing analogous properties to the "state of knowledge" that the production set describes. I have proposed here that this indirect approach may be understood as a reflection of the historical development of the theory. In the modern synthesis of the subject, production sets are a fundamental concept, production functions are a derived construct, and marginal productivity schedules are an implied attribute of production functions. Historically, however, it happened in the opposite order. Marginal productivity came first (Ricardo), then production functions (Wicksteed), then production sets (Koopmans, Arrow and Debreu). In the "finished" structure of modern theory, the concepts that developed later are logically antecedent to those that appeared earlier. The development of the more recent arrivals has been strongly influenced by their logical role in an already extant theoretical structure; they did not have much chance to develop a life of their own.

Thus it happened that it became much easier for the theorist to describe the logical connection between the production set and the production function than to explain the substance of what the production set supposedly represents—a state of knowledge. This neglect of the independent conceptual anchoring of the production set idea has inhibited both the recognition of its limitations and the development of alternative and complementary theoretical treatments of productive knowledge. The following section initiates the discussion of such treatments by exploring the central concept itself.

2. The Nature of Productive Knowledge

It may be helpful to begin by pointing out that the word "productive" is here serving as something other than a simple modifier in the term "productive knowledge". In fact, while it is certain that this discussion is about production, whether it is exclusively about "knowledge" is a semantic issue that may be open to dispute. To many people, "knowledge" and "knowing" denote things whose characteristic locus is the individual human mind. Books, manuals, computer files and other symbolic records supplement and leverage human memories, but do not actually engage in "knowing." Neither are groups or organizations conceived as "knowing." However, an important implication of the discussion to follow is that a narrow focus on what goes on in human minds can seriously impede understanding what goes on when organizations produce things. That sort of understanding is the true objective here, and the scope of the term "productive knowledge" is therefore deemed to be expandable as necessary to cover whatever needs to be understood.

There are several important differences that separate the subject of productive knowledge from other knowledge domains. These will be discussed in turn.

Pragmatic Validity. Over the span of human history, a quest for knowledge has been a shared concern of philosophers, mystics, scientists, engineers, captains of ships and captains of industry. What these different sorts of seekers have meant by knowledge, and the uses they made of it, have been quite different. All have been aware, however, of the question of validity. It has been recognized that, in every realm of knowledge, there is some false coin in circulation and it can be a problem to distinguish it from the real thing. When it comes to the level of concern with validity, and the sorts of tests used to establish it, the different sorts of knowledge seekers part company again. The philosophers have generally been the most concerned with the quest for certainty and the critical scrutiny of claims to validity, asking whether and how we can know, and how we can know that we know. Overall, their conclusions have not been reassuring. Meanwhile, the specific concerns of the other seekers have generally led them to make lesser demands for certainty, and to employ tests of validity that the philosophers would scorn. Again, the differences of view among the other seekers are notable; experimental scientists see the validity issue in quite different terms than mystics.

For engineers, production managers and corporate strategists, the visible face of the validity problem is the question of transferability across time and space. The process worked well today, will it also work well tomorrow? If we build a similar facility in a remote location, can we confidently expect that one to work as well? The salience of these questions depends critically on the degree to which the answers seem to be in doubt. When experience seems to confirm the temporal and spatial transferability of the knowledge in use, it quickly acquires a "taken for granted" character. When problems arise, attention is directed to them until they are solved "for practical purposes." Under both conditions, the judgments made are not those of philosophers or scientists who care about the validity question for its own sake, but of practical people who need to get a job done and have other urgent demands on their attention.

In some areas, production activity is indeed akin (as Wicksteed implied) to repetition of a well-controlled scientific experiment, and the validity of productive knowledge is there akin to the validity of knowledge based on experimental science. The domain of productive activity is broad, however, and within that broad domain the question of validity plays out in diverse ways. In some areas, the metaphor of the scientific experiment is very wide of the mark.[8] Consider agriculture: from before the dawn of history, the unpredictability and variability of results achieved in agricultural production has been an important, often central, fact of human existence. That uncertainty has been dealt with in diverse ways, and methods of production have reflected those diverse understandings. Is it important to appease the local gods with appropriate rituals? To choose the best time to plant with the aid of an accurate calendar? To burn the stubble after the crop is harvested? To attempt to produce plant hybrids or control how domestic animals pair up in mating?" To choose which crop to plant on the basis of the government's long range weather forecast, or of the predictions in the *Farmer's Almanac*?

While modern thinking may dismiss some beliefs and related practices as plainly superstitious and others as ill-founded, the line between superstition and practical knowledge is oftentimes difficult to draw. The traditions of folk medicine have contributed some valid treatment approaches, such as quinine for malaria symptoms, along with much mystical obfuscation and harmful error. Similarly, a specific folk practice in agriculture may reflect genuine empirical regularities that, *at least in its actual context*, lend it practical value. For example, a primitive understanding of the phenomena of genetic inheritance has shaped the way humans use plants and animals for many millennia. The actual contribution of a folk practice may or may not take the form contemplated in the culturally approved rationale of the practice. Contrariwise, the fact that modern scientific and engineering understanding were applied in the creation of a new pesticide or a new seed variety does not provide a guarantee against adverse effects arising, for example, from unintended consequences for the local ecology.

A striking and well-documented example of these issues concerns the role of the water temples in the irrigated rice agriculture of Bali. In the traditional system that had developed over a period of more than a thousand years, the allocation of water among hundreds of farming communities was governed by the priests of the water temples. The temple system was responsive to the variation of rainfall by elevation, seasonally, and from year to year. Implicitly, it dealt with an underlying tradeoff between the requirements of pest control, which is facilitated by synchronized planting and harvesting among the farms, and the problem of water allocation, which is complicated by synchronized decisions. This traditional system was disrupted when the government promoted change in agricultural practices at the time of the "Green Revolution" in the early 70s. The result was a brief period of increased productivity, followed by a collapse caused by increasing pest problems and water shortages. Fortunately, the traditional system had been under scholarly examination by anthropologist J. Stephen Lansing, who extended his investigations into a systematic comparison of the ecological consequences and economic effectiveness of the traditional and officially-promoted systems. Ultimately – but only after many years – this research led to a reversal of policy and an ensuing recovery of productivity (Lansing 1991; Lansing, Kremer et al. 1998). Professor Lansing commented, "These ancient traditions have wisdom we can learn from."[9]

The unifying generalization here is that agricultural production is highly exposed to contextual influences arising in imperfectly understood natural systems of great dynamic complexity. Because of that exposure, the character of productive knowledge in agriculture has traditionally been remote from the model of repeated scientific experiments, and modern agricultural research has not eliminated that gap.[10] Beliefs about sound practice have at best a contingent validity that depends on the context being constant in ways that certainly cannot be assured by the available tools of "experimental control," and may in varying degrees be unidentified, unanalyzed and unobservable. At worst these beliefs may entail significant misunderstandings, and sustain ways of doing things that are significantly counter-productive.

Although agricultural production is unusual in its obvious exposure to vigorous environmental influences, the issue of context dependence is actually pervasive. Semiconductor production, for example, is at least superficially at an opposite extreme from agriculture in its exposure to context. A semiconductor factory (a "fab") and its operating procedures can be viewed as an enormous and costly effort to achieve strong "experimental control" on the production process for semiconductor devices, made in the interest of attaining consistently high yields. But that extraordinary effort is made because it is economically reasonable to make it, and it is reasonable because of the extraordinary sensitivity of semiconductor production processes. Extreme efforts at control notwithstanding, that sensitivity continues to present serious challenges to production engineers (Flaherty 2000). Thus, while one might plausibly suppose that such a tightly controlled context would be one in which process outcomes were fully understood in terms of underlying physical principles and "recipes" would be highly reliable, in fact the opposite is the case. The struggle to increase and sustain yields has sometimes been characterized in terms of superstitious belief rather than scientific principles.[11] Elements that might (superficially) appear to be superstitious even appear in codified organizational practice, as in Intel's "Copy EXACTLY!" technology transfer policy:

> Everything which might affect the process or how it is run is to be copied down to the finest detail, unless it is either *physically impossible* to do so, or there is an *overwhelming competitive benefit* to introducing a change.[12]

This policy establishes a burden-of-proof context for transfer decisions that could easily lead to careful "ritualistic" copying of details that don't matter. Of course, its true basis is not superstition, but a very rational adaptation to the reality that understanding of what *does* matter is limited. As in the case of folk practices in medicine or agriculture, the absence of a coherent scientific rationale for the details of the practice does not suffice to establish the absence of pragmatic validity.

Of course, common sense finds the grounds of confidence not in science, but in experience. This, however, is the kind of confidence that is placed in what is "taken for granted". As Hume (1748) classically observed, natural experience by itself cannot resolve causal ambiguity; and quantity of experience does nothing by itself to remedy its qualitative incapacity to speak to causation.[13] Although overstated as a conclusion on induction in general, this point has special force regarding the knowledge displayed in complex production processes. The knowledge invoked in such productive performances resides for the most part in individual skills and in organizational routines that are, at the multi-person level, the counterpart of individual skills. Skills are formed in individuals, and routines in organizations, largely through "learning by doing." They are developed in particular environments and meet standards of pragmatic validity that are characteristic of those environments. As is discussed below, the validity of a skill or routine shaped by such a learning process is typically dependent in a variety of ways on features of the environment in which it developed. Many of these dependencies may not be obvious and go un-remarked in the learning process—particularly in the case of dependence on conditions that are constant over time in the environment of learning but may vary over the environment of application. Because skills and routines reflect learning processes that are context-dependent and in part unconscious, they are inevitably ambiguous in their temporal and spatial scope. Whether the new circumstances of a new site – or simply a new morning – are different enough to be disruptive is a matter that is hard to be confident about *ex ante*.

Finally, there is one major consideration limiting the validity of productive knowledge that the examples of agriculture and semiconductor production may not adequately suggest: people are involved. The work done by human beings in productive roles involves not only the physical manipulation of tools and work in process, but also human information processing of both simple and complex kinds. The specific physical effects the actors can produce depend on physical attributes like their strength and stature; the information processing they do depends on their sense organs and their minds. Needless to say, human beings are quite heterogeneous with respect to these attributes, and in many cases the attributes are quite difficult to observe. Further, the behavior of humans in work situations is molded into characteristic patterns by cultural and institutional factors that vary geographically, particularly among nations (Kogut 1991; Dosi and Kogut 1993). These considerations limit very seriously the relevance of the validity model provided by repeated, perfectly controlled scientific experiments.

Further, “ people are involved” in production in other ways besides their roles as sources of input services. People are also involved as the customers, the consumers, the ultimate arbiters of productive achievement. When the product is corn or computer chips, it may be reasonable to consider that the “experiment” ends when the product appears, and set the customer response aside as a separate problem. But what if the product is education, business consulting, health care, elder care or day care, entertainment, or just “the dining experience”? Certainly the arrangements and processes that yield these products ought to be within the scope of a useful theory of productive knowledge: not only do the national income statisticians regard them as production, they account collectively for a substantial fraction of the GDP! Yet if the true definition of the outcome of the productive “experiment” lies somewhere within the being of the customer, it is clear that another broad range of contextual influences is potentially relevant – and that the prospects for high validity are dimmed accordingly.

By way of conclusion to this discussion of pragmatic validity, it is important to underscore the fundamental difference in the operative incentives that distinguishes the economic production context from experimental science. In the former, the direct payoff is to output, not to depth of understanding. Of course, the process of scientific inference from experimental results is itself subject to well-known hazards, even when a strong theory of what might affect the result guides a diligent effort at control. The hazards of inference from "this has worked up to now" are notably larger in the context of productive knowledge, where the dominant motivational focus from the very start is on *whether* something works, and not why. Quite rational satisficing principles dictate that investment in the quest for understanding be deferred until there is a symptom of trouble to deal with. When the pace is fast and the competitive pressure intense, such deferral may even involve suppression of ordinary skepticism about the rationale for prevailing ways of doing things. Paradoxically, "practical men" are constrained to be gullible, while high standards of proof are a luxury reserved to certain cliques among the inhabitants of the ivory tower.

Hazy frontiers.[14] The operational aspect of validity was identified above as the transferability of existing knowledge across time and space. When the activity attempted is of a novel kind in itself, judgments about feasibility are subject to hazards and uncertainties well beyond those that attend such transfer. The part of the knowledge territory that is near the advancing frontiers has special importance and a special character, and requires separate attention.

The special importance of this area arises from its place in the linked realities of economic competition and historical change. Competition, for gain and for survival, pushes organizations to seek out new and better ways of doing things. Even for similarly situated business firms that are doing the same thing in roughly the same way, competitive success or failure often rides on the small differences among them, and those differences often reflect efforts at incremental innovation. In other contexts, innovation competition is the central competitive fact. Attempting major advances, firms necessarily find themselves operating at the fringes areas of known technology. In these areas, in the nexus of economic motivation and advancing knowledge, lie the sources of economic growth (Nelson 1995; Nelson 2000). It is in these areas that the vision offered by mainstream production theory, with its image of a static "state" of valid knowledge represented by a sharply defined production sets, is most distorted and misleading. As economists always emphasize, "incentives matter" – but in this area they matter in ways that cannot be cleanly separated from the ongoing, path-dependent evolution of the technological opportunities.

The resource allocation decisions that shape the path of advancing knowledge are necessarily made, at each point of the path, on the basis of the limited knowledge then available. Surveying available knowledge from such a vantage point is like surveying a landscape on a hazy day. Some things are close enough to be seen clearly, others remote enough to be totally invisible. The intermediate conditions cover a broad range and are complex to describe. In that range, the verdict on the visibility of a particular feature depends crucially on what definition of "visible" is chosen from a range of plausible alternatives. Similarly, whether something is to be regarded as "known" or not will often depend on what one chooses to mean by "known," and why one cares; ultimately it depends on the degree of indefiniteness concerning details that is regarded as consistent with the thing being "known."

Visibility across the landscape of productive knowledge is arguably an even more complex phenomenon than ordinary visibility. A particularly complex type of haze arise from the fact that, in all spheres, practical knowledge consists in large measure of knowledge of how to *try* to find solutions to problems that have not previously been encountered. It includes techniques for patching together solutions of familiar problems to solve, or partly solve, new ones. Knowledge of this sort cannot yield any clear *ex ante* distinction between problems that can be solved and those that cannot, such as would exist if the available knowledge consisted merely of an easily accessible file of achieved solutions, with no possibility of expansion. Rather, it permits a range of judgments of widely varying reliability as to the likelihood that available problem-solving techniques and solution fragments will prove sufficient in the various cases posed.

Thus, decisions near the knowledge frontier are made in the face of a lot of haziness about the details and significant uncertainty as to how the general picture will develop down the road. Decision makers are generally aware that the best, and often the only, way to get a clearer picture of what lies ahead is to move further down the road.. Effort devoted to assessing the possibilities for progress from a given vantage point therefore has an opportunity cost in terms of progress itself – because attention and cognitive powers are scarce, if for no other reason. Thus, decisions are made in a haze that arises partly from exogenous constraints, but is partly a chosen response to a recognized tradeoff between thinking and doing, between analyzing prospects and making progress.

A closely related issue here is the relationship between knowledge of possibility and of impossibility. In standard theory, the production set apparatus answers both types of questions. The absence of the affirmative knowledge required to produce **q** from **x** is logically equivalent to "(**x**, **q**) is not in the production set" and hence to "(**x**, **q**) is technically impossible." Change the sharp boundary to a hazy frontier and some new issues appear. A high-confidence claim of possibility is one thing, a high-confidence claim of impossibility is quite another, and there can be a lot of room in between.

The distinction is significant because some economic propositions depend on affirmative assumptions of possibility, while others rest on assumptions of *im*possibility. For example, using arbitrage arguments to show that some price relationships will not be observed is a familiar exercise that depends only on the availability (to numerous actors) of the arbitrage actions referred to – e.g., buy here, ship from here to there, sell there. Consider, by contrast, the prediction that cost increases will result from a regulatory intervention that requires products to be modified in ways that buyers generally desire. Confidence about this conclusion rests on confidence about what firms *can't do*. The claim is, if firms could achieve these modifications at zero incremental cost, "they would have done it already" – since buyers attach positive value to the change. If one cares about the level of confidence that predictions deserve, and also acknowledges that the frontiers are hazy, one notices that quite different approaches are involved in assessing the confidence to be placed in the two different sorts of propositions.

In particular, the regulatory example demands some assessment of how the change affects the incentives to explore previously untried approaches that lie in the hazy zone. In that region, analysis of incentive effects falls under the heading of "induced innovation." The shaping effect of economic incentives is not limited to their effect on the optimal choice from among existing production methods. Rather, relative prices and other incentive considerations shape the emergence of new production methods in a variety of ways and at a variety of time horizons. Energy economics is a domain broad enough to illustrate the full spectrum of possibilities here: consider the impact of changing views of the future prices of fossil fuels on the choices made by both public and private decision makers. Near the existing knowledge frontiers, auto manufacturers making design and development choices attend to the implications of gasoline prices and vehicle gas mileage for the demand for their products. At the outer extreme, decisions regarding expenditure on basic research relevant to alternative energy sources, e.g. solar, nuclear, fusion, tidal, are likely to be affected – though of course it may take a substantial change in images of the future to produce a measurable response.

There is a substantial theoretical literature on induced innovation, though the attention devoted to it falls far short of what it deserves. Most of it models the problem in a recognizably mainstream style, with a logic that parallels quite closely the standard analysis of choice of technique.[15] This is unfortunate, since it means that the limitations of standard production theory are extended into a domain that could potentially afford an escape from them. Evolutionary modeling in the area has mostly been done incidental to other purposes, and has not been featured in its own right. It at least illustrates the possibility that the sorts of incentive effects that common sense makes "obvious" can by supported rigorously under weak assumptions about rationality and the character of the alternatives. Much, however, remains to be done in this area, as is further discussed in the concluding section of this paper.[16]

Distributed character. A third distinctive attribute of productive knowledge is that it frequently resides in work groups or organizations rather than in individuals. This is not simply a matter of the same knowledge being held by several individuals, although such common knowledge may play an important role. Neither is it adequately captured by the image of complementary specialized skills being coordinated in the execution of a "recipe" – because the recipe, or at least much of the knowledge required in its execution, often exists only as it is evoked in the actual activity of the group. It is crucially a matter of *distributed* knowledge, i.e., of complementary parts of the same functional unit of knowledge being held by several individuals and applied in a coordinated way – the coordination itself being a crucial aspect of the knowledge. Such a "functional unit" might comprise, for example, the ability of a cockpit crew to get an airliner safely to its destination,[17] of a surgical team to perform an open-heart operation, or of a string quartet and a piano player to perform "The Trout". It is the obvious significance of these sorts of functional units of organizational knowledge and performance that makes the idea of distributed knowledge inescapable. We would find little use for a theory that was either inapplicable to such performances or pretended that they involved no knowledge, or admitted the knowledge content of the performance but offered no answer to the question of where the knowledge resides, or obstinately pretended that in such cases all of the knowledge is held in the mind of a single individual, or posited that the successful coordination is the fruit of a "recipe" that somehow exists apart from the individual skills. Accepting the view that knowledge can reside in a group is in this sense a "forced move."

There is nothing either novel or mysterious about the notion of productive knowledge residing in a group rather than in individual minds. Primitive peoples have long used collective hunting methods involving specialized individual roles. The question of how such distributed knowledge can arise in a group has multiple answers, but one is particularly clear and straightforward: such knowledge is the product of shared experiential learning. Just as practice allows an individual to improve in the performance of a complex skill through improved coordination, so the shared practice of a group permits patterns of multi-person coordinated action to arise, improve and stabilize. In both cases, "procedural memory" – in which micro responses are evoked sequentially by cues arising in large part from the emerging action itself -- is a key mechanism of learning (Squire 1987; Cohen and Bacdayan 1994). In the individual case, those cues move intra-personally as nerve impulses from sensory systems, including the kinesthetic sense, to the brain. In the group case, cues move interpersonally and include verbal messages as well as mutual observation. Although individual minds are a principal locus of storage for distributed knowledge, the coherent enactment of the functional unit of knowledge is a phenomenon that occurs at the level of the group. The more sharply differentiated the individual roles in the group performance, the more obvious it is that there is no individual who knows what the group knows. Whatever standing a group of diverse specialists may have achieved in the development and retention of a functional unit of some type of knowledge, it cannot be contested by any individual mind but only by another experienced group encompassing the same specialties.

The reality of distributed knowledge, and of the shared experiential learning that typically produces it, seems obvious and incontrovertible. The aesthetic appreciation of high levels of interpersonal coordination is, for example, part of the enjoyment of watching a great team play soccer or basketball. Yet this obvious and important point about productive knowledge is often resisted. Intellectual inclinations toward reductionism are no doubt part of the reason.[18] Strongly amplifying those philosophical dispositions is the practical reality that distributed knowledge is a phenomenon that is very awkward to capture in a simple analytical framework. It brings history into the picture and implies that an adequate representation of individual actors would have to record not only what they have done in the past, but with whom they have done it. This complexity contrasts sharply with the standard economic theory of production, which adopts the very convenient fiction of homogeneous input categories, but is thereby forced to adopt the awkward complementary fiction that a crucial part of the distributed productive knowledge in an experienced team of input suppliers resides someplace outside of the team.

In recent years, the fact that productive knowledge often resides in groups rather than in individuals has received increasing attention both from business firms and from scholars outside of the economics discipline. There has been a striking degree of mutual reinforcement between the interest in these issues on the business side, driven by the practical concerns of an increasingly knowledge-based economy, and academic scholarship.

As is true of knowledge held by individuals, the productive knowledge distributed in groups is of a variety of types. There is a spectrum ranging from highly "automatic" performance of organizational routines, analogous to the exercise of tacit psycho-motor skills in an individual, through coordinated attacks on well-posed problems, and on, to solutions of ill-posed problems and attempts at high creativity. Verbal communication among group members plays an increasing role as the attributes of the situation call increasingly for highly deliberate and/or creative response; it may play no role at all when response is automatic. In the latter case, the reality of knowledge at the group level is manifested in the quality of the coordination visible in action, but at the more creative end of the spectrum the efficacy of verbal communication is an important key to both the enactment and the assessment of group knowledge. Groups that communicate well tend to make better choices of action; the reality of group knowledge is manifest in the communications that precede the choices as well as in the choices themselves.[19]

Ethnographic research has produced significant evidence in support of this assessment – although, by their nature, ethnographic studies are qualified by the absence of experimental control. A particularly influential contribution was Julian Orr's detailed ethnographic study of the work of service representatives for Xerox Corporation (Orr 1996). The specific activities of the service representatives are triggered largely by the urgent need for repairs to malfunctioning copiers, implying that their work is highly unpredictable in timing and content. At that level of analysis, it is not governed by routine. A majority of the individual repair tasks are, however, non-problematic in that they are covered by the company's directive documentation, are within the high frequency domain of the technicians' skills, or both. The knowledge applied in addressing the less common failure modes is of a different character. Orr recounts how the technicians use stories to organize and deepen their collective understanding of the ambiguous behavior of the complex machines they deal with. The totality of that specialized knowledge resides in the community of technicians, and is mobilized by the individual technician or a small group when attacking difficult problems. Status in the community depends in part on making contributions to the common knowledge pool and demonstrating awareness of what it contains.[20]

This image of knowledge residing in a "community of practice" has been highly influential, in part because of its extension and effective explication to a broader audience by the work of (Brown and Duguid 1991; Brown and Duguid 2000).[21] Its significance derives partly from its relation to a particularly important segment of the knowledge problems that organizations encounter, and partly from its rich psychological, sociological and organizational aspects. An important open issue is whether an effective community of practice is can be created deliberately by management – as many companies have tried to do – or whether they are necessarily spontaneous and informal, open at most to acceptance and some facilitation by management. In any case, communities of practice represent only a subset, albeit an important one, of the broader phenomenon of distributed knowledge.

Apart from displaying the phenomenon of organizational knowledge distributed in a community of practice, Orr's study documented several other points consistent with the general view of productive knowledge presented here. The work of the technicians is not understood at higher levels in the organization, or, in particular, by the creators of the repair documentation. In effect, the organization's functioning depends on productive knowledge of which management is unaware and on practices that often contravene management's efforts to control behavior.[22] Particularly striking is the fact that the management seems not to recognize that keeping the customers happy is an important and challenging part of the job of a technician, which Orr observed to be quite distinguishable from the challenge of repairing machines. Per the discussion above, the true end point of the service production process is not in the state of the machine, but in the customer's head. Perhaps if managers thought about it carefully, they would agree that keeping the customers happy warrants a lot of attention and even a page or two in the documentation – but in fact they apparently do not think about the technician's job in that light.

Although the work group or community of practice is an appropriate focal point for arguing the reality of distributed knowledge, it is clear that knowledge in business firms is also "distributed" on a larger scale. Neither the theoretical understanding nor the empirical evidence about this is particularly strong overall, but there are some areas of strength. A very substantial research effort has gone into assessing production practices in automobile assembly operations, with particular attention to the diffusion of "lean production" and to the practices of Toyota in particular (e.g., (Womack, Jones et al. 1991), (Adler 1993); (Adler and Borys 1996); (Coriat 2000; Florida and Kenney 2000; Fujimoto 2000). While this research provides further illustration of the points about micro-level knowledge made in the work of Orr and others, it also shows that there are large systematic differences among firms in their ability to orchestrate the productive application of distributed knowledge. Indeed, it seems that the important overall implication of this work is that a multi-level analysis of productive knowledge is ultimately a necessity. To put it another way, it is hard to find a potential barrier to the movement of knowledge that does not function significantly as an actual barrier: national boundaries matter, firm boundaries matter, plant boundaries matter and even within-plant boundaries matter (Dyer 1998). To explicate the functioning of such a complex, multi-level system of distributed knowledge stands as a major challenge for theorists.

3. Replication

As noted above, the question of the pragmatic validity of productive knowledge is, at its most basic level, a question of whether the use of a functional unit of knowledge can be extended in time and in space – that is, whether a productive performance can be repeated, or whether a functionally equivalent performance can be accomplished at a different site. It is tempting to jump to an affirmative conclusion: standard production theory certainly suppresses doubt about this, and perhaps "common sense" does as well. There is, however, much to be learned by letting the doubts resurface.

The "obvious" feasibility of temporal and spatial replication is a conclusion that is derived, first, by abstracting away complications that are often significant in real production situations. The standard theory dispatches many of these complications by the assumption that inputs are conveniently sorted into categories that are both homogeneous and readily recognized, others by the assumption that productive knowledge has a central (though unspecified) locus that is somehow inherently associated with the firm itself, and then by simply ignoring other issues. In the background is, perhaps, an implicit reliance on the "repeated scientific experiments" view of the nature of productive knowledge, or (more realistically) on the unmentioned problem-solving capabilities that fill the knowledge gaps when conditions change. As for the common sense appraisal, it is probably dominated by the familiar experience of seeing similar operations in different places – without looking too closely and without, of course, any real assessment of whether it is "the same knowledge" that is being used. More careful scrutiny of the way knowledge figures in these "simple" issues of temporal and spatial replication deserve a prominent place on the initial agenda of an evolutionary approach to production.

Standard theory. In standard production theory, something akin to explicit discussion of replication arises in the context of formal axiomatic treatment of production set properties. More specifically, it appears in statements that seek to motivate the assumption of *additivity,* and a review of this standard analysis provides both a starting point and a contrast for an evolutionary approach.

Additivity says that the sum of feasible production plans is itself feasible. If it is known how to produce output vector $\boldsymbol{q^a}$ from input vector $\boldsymbol{x^a}$, and also how to produce $\boldsymbol{q^b}$ from $\boldsymbol{x^b}$, then it is known how to produce $(\boldsymbol{q^a} + \boldsymbol{q^b})$ from $(\boldsymbol{x^a} + \boldsymbol{x^b})$. In the special case when $\boldsymbol{x^a} = \boldsymbol{x}^b$ and $\boldsymbol{q^a} = \boldsymbol{q}^b$, the conclusion is that $2\boldsymbol{q^a}$ is producible from $2\boldsymbol{x}^a$—doubling of a feasible plan yields a feasible plan. It is this form of the proposition that has often been advanced in less formal discussion, as in, for example, the Wicksteed passage about exactly repeating a particular production plan for wheat, or in the following statement by Hahn:

> The common sense proposition that the duplication of a given industrial process (together incidentally with a duplication of entrepreneurs or entrepreneurial function) will lead to a doubling of the output is unassailable....
>
> If two identical entrepreneurs set up two identical plants, with an identical labor force, to produce an identical commodity x, then the two together will produce twice the amount of x that could be produced by one alone.[23]

In a similarly confident vein, (Arrow and Hahn 1971) more recently said "the assumption of additivity ... can always be defended and indeed made essentially tautologus" (p. 61).

Lurking well beneath the surface here are some significant issues involving the relationship of additivity to knowledge and to the spatial distribution of productive activity. Note first that the image suggested by Hahn's two identical plants or Wicksteed's notion of conditions "exactly repeated" is an image of *spatially distinct* but otherwise identical production activities. Since production is understood to involve input and output flows, or amounts over a designated time interval, doubling of production must be understood as a concurrent replication of activity – not, for example, doing the same thing over again. It is then plain enough that a second manufacturing plant or a second wheat field cannot be in precisely the same location as the first, and also that nothing resembling a replication of the original activity could be accomplished by trying to squeeze twice the activity into the same physical space. In short, the additivity axiom must be understood as relating to the aggregation of the results of activity occurring in different places. Hahn's example of the two plants makes it particularly clear that this is the correct interpretation . In this sense, the additivity axiom seems to say little more than "addition works," and perhaps that is why Arrow and Hahn said that it can be "made essentially tautologus."

There is, however, a limit on how "different" the places can be, which Hahn does not emphasize. Since location generally matters, the spatial distinction between the sites is potentially of economic significance, in which case it might not be meaningful to add quantities across sites. Adding California lemons and Italian lemons is as erroneous as the proverbial adding of "apples and oranges," so far as economics is concerned. If transportation costs are small relative to the input ant output values, because distances are short or for other reasons, the addition is justified.[24] Yet, if the distances are very short, it seems possible that the validity of the conclusion might be undercut because of physical interference between the sites. Might not production do something to the local environment in terms of air quality, noise, vibration or other considerations that could affect similar activity nearby? It seems that, in addition to its implicit relationship to location economics (implying distances are short), the additivity axiom also to has an implicit relationship to the topic of "externalities" (implying perhaps that distances can't be *too* short).

This is a lot of implicit theorizing to pack into one little axiom. The hidden complexities are, however, aspects of a more basic point: the usual production set formalism is simply not a very helpful tool for probing the relationship between "states of knowledge" and production possibilities. (The historical discussion above argues that this is an understandable consequence of the developmental trajectory of standard production theory.) The proposition that is really at issue is that the use of existing productive knowledge can be extended geographically. This is a proposition with great plausibility on its face, such that it is often taken for granted, and one that seems crucially important to understanding the historical development and current functioning of the economy. It is hence the sort of proposition that is worth deep and careful examination. The additivity axiom is a shallow and imperfect abstraction that does more to obscure the issues relevant to this proposition than to illuminate them.

The challenges of replication. To explore the issues more carefully, it is best to face the central issue directly. Abandoning the grandiose ambition of trying to characterize the structure of everything that is feasible (regardless of whether it is happening), begin with the supposition that some productive activity corresponding to (**x**, **q**) is *actually being conducted* *successfully* at site A. What does that imply about the possibility that a similar success can be achieved at site B? More specifically, the question is whether the cost of achieving similar success at B is significantly reduced by the fact that (in some sense) the requisite knowledge has apparently been developed and is in actual use at A.

The possibility of such cost reduction seems to be implicit in the most basic teaching about information (or knowledge), which is its property of being "non-rivalrous in use". The knowledge required to produce at A is not used up or degraded by using it also at B. The question is what this proposition has to do with the concrete task of making success at B happen, and with the cost of accomplishing that task. To isolate analytically the specific issues of extending the use of knowledge, the assumption here is that, perhaps after some initial period, the specific resources involved in production at B are entirely different from those at A. (It is not a case, for example, of having key workers at A work a second shift at B.) Also set aside for this discussion is the issue of temporal replication -- in general, temporal replication presents attenuated versions of the same issues that arise in spatial replication. Finally, recognize that such replication will not happen without some organized managerial effort; the relevant actor will be called "top management". The word "top" serves only to flag the point that there are typically other management layers below the level from which the replication effort is initiated, and these generally attenuate the knowledge of operational details held at the "top."

As the above discussion of pragmatic validity suggests, the premise that productive activity is going forward successfully at A actually has only limited implications for the state of productive knowledge at A, particularly as regards depth of causal understanding. It has still less implication for the feasibility of successful transfer from A to B for, given the distributed character of productive knowledge. Much of organizational knowledge is largely tacit, in a sense analogous to the tacit nature of individual skill: a great deal gets done without the "conscious awareness" of top management, even if the knowledge at lower levels is quite explicit. Even if systematic and sophisticated attempts at codification are made, there is inevitably a limit to the coverage of details. Many of these may be known to individual workers and describable by them, but not known to fellow workers, supervisors or managers. Others may be known but tacit, impossible to describe in a way that is helpful to a learner. Some may be secrets, cherished for their contribution to making work easier, and some may be contrary to the prevailing formal rules. Finally, it is well established in the psychology literature that there can be "learning without awareness;" people can learn to adapt their behavior effectively to varying stimuli without ever being consciously aware either of the stimulus conditions to which they are responding or of the fact of their adaptation (Lewicki, Hill et al. 1988; Reber 1989) There is little reason to doubt that this type of learning goes on in the natural settings of business activity as it does in the psychology labs.

Thus, if management wishes to replicate at B the success it is having at A, a first challenge it faces is to devise methods to evoke the transfer of details of which it is unaware, including some of which nobody is aware, and some that represent information that is in various ways "impacted."[25] This is a tall order that can never be filled completely; the gaps in the transfer will be filled by independent re-invention.

The literature of situated cognition points to the fact that the productive knowledge exercised at A exists in an intimate interconnection with the context of activity at A, both in its designed and coincidental features. Of particular importance here are the "artefacts," tools and/or equipment used in the work (Hutchins 1995; Hutchins and Klausen 1996). In contemporary work settings, crucially important knowledge is often embedded in the equipment, and the local understanding of its functioning is narrowly instrumental. Computers and software are the most ubiquitous and familiar examples of this issue, but there are "low tech" examples as well. The spatial organization of activity at the micro level affects patterns of interaction and communication, and this can affect outcomes in ways that are not fully recognized or understood (Bechky 2001).

Here, then, is a second challenge. While top management can attempt to establish a local context at B identical to that at A, a variety of obstacles can arise. The site layout may have to be different, perhaps because of its position relative to transportation corridors. Perhaps new generations of computers and software have succeeded those installed at A. The latter may represent progress and ultimately be economically advantageous overall, but it is not advantageous in the specific terms of achieving economies by re-using the productive knowledge that exists at A. From that viewpoint, new equipment is a complication. Again, a need for reinvention, relearning and perhaps creativity is implied.

Next, as discussed in relation to validity, there are the various relationships of the activity to its external physical environment. Depending on the character of the specific activity, there may be significant effects from the weather, from air quality and atmospheric pressure, from noise and vibration levels, the electromagnetic radiation environment, and so forth.[26] To the extent that these influences remain constant at A, their causal significance is likely to have gone unrecognized. If the environment at B is significantly different, the relevance of productive knowledge imported from A can be undercut in ways that are mysterious, at least initially. Learning and problem solving are called for.

Finally, the success of the activity at A depends on the attributes and behaviors of exchange partners encountered there – customers, employees and suppliers. While these may be influenced through the prices (wages), specifications and monitoring arrangements put forward, the specific responses to these influences are ultimately up to the exchange partners and not to the focal enterprise. Some of the causal principles governing such responses may be understood at a crude level, and careful monitoring may be effective in limiting variability. Nevertheless, important details often remain quite obscure – notwithstanding the substantial research efforts devoted to such questions, both in business firms and academe. Do customers read the warning labels on hazardous chemicals and behave accordingly? Do they use products only in recommended ways? Do the assemblers follow the drawings? Can they?[27] How serious is the adverse selection that inevitably occurs when the organization is in some distress and employees with good alternatives quit? To the extent that operations at B entail interactions with different exchange partners – which they necessarily do, at least so far as employees are concerned – differences in the populations of partners encountered consist of another reason why things might not function at B the way they do at A. Once again, the ability to operate successfully at A clearly does not entail possession of the sort of knowledge required to cope with the different circumstances at B. And once again, there is an implied requirement for new learning.

In some cases, top management may be totally unaware of the fact that considerations in these various categories are relevant to success. For example, this is inevitably so for procedural details that have been learned without awareness. In many cases, however, it is fairly obvious that the success of an activity depends in various ways on features of its context. Management may understand very well, for example, that the skills and personalities of employees have considerable bearing on the results achieved. What is unlikely to be known, however, is the size and interrelations of the tolerance intervals – how slow or fractious can the employees before success is actually undermined? An abstract statement of the general issue here is this: there is a region in the multidimensional space of contextual variables in which success can be achieved, but operations at the initial site neither entails nor reveals more than a little knowledge of where the boundaries of that region lie. From a performance point of view, it is of course a good thing if contextual variables stay, with high reliability, well within the region where success is possible. But the reliability that sustains successful performance also sustains causal ambiguity. If understanding were the goal, it would actually be better to experience the response to wider variation in the variables.

Frequently, new sites are chosen that are highly dissimilar from the original in some dimensions, because they are expected to be highly advantageous in other dimensions -- e.g., as suggested above, plant sites of preferred shapes may not exist in preferred locations, or retails locations that are highly attractive in terms of customer flows and income levels may be ones where it is difficult to attract employees at acceptable wage levels. This means that the shallowness of causal understanding, which remains latent in the successful operations at site A, can easily frustrate the necessary efforts to adapt to the different circumstances at B.

Replication in practice. The above discussion suggests that replication is potentially quite challenging. It is not necessarily viewed as such, however. Often, managements seem to take quite a relaxed attitude toward such challenges. While the relaxed attitude may be justified in some cases, it is arguable that many managers still have a good deal to learn about the subject.[28] Where special circumstances force the issue to management attention, such as the technological peculiarities of semiconductor production (Intel), or McDonalds' strategic devotion to a uniform customer experience, we do see managerial practices that are consistent with the general picture offered above.

The example of Intel's Copy EXACTLY! policy (McDonald 1998) is particularly valuable because it rests precisely on the recognition that there is more productive knowledge implied in achieved high yields than the organization can capture in the form of comprehensive causal understanding of the method in use. It exemplifies almost perfectly the notion of reliance on the knowledge embedded in the "template" of an existing achievement, as put forward by Nelson and Winter (Nelson and Winter 1982), pp. 119-120. Since it eschews adaptation and improvement as a matter of policy, this approach might be considered to be at odds with the above emphasis on the inevitable role of new learning and problem solving in achieving successful replication. This is not at all the case; the point is to control the *amount* of new learning and problem solving required. Every decision to create a difference between the sites makes an addition to an invisible list of unintended and hidden differences that will occur in spite of the policy. Interaction effects tend to make complexity rise exponentially with the number of discrepancies to be dealt with; it is better to keep the list as short as possible. There will be no shortage of problems to solve.

There are several reasons why the problematic aspects of replication often go unnoticed. A basic one, surely, is that exploiting existing knowledge from the original site is rarely an end in itself. The objective is successful operation at the new site, regardless of the mix of replication and new problem solving, and it is the time and cost of the combination that matters. The opportunity to exploit newer technology at the new site often tips the balance away from replication – or seems to. Further, the very fact that there are differences between the sites softens expectations concerning the level of performance achievable at the new one, so the occasion to question the success of the replication may not arise.

The empirical information on replication processes is relatively sparse and scattered, but it speaks quite uniformly on the point that significant costs and managerial challenges are involved. Szulanski's study of Banc One (Szulanski 2000) provides specifics concerning time, costs and challenges. More importantly, it provides a uniquely systematic and comprehensive view both of the specific processes involved in the development and application of a strategy of spatial replication in U.S. banking – in this case, via the conversion of newly affiliated banks to the common systems and products of the bank holding company. In the relevant time period, Banc One was an industry leader both in its ability to carry out conversions successfully and to improve earnings by the application of its common systems. The overall conversion process for a new affiliate took about six months, but culminated in a single frenetic weekend that separated the "before" and "after" identities of the affiliate from the customers' point of view. A successful conversion is one that is a "non-event" in the perspective of all but those concerned with carrying it out, but particularly from the viewpoint of customers. Banc One had many successes, but a few traumatic conversions were key learning episodes. Szulanski observed that Banc One became increasingly ambitious as it learned and routinized its conversion process, taking on conversions of increasing scope while failing to fully address the implied scheduling problems for its conversion capabilities. A major conversion effort ultimately hit multiple snags, triggering new learning efforts directed to scope and scheduling issues.

4. Production Theory Evolving

That knowledge and information are not exhausted by use is a kind of economic magic, a cheerful exception to the manifold scarcities that give the dismal science its name. The insight that accords this magic a central role in the processes of economic development is fundamentally on target – but the magic is not as simple as some suggest (e.g., (Romer 1993). To extend the use of existing knowledge in time and space is not at all the trivial matter it is often made out to be. A first point illustrated by the above account of replication is the way words like learning, problem solving, creativity and management inevitably come into play in a realistic discussion—even in relation to this, perhaps the simplest of all major issues in the economics of productive knowledge. Those words signal, in turn, the presence of costs and organizational capabilities that are typically overlooked by economists, and hence of potentials for competitive superiority and inferiority that are also neglected.

To recognize the scope for new learning, problem solving and creativity implicit in the "simple" idea of spatial replication is to recognize the magnitude of the conceptual gap that separates the evolutionary formulation of these issues from the standard apparatus, with its sharp distinction between the "known" and "unknown" techniques. There is a very basic contrast between, on the one hand, behavior that is learned and practiced, and on the other, behavior that may be improvised or planned – but in any case is not practiced. This contrast is fundamental to evolutionary economics in general, and to evolutionary production theory in particular. Representing it, reflecting on it and exploring its many intricacies is a large part of what the evolutionary program is about.

A second point that is highlighted by the replication discussion is the role of the individual, functioning productive unit – the "establishment", in statistical parlance – as a locus and repository of productive knowledge. Near the beginning of their discussion of organizational capabilities, Nelson and Winter suggest in a cautionary footnote that the concepts presented there may be more relevant at the establishment level than they are at the level of the multi-unit business firm (Nelson and Winter 1982), p. 97. In other contexts, however, evolutionary theory is (quite appropriately) associated with the idea that *business firms* are repositories of productive knowledge, or "*business firms* are organizations that know how to do things" ((Winter 1982; Winter 1991) respectively, emphasis supplied). There is an issue here, and an important agenda on which only limited progress has been made. On balance, that progress tends to support the validity of the footnote warning, and suggests that the relative importance of the establishment and firm levels as repositories of knowledge is something that evolutionary economics badly needs to address. Intra-firm homogeneity of method across establishments is definitely not something that just happens; it is something that happens when managements work hard to achieve it and devote substantial resources to the task. This is the lesson equally from Intel and McDonalds, and it is the lesson in the recent enthusiasm for "internal transfer of best practices." Regarding the latter, the starting point is the observation that valuable intra-organizational variety arises naturally; the experience is that it is often possible, but rarely easy, to exploit that variety and improve the average by transferring practices (Szulanski 1996). Giving the establishment level its appropriate place in the productive knowledge picture does not deny the importance of the firm level, but it certainly implies that there is a major challenge in creating a picture that is appropriately balanced and indicative of the great diversity of situations. The fact that there is an active market in establishments, sometimes grouped into business units or whole firms, only enhances the interest and difficulty of that challenge.

The overall agenda facing evolutionary production theory is of course much larger than the territory explored in this paper. A plausible next step, much in the spirit of the above discussion of additivity and replication, would be to examine the issue of returns to scale in production – meaning, in substance, the issue of *increasing* returns in production. These phenomena are excluded in axiomatic treatments of "well behaved" production situations by the assumption of divisibility (or, more directly "non-increasing returns to scale.) From the same analytical attitude at which one can observe that the additivity axiom has a certain common-sense plausibility to it, one can observe also that the divisibility axiom does basic violence to the truth. Human beings, structures, machines and the process of getting from one place to another all have a fundamental indivisibility. The economic consequences of such indivisibility may be diminished in some cases because, e.g., of the time-divisibility of service flows, or because similar equipment can be designed at different scales – but rarely are the economic consequences of indivisibility banished entirely. Nevertheless, the assumption of divisibility can be accepted as a method for approximate analysis, and it is clearly a better method in some cases than others.

The problem is to understand the "others." A counterpart and complement to the above replication discussion would take as the starting point an ongoing activity successfully producing **q** from **x**, and ask how output capacity can be increased to 3**q.** One answer would be: by creating two replicas of the original and producing 3**q** from 3**x**. The crucial point is that there may be a better alternative, because additional design freedom is conferred when scale is increased. In particular, it is often possible to achieve higher efficiency in equipment units of greater physical size, for underlying reasons of physics and geometry. If replication or partial replication is also possible, the additional design freedom should tip the balance to increasing returns and allow the additional 2**q** to be produced from something less (or cheaper) than 2**x**. Although this picture is subject to necessary qualifications, it is hardly contestable that there is substantial reality to it (Levin 1977),(Scherer 1975). A major part of the analysis that is needed here involves exploring the sources and nature of the design freedom, and the way its exploitation is complemented by replication. The major objectives include better understanding of the contribution of increasing returns to rising productivity, its relation to technical change, and the reasons for the diversity in its role across activities.

Beyond the analysis of increasing returns, there lie many similar projects of narrower scope. As suggested in the introduction, the overall objective is to develop an economic understanding of production that is deeply grounded in the engineering and managerial realities of the phenomenon. The possibility of doing this, without getting bogging down in excessive detail, derives from the existence of broadly applied heuristic approaches to production challenges – a category of which"replication" and "exploiting the design freedom conferred by larger scale" are prominent examples. There are heuristic approaches of somewhat narrower scope – variously known as "trajectories," "paradigms," or "industry recipes" – that shape progress in particular sectors in specific historical periods (Rosenberg 1969; Nelson and Winter 1977; Dosi 1982; Spender 1989). These too can be the object of useful analysis. By understanding the logic and practice of a wide range of these heuristics, evolutionary economists could make progress toward sophisticated methods for appraising the likely course of future technological change. Such methods should facilitate the joining of economic analysis and expert appraisal of trends and emerging opportunities in technology – analysis of a type that is obviously needed but rarely, if ever, attempted.

Skipping forward across a wide range of subjects of intermediate size and complexity, I conclude by sketching one very large problem that builders of evolutionary production theory ought to have in view. Environmental concerns, and particularly the problem of global climate change, raise the related problems of designing regulatory interventions and estimating the economic response to such interventions. The relevant time horizons for most of these concerns are (we hope) quite long – though not long enough to warrant complacency. The problem posed involves "induced innovation" on a grand scale, operating at many levels and loci in the economic, technological and scientific systems, and posing challenges of policy, management and institutional design. The capacity to analyze complex problems that lie in the domain of the "hazy frontiers," where evolving technological prospects mix with uncertain and changing economic incentives, would be very valuable in dealing with this problem. Perhaps it is indispensable. Economics has thus far done little to help build this sort of capacity; the way standard economics defends its boundaries and seals out alien problems has been a factor limiting its contribution. By contrast, evolutionary theory has few, if any, commitments that are put seriously at risk by ideas or facts from other disciplines. As this paper attempts to illustrate, it has open frontiers, lives with other disciplines in what is recognizably the same world, and has much to gain from trade. The evolutionary approach offers the kind of economics, and should provide the kind of production theory, needed to address the problems that progress brings to the planet.

References

Adler, P. (1993). "The learning bureaucracy: New United Motor Manufacturing, Inc." Research in Organization Behavior **15:** 111-194.

Adler, P. S. and B. Borys (1996). "Two types of bureaucracy: Enabling and coercive." Administrative Science Quarterly **41**: 61-89.

Arrow, K. J. (1974). The Limits of Organization. New York, Norton.

Arrow, K. J. and G. Debreu (1954). "Existence of equilibrium for a competitive economy." Econometrica **22**: 265-290.

Arrow, K. J. and F. H. Hahn (1971). General Competitive Analysis. San Francisco, Holden-Day.

Atkinson, A. and J. Stiglitz (1969). "A new view of technical change." Economic Journal **79**: 573-578.

Bechky, B. A. (2001). Sharing meaning across occupational communities: The transformation of knowledge on a production floor. Philadelphia, Working Paper, The Wharton School.

Brown, J. S. and P. Duguid (1991). "Organizational learning and communities-of-practice: Toward a unified view of working, learning and innovation." Organization Science **2**: 40-57.

Brown, J. S. and P. Duguid (2000). The Social Life of Information. Boston, Harvard Business School Press.

Carlson, S. (1956). A Study on the Pure Theory of Production. New York, Kelley and Millman.

Cohen, M. and P. Bacdayan (1994). "Organizational routines are stored as procedural memory." Organization Science **5**: 554-568.

Collins, H. M. (1985). Changing Order: Replication and Induction in Scientific Practice. London, Sage.

Coriat, B. (2000). "The "abominable Ohno production system": Competences, monitoring and routines in Japanese production systems." The Nature and Dynamics of Organizational Capabilities. G. Dosi, R. R. Nelson and S. G. Winter, Eds. Oxford, Oxford University Press**:** 213-243.

Debreu, G. (1959). Theory of Value. New York, Wiley.

Dosi, G. (1982). "Technological paradigms and technological trajectories." Research Policy **11**: 147-162.

Dosi, G. (1997). "Opportunities, incentives and the collective patterns of technological change." Economic Journal **107**: 1530-1547.

Dosi, G. and B. Kogut (1993). "National specificities and the context of change: The co-evolution of organization and technology." Country Competitiveness: Technology and the Organizing of Work. B. Kogut, Eds. New York, Oxford University Press**:** 249-262.

Dyer, J. H. (1998). Interorganizational learning and competitive advantage: Does supplier performance differ for different customers? Philadelphia, Wharton School Working Paper.

Engestrom, Y. and D. Middleton, Eds. (1996). Cognition and Communication at Work. Cambridge, Cambridge University Press.

Flaherty, M. T. (2000). "Limited inquiry and intelligent adaptation in semiconductor manufacturing." The Nature and Dynamics of Organizational Capabilities. G. Dosi, R. R. Nelson and S. G. Winter, Eds. Oxford, Oxford University Press**:** 99-123.

Florida, R. and M. Kenney (2000). "Transfer and replication of organizational capabilities: Japanese transplant organizations in the United States." The Nature and Dynamics of Organizational Capabilities. G. Dosi, R. R. Nelson and S. G. Winter, Eds. Oxford, Oxford University Press**:** 281-307.

Fujimoto, T. (2000). "Evolution of manufacturing systems and ex post dynamic capabilities: A case of Toyota's final assembly operations." The Nature and Dynamics of Organizational Capabilities. G. Dosi, R. R. Nelson and S. G. Winter, Eds. Oxford, Oxford University Press**:** 244-280.

Gould, S. J. (1980). The Panda's Thumb: More Reflections in Natural History. New York, Norton.

Grant, R. M. (1996). “Toward a knowledge-based theory of the firm.” Strategic Management Journal **17**(Winter special issue): 109-122.

Hahn, F. H. (1949). “Proportionality, divisibility and economies of scale: Comment.” Quarterly Journal of Economics **63**: 131-37.

Hutchins, E. (1995). Cognition in the Wild. Cambridge, MA, MIT Press.

Hutchins, E. and T. Klausen (1996). "Distributed cognition in an airline cockpit." Cognition and Communication at Work. Y. Engestrom and D. Middleton, Eds. Cambridge, Cambridge University Press**:** 15-34.

Kogut, B. (1991). “Country capabilities and the permeability of borders.” Strategic Management Journal **12**: 33-47.

Koopmans, T. C. (1977). “Concepts of optimality and their uses.” American Economic Review **67**: 261-274.

Lansing, J. S. (1991). Priests and Programmers: Technologies of Power in the Engineered Landscape of Bali. Princeton, Princeton University Press.

Lansing, J. S., J. N. Kremer, et al. (1998). “System-dependent selection, ecological feedback, and the emergence of functional structure in ecosystems.” Journal of Theoretical Biology **192**: 377-391.

Lave, J. and E. Wenger (1991). Situated Learning: Legitimate Peripheral Participation. Cambridge, Cambridge University Press.

Levin, R. C. (1977). “Technical change and optimal scale: Some implications.” Southern Economic Journal **2**: 208-221.

Lewicki, P., T. Hill, et al. (1988). “Acquisition of procedural knowledge about a pattern of stimuli that cannot be articulated.” Cognitive Psychology **20**: 24-37.

Lippman, S. and R. Rumelt (1982). “Uncertain imitability: an analysis of interfirm differences in efficiency under competition.” Bell Journal of Economics **13**: 418-438.

Maanen, J. v. and S. R. Barley (1984). "Occupational communities: Culture and control in organizations." Research in Organizational Behavior. B. M. Staw and L. L. Cummings, Eds. Greenwich, Conn., JAI Press. **6**.

Marshall, A. (1920). Principles of Economics, 8th ed. New York, Macmillan.

McDonald, C. J. (1998). “The evolution of Intel's Copy EXACTLY! policy.” Intel Technology Journal (4th quarter).

Narduzzo, A., E. Rocco, et al. (2000). "Talking about routines in the field: The emergence of organizational capabilities in a new cellular phone network company." The Nature and Dynamics of Organizational Capabilities. G. Dosi, R. R. Nelson and S. G. Winter, Eds. Oxford, Oxford University Press**:** 27-50.

Nelson, R. (1995). "The agenda for growth theory: A different point of view." Technology, Trade and the New Economy. J. d. l. Mothe and G. Paquet, Eds. Ottawa, University of Ottawa, Program of Research in International Management and Economy.

Nelson, R. R. (2000). The Sources of Economic Growth. Cambridge, MA, Harvard University Press.

Nelson, R. R. and S. G. Winter (1977). “In search of useful theory of innovation.” Research Policy **Vol. 6**(January 1977): 36-76.

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

Orr, J. E. (1996). Talking About Machines: An Ethnography of a Modern Job. Ithaca, NY, ILR Press/Cornell University Press.

Reber, A. S. (1989). “Implicit learning and tacit knowledge.” Journal of Experimental Psychology: General **118**: 219-235.

Romer, P. (1993). “Idea gaps and object gaps in economic development.” Journal of Monetary Economics **32**: 543-73.

Rosenberg, N. (1969). “The direction of technological change: Inducement mechanisms and focusing devices.” Economic Development and Cultural Change **18**: 1-24.

Ruttan, V. W. (1997). “Induced innovation, evolutionary theory and path dependence: Sources of technical change.” Economic Journal **107**: 1520-1529.

Scherer, F. M. (1975). The Economics of Multi-plant Operations: An International Comparisons Study. Cambridge, MA, Harvard University Press.

Shaiken, H. (1996). "Experience and the collective nature of skill." Cognition and Communication at Work. Y. Engestrom and D. Middleton, Eds. Cambridge, Cambridge University Press**:** 279-295.

Simon, H. (1995). "Bounded rationality and organizational learning." Organizational Learning. M. D. Cohen and L. S. Sproull, Eds. Thousand Oaks, CA, Sage**:** 175-94.

Spender, J.-C. (1989). Industry Recipes: An Inquiry into the Nature and Sources of Managerial Judgment. Oxford, Blackwell.

Squire, L. R. (1987). Memory and Brain. New York, Oxford University Press.

Szulanski, G. (1996). “Exploring internal stickiness: impediments to the transfer of best practice within the firm.” Strategic Management Journal **17**(Winter Special Issue): 27-43.

Szulanski, G. (2000). "Appriopriability and the challenge of scope: Banc One routinizes replication." The Nature and Dynamics of Organizational Capabilities. G. Dosi, R. R. Nelson and S. G. Winter, Eds. New York, Oxford University Press**:** 69--98.

Szulanski, G. and S. G. Winter (2002 forthcoming). “Getting it right the second time.” Harvard Business Review.

Wenger, E. (1998). Communities of Practice: Learning, Meaning and Identity. New York, Cambridge University Press.

Wicksteed, P. (1894). An Essay on the Coordination of the Laws of Distribution. London, Macmillan.

Williamson, O. E. (1975). Markets and Hierarchies: Analysis and Antitrust Implications. New York, Macmillan.

Winter, S. G. (1982). "An essay on the theory of production." Economics and the World Around It. S. Hymans, Eds. Ann Arbor, MI, University of Michigan Press**:** 55-91.

Winter, S. G. (1991). "On Coase, competence and the corporation." The Nature of the Firm: Origins, Evolution, and Development. O. E. Williamson and S. G. Winter, Eds. New York, Oxford University Press**:** 179-195.

Winter, S. G. and G. Szulanski (in press). Replication as strategy. Organization Science.

Womack, J. P., D. T. Jones, et al. (1991). The Machine That Changed the World: The Story of Lean Production. New York, Harper Collins.

Zadeh, L. A., Ed. (1975). Fuzzy Sets and their Applications to Cognitive and Decision Processes. New York, Academic Press.

Zollo, M. and S. G. Winter (in press). “Deliberate learning and the evolution of dynamic capabilities.” Organization Science.

Endnotes

* Particularly in its history-of-thought aspects, this work draws extensively from my 1982 paper, "An Essay on the Theory of Production" (Winter 1982). It also draws on more recent joint work with Gabriel Szulanski, some reported in (Winter and Szulanski in press) and some still in progress. I am indebted to Giovanni Dosi and Kurt Dopfer for the encouragement and guidance they provided for this undertaking.

---

[1] One example of such attention is (Atkinson and Stiglitz 1969). Their analysis, however, tends to further reduce one's confidence in the idea that the standard neutrality concepts for technological change correspond to categories of real economic mechanisms that might produce the corresponding types of effects.

[2] (Wicksteed 1894), p. 4.

[3] Ibid., p. 33.

[4] Carlson attributes the idea to Edgeworth, but the cited passage in Edgeworth is rather unclear.

[5] (Carlson 1956), pp. 14-16, emphasis in original.

[6] For citations and discussion, see (Koopmans 1977)

[7] (Arrow and Hahn 1971), p. 53.

[8] In using experimental science to anchor the high-validity end of a notional continuum, I am oversimplifying the situation in the interests of brevity. There is a substantial literature in the sociology of science that documents the fact that the validity issues that I here associate with (parts of) the productive knowledge domain also arise in (parts of) the domain of experimental science. (See, e.g., (Collins 1985)) The contrast between the domains is real, but as not as sharp as I here pretend.

[9] (www.nsf.gov/sbe/nuggets/015/nugget.htm, 8/14/01.)

[10] Of course, outcome variability *per se* is not necessarily inconsistent with valid and reliable productive knowledge. It could be the case that the variability takes the form of an additive random disturbance, complicating the problem of determining the best method without being causally intertwined with it. This is the sort of assumption statisticians like to make, but it is not the way that "randomness" actually appears in the sorts of deterministic but complex dynamical systems that constitute so much of our familiar reality.

[11] For example: "Despite the fact that the semiconductor silicon is likely the most intensively studied and well-understood solid material on earth, when it comes to making integrated circuits in large quantities, the pure white gowns and caps of production line workers might as well be covered with moons and stars of the type that bedeck the robes of the conjuror ... the causes of defects are not always understood, nor are they easy to control when they are understood." (A. L. Robinson, "New Ways to Make Microcircuits," *Science* 208 [May 1980]: 1019.

[12] (McDonald 1998), emphasis in original.

[13] The useful concept "causal ambiguity" was introduced into the economics and management literatures by (Lippman and Rumelt 1982), discussing obstacles to imitation.

[14] Or perhaps it should be "fuzzy frontiers". "Hazy" provides a better metaphor, but "fuzzy" would evoke, quite properly the concept of a "fuzzy (production) set." See, e.g., (Zadeh 1975).

[15] See (Ruttan 1997) for references and discussion.

[16] Responding to Ruttan, Dosi upholds the evolutionary approach to induced innovation and summarizes it effectively(Dosi 1997).

[17] See the ethnographic study of a cockpit crew by (Hutchins and Klausen 1996) in the volume edited by (Engestrom and Middleton 1996). Among other helpful studies in the same volume, see particularly (Shaiken 1996) on the "collective nature of skill."

[18] Herbert Simon (Simon 1995) stated "All learning takes place inside individual human heads; an organization learns in only two ways: (a) by the learning of its members, or (b) may ingesting new members who have knowledge that the organization didn't previously have." (1991, p. 176). He then goes on to qualify this substantially (though inadequately) by noting the relationships among the contents of the heads. But see Grant , who gives this quotation without the qualification, and says he is "dispensing with the concept of *organizational knowledge*." ((Grant 1996), pp. 112-113).

[19] (Arrow 1974) emphasized the importance of organization-specific "codes" as a factor in organizational efficiency.

[20] There are substantial parallels between Orr's account of the repairmen and the study of cellular network technicians by (Narduzzo, Rocco et al. 2000).Consistent with a point about frequency made by Zollo and Winter(Zollo and Winter in press), Orr notes (p. 109) that experienced technicians do not refer to the documentation when solving the most common problems. Thus, in the context of continuing activity, the role of codified knowledge is bounded on one side by a domain where it is unnecessary and on the other by a domain where it is inadequate. The situation is different in the context of new learning or knowledge transfer, where the "unnecessary" domain is replaced by one where codification is helpful.

[21] The term "community of practice" is usually associated with (Lave and Wenger 1991; Wenger 1998); see also (Maanen and Barley 1984) on "occupational communities." In his introduction to the Orr book, Stephen Barley comments that "…Orr puts the flesh of everyday life on Lave and Wenger's idea of a community of practice …." (p. xiii).

[22] For another empirical perspective on this important point, see the discussion in (Adler 1993; Adler and Borys 1996) contrasting the GM and NUMMI approaches at the Freemont, CA auto assembly plant taken over by NUMMI from GM. Clearly, efforts to evade control are not all equally meritorious from an efficiency point of view !

[23] (Hahn 1949)

[24] Awareness of this point is reflected in Debreu's classic *Theory of Value* (Debreu 1959) in his reference to "elementary regions" within which transportation costs are negligible (p. 29). When they offer data on, e.g., national production of Portland cement, economic statisticians are in effect treating the nation as such a region. Probably Debreu would be skeptical about this.

[25] The term "impacted information" is generally applied to situations where it is difficult to provide appropriate incentives for disclosure, as in the case of a piece rate worker who has developed a superior technique (Williamson 1975). Certainly this incentive issue is important, but it does not dominate the knowledge transfer scene to the extent that economists seem to imagine. For evidence, see (Szulanski 1996).

[26] Alfred Marshall's characterization of "land" as a factor of production is on point here. He recognized that it is not just the soil but the whole bundle of physical attributes specific to a location that is involved: "… the space relations of the plot in question, and the annuity that nature has given it of sunlight and air and rain." (Marshall 1920), p. 147). To evoke the potential relevance of environmental influence to non-agricultural production requires a substantially longer attribute list.

[27] In her ethnographic study of a semiconductor equipment firm, Bechky (Bechky 2001) noted that the assemblers were actually guided by the visible example of their objective, a prototype machine constructed by the technicians. Nominally, the assemblers were working from engineering drawings. Management did not appear to be cognizant of the important role the technicians were playing in mediating the technology transfer from engineers to assemblers.

[28] See (Winter and Szulanski in press). A further statement on this, with a stronger prescriptive message, is (Szulanski and Winter 2002 forthcoming).