10— The Division of Inventive Labor and the Extent of the Market

Timothy Bresnahan and Alfonso Gambardella

10.1— Introduction

Vertical disintegration of economic activity is the core of the division of labor as Smith (1776) and Stigler (1951) saw it. Stigler's conjecture was that *general specialties* were the economically important examples of division of labor. Not an oxymoron, the phrase has instead Stiglerian compactness and precision. A division of previously unified labor can create a "specialty." That specialty will be of economywide importance when its scope of application is "general," meaning that it is used by many kinds of customers. In the first and second industrial revolution, many general specialties exploited scale economies in production. Stigler thought of transportation examples like railroads and shipping, of financial examples like the London banking center, and of specialized production of intermediate materials (steel, chemicals) or capital goods (machine tools, electric motors, and lights). In modern times the creation of new general specialties continues unabated. Specialized science and engineering-based high-tech industries, broadly useful through the economy, lead this trend. Electronics, for instance, has been accompanied by sustained increase in specialization as hardware, software, networking, and so on, have become separate engineering subdisciplines practiced by specialists working in separate industries. Yet these are general specialties, selling the fruits of their invention to a wide variety of distinct types of customers. In short, both today and in the past, one of the most apparent effects of the division of labor has been the creation of whole new bodies of specialized

---

We thank the Alfred P. Sloan Foundation and Italian Consiglio Nazionale delle Ricerche under contract number 94.02070.ct10 for support, an editor of this volume, a referee, Ashish Arora, Victoria Danilchouk, Nathan Rosenberg, and Manuel Trajtenberg for comments and discussion.

knowledge and frequently whole new industries selling to many others in the economy.[1]

These general specialties, or GPTs, are an important source of economywide scale economies and growth (Bresnahan and Trajtenberg 1995; Helpman and Trajtenberg 1995).[2] Moreover the fact that GPTs, or more generally scale economies in invention, can drive growth, is a clear implication of the new growth theory (Romer 1986, 1990). But while the causal link from division of labor to growth through GPT seems clear, the modern theoretical literature has cast some shadows on how growth and the attendant increase in scale economies permit exploitation of the division of labor. Smith and Stigler argued that division of labor is limited by the extent of the market. By contrast, the main theorem in the modern literature shows that a large market will lead to vertical integration, *not* vertical disintegration and specialization (Perry 1989, sec. 7.3.) The literature has concluded that the problem lies in the Smith-Stigler treatment of *specialization*. In order to obtain a theory in which the Smith-Stigler conjecture is correct, it has then proposed a definition of specialization based not on scale economies but on scope diseconomies.[3] While this saves the theory formally, it has the very unhappy implication that division of labor is not a mechanism for positive feedback from growth through scale economies to more growth but rather a mechanism for ameliorating some of the costs of growth under *convex* (no scale economies) technology.

Another literature, in growth theory, has changed the definition of "specialization" in a different way in order to get a result with a Smith-

---

[1] Despite the fact that his famous example of the "pin factory" concerned the division of labor across workers within the firm, Smith himself clearly envisaged a much broader division across industries or activities based on deep and generalized knowledge bases. He first argued that improvements in machinery are sometimes made by philsophers or men of speculation: "who . . . are often capable of combining together the powers of the most distant and dissimilar objects." Then he noted that ". . . in the progress of society, philosophy or speculation becomes, like every other employment, the principal or sole trade and occupation of a particular class of citizens . . . it is subdivided into a great number of branches . . . and this subdivision of employment in philosophy, as well as in every other business, improves dexterity, and saves time." (Smith 1776, ch. 1)

[2] We will treat "general specialty" and "general purpose technology" as nearly identical in meaning. To the extent that there is a distinction, a GPT's body of knowledge or the people who know it are a general specialty.

[3] "The difficulty with the Stigler model occurs from trying to capture the notion of specialization. . . . An alternative view of specialization is that there are economies from doing a limited set of activities, rather than economies of scale from simply doing a lot of any one activity. . . . Specialization would mean diseconomies of scope across vertically related production processes" (Perry, 1989, p. 232). Perry directs our attention to the agency costs which might cause the scope diseconomies, a direction that the information-based theory of the firm has followed up carefully (Milgrom and Roberts 1992.)

Stigler flavor. Using monopolistic competition models, papers such as Romer (1992) show that a larger economy will have more product variety as the costs of specialized inputs can be spread out over more units. In this way they show that large market size causes *horizontal* specialization, a division of labor among firms or workers that make substitute goods. These models are obviously right, and the increasing product variety of modern consumer economies is clearly an important part of the growth process. Yet this analysis is unrelated to the *vertical* "specialization" and "division of labor" noted by Smith and Stigler and so important in modern high-tech industry.

In this chapter we argue that the problem does not arise from an incorrect definition of specialization but lies instead with the modern theory's definition of the *extent of the market*. Vertical integration theories have formalized the extent of the market as the output of a single good, asking when two production tasks will be vertically disintegrated. Our alternative definition views vertical disintegration as the introduction of a general specialty and the founding of a GPT. The point is that the extent of the market for a GPT is not only the volume of production in one of the sectors that applies it but also the number of distinct applications sectors that might apply it. Taking the extent of the market to mean breadth as well as depth connects the theory to the world. The general specialties noted by Stigler all had a distinctly "infrastructural" flavor, as do many of the modern GPTs. Their customers are widely dispersed through the economy, using their products, services, and technologies in a wide variety of distinct ways.

In our theory the forces leading to the division of inventive labor are quite simply those that make a GPT part of the cost-minimizing organization of inventive effort. These forces call for general specialists, and the inventions by these specialists are facilitated by economies of scale because, within the specialty, there is sharing or re-use of concepts and tools. But the other important point is that the broad applicability of a general specialty is not free. An inherent tension in any division of labor is that the distinct users of a technology, or for that matter of a good or service, employ it for different purposes. Consequently they have different needs, and these needs would be best satisfied by producing, adapting, or using the technology or input according to their special goals and demands. This is a force for localization. Standardizing the technology or input allows exploitation of the gains from specialization, while localizing it permits superior matching. A general specialty is a compromise between the scale economies inherent in specialization and the failure to localize inherent in generality.

In the next section we build a very simple model of this trade-off. The only open issue is whether to create a general specialty input or a set of localized specialty inputs. The model has increasing returns in the specialty, a cost of mismatching, and no other elements. In section 10.3 we show how two different definitions of "the extent of the market" lead to opposite results for vertical specialization. Increasing the size of all applications sectors leads to vertical integration, while increasing the variety in applications sectors leads to division of labor. This model is the simplest one in which the Smith-Stigler conjecture is correct and can be distinguished from the confusing modern alternative. The existence of a general purpose technology in this model is endogenous in equilibrium.

In section 10.4 we examine several contemporary and historical GPTs. This leads us to, in section 10.5, a second model with extra features that we think capture some of the essence of science- and engineering-based general specialties. In particular, the second model has an extremely simple treatment of how science's abstract and explicit representation of knowledge becomes useful. The second model suggests a higher economic return to formal science and engineering in a larger and more diverse economy because generality in the representation of knowledge itself lowers the costs of broad and general use. The main result of the model, reported in section 10.6, is that the same force, broad application, that leads to general specialties is also a force for abstraction, science, and formal engineering. Just as the division of labor is limited by the extent of the market, so too is the division of inventive labor. In the last section we examine some very specific cases of general specialty industries. Extensive markets in the "many-uses" sense are a far better predictor of the division of inventive labor than are extensive markets in the "large-uses" sense.

We argue that our first model explains the role of general specialists, at an abstract level, in economic growth generally. Like any scale economies model, it faces the very real challenge of explaining why the scale economies are not exhausted in large economies. Our second model shows how opportunities for increasing the formality, abstraction, and generality of scientific and engineering knowledge have preserved an economic role for GS scale economies. Exploitation of these opportunities has called for further specialization of knowledge creators, further vertical disintegration of inventive effort, and the creation of new GPTs. Two important forces are at work here. First, science and engineering have been amenable to specialization, refinement, and formalization. Smaller and smaller subspecialties are created, which are potentially but by no means necessarily available for commercial application. Second, modern economies are

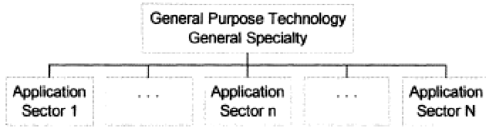

Figure 10.1
Industry structure with a GPT

diverse and complex as well as large. Consumerist societies have a wide variety of kinds of products and services, and international trade (supported by better transportation and communication) means that a number of distinct societies can be served by the same technology. These two forces imply that new general purpose technology industries can be founded and can evolve into new engines of growth.

10.2—
General Specialists and Cost-Minimizing Industry Structure: The Simplest Model

We consider an activity, like the creation of new knowledge or the provision of intermediate inputs, which has a tension between generality and localization. The knowledge or inputs are broadly useful and scale economies in their creation or provision are important, so there is a value to generality. Yet different uses need different varieties of the knowledge or features of the input, a force for localization. In this context, we consider two alternative systems of organization of economic activity, shown in figures 10.1 and 10.2.

The first has a general purpose technology based upon a general specialty. The industry structure is vertically disintegrated, with a GS-based GPT selling to many application sectors (AS) sectors each of which performs its own localization. In the other, there is complete localization of invention within each AS. No independent GPT sector exists. Existing theory has done a good job of understanding the incentive, welfare, and growth implications of a GPT but has not yet explained when a GPT will come into existence. Equilibrium in our model will determine which of the two systems, one with, the other without, a GPT, will organize activity.

Abstracting away from incentives issues, we ask which system minimizes costs. This leads immediately to the (true) Smith-Stigler theorem with the right definition of "the extent of the market" and also to the counterintuitive modern theorem with the other definition.

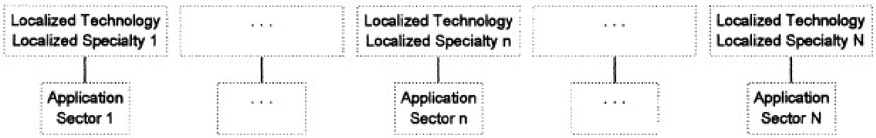

Figure 10.2
Industry structure with localization

10.2.1— Setup of Problem and Assumptions about Costs

Many of the elements of our very simple model are the same under either system. We posit N distinct uses of the knowledge or input—applications sectors. Each use has a desired rate of production, Q_n. Since ours is a theory of cost, we take N and Q_n, to be given and exogenous.

Each sector needs to undertake two activities, which might be understood as production processes or as inventions. One of the two activities is always localized and serves to define the identity of an AS. The other activity might be localized or might be served by a general specialty. These have cost functions given by c_n,(Q_n) and K (Q), respectively. We make no assumptions about the c_n( ), since they do not affect our results. The costs K( ) are the costs of providing the potentially general input. We assume that there are increasing returns in this activity—that is, the elasticity of K( ) is smaller than 1. Moreover we assume that the degree of scale economies is falling with the rate of output, namely $K(Q)/Q - K'(Q)$ declines with Q. These assumptions on the shape of K( ) are crucial.

10.2.2— Costs under Localization and General Specialty

If there is no general specialty, the costs in sector n are given by

$$c_n^L = c_n(Q_n) + K(Q_n), \qquad (1)$$

where the superscript L refers to localization. The first term is the costs of the activity that occurs only in this AS. In this localized industry structure the costs of providing the potentially general input are borne separately by each use.

If there is a general specialty (a GPT sector), costs in it are

$$C^{GS} = K\left(\sum_{n=1}^{N} Q_n\right). \qquad (2)$$

Note that the function K( ) of the GPT sector is the same as for each individual AS under localization. However, they will be lower than the corresponding costs in (1) because of the scale economies inherent in K( ).

Under a general specialty system, however, costs in the applications sectors are raised by "mismatches" and absence of variety. We represent these costs as

$$C_n^{GS} = c_n(Q_n) + d_n Q_n. \qquad (3)$$

Here d_n 0 is the "distance" of the applications sector from the general input or knowledge. The greater this distance, the greater the penalty for failure to localize the input or knowledge. Our model is one of overall cost minimization, so we ignore transfer prices, contracts, and so on.

10.3— Division of Labor and Extent of Market, I

We now examine the conditions under which our first model will have division of labor in cost-minimizing industry equilibrium. We examine the comparative statics of cost-minimizing industry structure as we increase the extent of the market. We use two definitions of the extent of the market: Q, the desired production in each use, and N the number of uses. For simplicity we offer proofs only for the symmetric case where all N sectors have the same $Q_n = Q$ and the same $d_n = d$.

The cost-minimizing structure is general specialist whenever

$$\sum_n (c_n(Q) + dQ) + K\left(\sum_n Q\right) < \sum_n (c_n(Q) + K(Q)).$$

In the vertical integration literature the theorem below contradicts the Smith-Stigler conjecture.

Theorem 10.1

Increases in the size of each use, Q, can change cost minimizing industry structure from general specialist to localized and vertically integrated but not the other way.

Proof

We differentiate the difference in costs between the two structures with respect to Q and evaluate the derivative at a point where the two regions are equal cost (if such a point exists). Define   to be the difference between industry costs under GS and localization. Then $NdQ + K(NQ) - NK(Q)$, and

$$\left(\frac{\partial \Delta}{\partial Q}\right) = Nd + NK'(NQ) - NK'(Q).$$

At   0, $Nd = N[K(Q)/Q - K(NQ)/NQ]$. Replace this in the expression for (/ Q). One obtains

$$\left(\frac{\partial \Delta}{\partial Q}\right) = N\left[\frac{K(Q)}{Q} - K'(Q)\right] - N\left[\frac{K(NQ)}{NQ} - K'(NQ)\right].$$

But the degree of economies of scale is falling with the rate of output. Then $K(x)/x - K'(x)$ declines with x, and therefore (/ Q) > 0. ■

This theorem goes the reverse way of the Smith-Stigler theorem because it has a local definition of the extent of the market. As each use grows large, the fixed costs of inventing technology for it grow small by contrast. But the costs of using inappropriate technology or infrastructure scale up along with the market. This is by far the most common result in the literature. It is almost always demonstrated for the $N = 1$ case because large Q is the only definition of "extent of the market" considered in the papers.[4] It is obvious that the result extends to such alternative comparative statics as increasing the Q, of a single sector—that sector can switch from but not to using the GPT. Similarly extensions of the theorem to the case of heterogeneous d_n and $K_n(\ )$ call only for a treatment of the possibility that some AS's will use the GPT, while others will not but will have the same kinds of results as here. We now show a result in which the Smith-Stigler conjecture is correct.

Theorem 10.2

An increase in the number of uses, N, can change cost-minimizing industry structure from localized and vertically integrated to general specialist but not the reverse.

Proof

By same method evaluate (/ N) at = 0. One obtains

$$\left(\frac{\partial \Delta}{\partial N}\right) = -Q\left[\frac{K(NQ)}{NQ} - K'(NQ)\right] < 0. \quad ■$$

This result is different because the margin of more uses (N) is economically quite different from the margin of larger uses (Q). Larger uses tend to call forth more localized and locally optimized inputs. More, or more diverse, uses call for the creation of general specialties, as the general specialist can be spread out over more units. This highly stylized model cannot reveal very much more than the distinction between the two margins, but it does show that distinction very clearly.

These two results clarify an important aspect of the Smith-Stigler conjecture that division of labor is limited by the extent of the market. If by the latter one means the number of firms of a given size, then division of labor is encouraged by an increase in the number of such firms. But, if increases in the size of the market occur through increases in the size of firms or sectors (i.e., in the potential demand of applications of a certain type), there is a discouraging effect on division of labor. The intuition is

[4] See Perry's (1989) discussion.

straightforward. Suppose that the market expands because of a larger potential demand for the particular good produced by each firm. These firms can then spread the fixed cost of producing a localized technology on a larger internal market, which encourages them to invest in the local technology. But, if the market expands because of increases in the number of potential uses of the GPT, then the value of a *general* solution rises.[5] Thus only under the latter definition of the extent of the market is the Stigler-Smith conjecture true. This is clearly what Stigler had in mind. The infrastructural industries he saw as general specialties provided a well-defined service or product to all comers in a market organization.

10.4—
Division of Labor and Resulting Economies in Several Industries

We now examine the real world of several historical and contemporary general specialty industries. This scrutiny shows two things. First, our assumptions about $K(\,)$ and d_n have different but related analogies in the real world in each industry, so that the logic of the creation of a general specialty emerges clearly. Second, once each general specialty was created, technologists and business people noted the rigidities associated with the exploitation of scale economies. This led to attempts to lower d, namely to make the general specialty more adaptable and widely useful. The importance of this kind of change, and the way it drew on formal science and engineering, are an important part of our second model.

For Stigler's example of the railroad, long-run increasing returns to scale have two distinct meanings. First, the physical capital of a railroad is efficiently shared across many classes of shipments. Second, the (very considerable) *invention* costs of improvements in steam power, steel rails, telegraph, and management structures in order to control large transport systems could be similarly spread out. The cost for any user of building dedicated transportation lines linking, say, a particular shipper's most habitual routes could be considerable. Relatedly, single shippers were unable to generate enough transportation demand to justify these setup costs.

The corresponding d's are not zero. Shippers of different kinds care differently about speed, reliability, smoothness, and cost. A railroad opti-

---

[5] Earlier work which has anticipated our point has noted the value of the generality—see, for example, Rosenberg (1963) on technological convergence—or seen the general market as a source of even more returns to scale than the largest specific market—see, for example, Romer (1996)—without noting that the emergence of a GPT itself is contingent on the economic return to broad rather than deep application.

mized to deliver fresh fruit or passengers differs from a coal or grain carrier. By relying on specialized suppliers, the users were giving up the opportunity to ship their freight at their most desired moments in time, along the optimal routes between departure and destination, or under any other very special condition. This happened for the obvious reason that the layout of the railroad system, and its scheduled routes and timing, had to be optimized according to the utilization of the network as a whole by its many customers rather than fitting the exact needs of individual users. Nonetheless, the combination of operational and invention scale economics swamped these modest benefits of diversity and a general specialty, railroading, emerged.

10.4.1— Making the General Specialty More Adaptable

In the eighteenth and nineteenth centuries, when railroads were a new and high-technology industry, a growing demand for railroad transportation by many customers with different shipping requirements (timing, destination, routes, etc.) gave a great impetus to improve and widespread coordination within and among railroad companies, thereby improving the efficiency of utilizing the network by any of its users. As Chandler (1990, pp. 53-56) points out, this was carried out by notable organizational and technological innovations. The U.S. railroad companies pioneered the techniques of modern management. They undertook considerable steps toward a "scientific" approach to the scheduling of movements of trains, freight, and passengers; and to the optimization of routes and connections between hundreds of locations and destinations; and to the maintenance of railroads and related equipment. This very exact scheduling, which was critical to enhancing the efficiency of transportation, was created by subdividing vast and very complex set of operations into a hierarchy of smaller and simpler tasks, which were supervised, monitored, and coordinated by different layers of managers. The effect of this advance was to progressively lower, but never to remove, the *d* costs associated with railroads as a transport system.

Fundamental change in the conditions of localization for transport awaited the invention of the automobile and the truck.[6] This technology shifted the boundary between the general and the localized. "Road" continued to be general (and indeed shifted to public provision), whereas "rolling stock" and "management" became specific to the using sectors.

[6] See Bresnahan and Gordon (1996) for a discussion.

Now a user would own its vehicles and, subject to a congestion externality, schedule its own shipments or travel. Even though motor vehicles are subject to vastly less scale economies than are trains, they have increased flexibility, breaking the rigidity of sameness and standardization. A less steeply sloped *K*( ) permitted considerable escape from generalist production and from the *d* costs imposed on users.

A different example comes from the twentieth-century chemicals industry. At the beginning of the century, each firm designed its own manufacturing process. There was very little sharing of process knowledge across makers of different products. The emergence of the chemical engineering discipline changed that radically. As noted by Rosenberg's (1997) chapter in this book, chemical engineering emphasizes the quantitative analysis of basic "unit operations" such as distillation, evaporation, drying, filtration, absorption, and extraction, and this provided a unified framework for thinking about the design of different processes in oil refining and in chemical production. While analysis of these operations varies with the material being operated upon, the analytical principles do not vary. The engineering discipline, with strong roots in science, was able to advance understanding of the general analytical principles and of their mode of application to different materials. Thus the invention of a general specialty involved a division between the general (here process knowledge) and the specific (the application of that process knowledge to particular products).

Most important, this development created enormous opportunities for specialization of the invention function itself. Especially with the high growth of chemical and oil refining markets immediately after World War II, the industry witnessed the formation of many companies with deep process knowledge, the specialized engineering firms (SEFs)[7] Their services assisted chemical and petrochemical companies in the design and engineering of chemical plants and oil refineries. The SEFs supplied process design and engineering services for a wide number of products (plastics, fibers, elastomers, etc.). Some of the most prominent SEFs have generated important *process* innovations, and significant improvements of existing processes.[8] Universal Oil Products (UOP), for instance, has been responsible for a tremendous number of inventions throughout this century. UOP acted as the R&D department of many small and independent oil refiners and chemical firms. Still today it holds a huge number of

[7] This story is discussed in detail in Arora and Gambardella (1997).

[8] See Arora and Gambardella (1997). See also Freeman (1982).

licenses in many oil refining and chemical processing technologies, and in several countries.

Two of UOP's technologies figure as quintessential examples of general purpose inventions—the first continuous cracking process for producing gasoline, the Dubbs process, developed in the 1910s, and the Udex process for separating aromatic chemical compounds from mixed hydrocarbons, developed in the 1950s. The value of the Dubbs process was twofold. It worked continuously, without stopping production. And it produced gasoline from either high-quality feedstocks or from low-quality "black oil." 'With the Dubbs process, UOP could live up to the 'universal' in its name by cleanly cracking *any* oil, regardless of coke-formation quantities."[9] Rather than vertically integrate forward into refining, UOP licensed its technology to the myriad local refiners, helping them specialize it to their particular feedstocks.[10]

During the 1950s UOP developed the Udex process to separate aromatic compounds (benzene, toluene, and xylene) from mixed hydrocarbon streams. These aromatic compounds are themselves general purpose inputs used in the making of many distinct chemical prospects. The Udex process was extremely flexible: "Generally, UOP has been able to assemble a combination of processing 'blocks' that would allow a producer to make any desired combination and relative quantity of benzene, toluene, and xylene isomers from every conceivable feedstock."[11]

These examples show that the SEFs' expertise was specialized in the sense that it was deep in particular processes but general in the sense that it cut across many products. Further the SEFs were able to generalize their process knowledge to a wide variety of distinct feedstocks.

Parallel to the SEFs, university chemical engineering departments played a crucial role in making fundamental process improvements. Many fundamental process improvements were made in universities or SEFs and shared across a wide variety of firms making a wide variety of products. Here the source of increasing returns in $K(\ )$ lies entirely in the *invention* of improved processes. Once the chemistry was understood, it was wasteful to separately advance knowledge of how to improve the same process for distinct chemical products. Further the general inventions lowered the costs of particularization, namely lowered d. The general

[9] Remsberg and Higdon (1994, pp. 50–51). Italics in the original.

[10] Note the incentives that arose because of a large market in the "many-uses" sense: "UOP's approach to process licensing was particularly applicable to the oil refinery business of that era, Practically every little town in the country with access to oil had a small refinery. . . ." (Remsberg and Higdon 1994, p. 50).

[11] Spitz (1988, p. 191).

specialists in universities and in the SEFs made inventions of general scope and value, with the knowledge represented in the manner of general scientific principles. Applications to local circumstances were just applications, analogous to students' problem sets rather than to the development of a whole new body of knowledge.

One feature of the chemicals history is that things that had been done distinctly came to be done in the same general way. Rosenberg (1963) identifies a process of technological convergence that creates general specialties. While the distinct applications sectors had been separately served by distinct technologies (e.g., industry specific machine tools), after convergence they are served by a common, general input. Our theory ignores this dynamic, only suggesting that large and diverse markets increase the return to this kind of generalization.

A related example is the invention of the microprocessor by Intel in 1971. Before then, integrated circuits were largely "dedicated" products, in the sense that their operations were defined by the physical wiring and interconnections designed and built by the manufacturer on the chip. An important consequence was that the circuits had to be produced by the manufacturers with specific applications in mind. By contrast, the microprocessor, or "programmable chip" as it was aptly named, could be programmed. As Braun and Macdonald (1982) note, it implied "software wiring" as opposed to "hardware wiring." It could then read and process more variable instructions, and perform a far larger number of operations. Most important, it could be produced without specific applications in mind. Its functions could be defined to a greater extent by the users themselves who could program the chip according to their needs.[12]

It is then not surprising that the device rapidly found extensive applications. Apart from its core use in microcomputers, it became a pivotal component in telecommunications, aerospace, and office equipment, in the control of industrial processes, in the automobile industry, among many others. Its utilization extended the range of applications of integrated circuits. In a sense this widespread application was the natural consequence of the fact that whether deliberately or not, the microprocessor was conceived, from its very invention, as a general purpose object. The impact

[12] Braun and Macdonald (1982) suggest that the distinction between the microprocessor and the earlier integrated circuits is really that between a computer and a calculator. The latter can perform only the functions that are "permanently" defined on its chip and that can be activated by pressing special keys (e.g., numbers or arithmetic operations). The computer instead can read instructions defined in many possible ways (logic, arithmetic, etc.) and therefore perform more elaborate and distinct operations. They also note some intermediate forms, like the erasable programmable read-only memory.

was to lower *d*. Ultimately the microprocessor also changed part of the integrated circuit business into one characterized by very steep *K*( ). A general microprocessor is a very complex device, and we now see a small number of firms making very long production runs of a few microprocessor designs.

The uses of integrated circuits vary along another dimension as well, the performance—cost trade-off. Hardware wiring, despite all the wasteful duplications of design costs (multiple *K*( )) it involves, offers superior performance in many applications. An "applications-specific integrated circuit," or ASIC industry, flourished in parallel to the microprocessor GS. Here the division of inventive labor is quite different. Manufacture of ASICs is performed by general specialists. But unlike Intel or Motorola, these are specialists in the manufacturing process only. They do not design the products they make. Applications sectors design ASICs and solicit manufacturing cost bids from these general specialists. A fundamental organizational innovation has arisen to lower the *d* costs in this industry. A language has emerged for describing ASIC designs. It is a computer language, spoken by two very different kinds of computers. The first are computer-aided design workstations used in the AS. The second are manufacturing-control computers used in the GS. By this mechanism even ASICs can have "software wiring."[13] The AS firm designs a logical chip, and the GS firm makes it. Thus both the substantial scale economies in the plant (steep *K*( )) and the substantial benefits of localization have been achieved by this *d*-lowering organizational invention.

In sum, the more important modern general specialties have been in the science- and engineering-based industries. These offer a distinct advantage, one in which the tension between localization and generalization is lessened by invention of lower *d* ways to organize inventive activity. The advantage arises in the use of regularized, systematized knowledge to make AS invention easier. This works to draw more AS's into the ambit of the GPT. Our model needs an extension to deal with these greater opportunities, but an extension that is consistent with the basic thrust of the model.

10.5—
Model with Endogenous Generality and AS Voluntarism

We add two features to our model in order to capture the tension between localization and scale economies more fully. First, we model the

---

[13] That is, they can have software wiring at design time if not at use time. ASICs are more suitable for use in a wide variety of special-purpose devices than in, say, computers.

general specialty as being able, for a fixed cost, to lower *d*. This captures the costs and benefits of a more abstract, portable, and reconfigurable technology, as described in the examples in the last section. We will interpret lower *d* equilibria in our model as the endogenous creation of a more general purpose technology. Second, we permit the AS to be of different sizes and to substitute in and out the use of the GPT. Then there can be variety in the technology choice of the AS's. Some users (the larger ones, from the analysis of the earlier section) develop a localized version of the generally useful technology. Others, and the boundary here will be endogenous to equilibrium, can choose to participate in the general purpose technology, using the goods or services of the general specialist. This involves two changes: We differentiate between the size of the market in the AS and the size of the market the AS offers to the GPT, and we permit sectors to vary in their participation.

10.5.1—
Setup of Model:
Localized or General Purpose Technology

The AS sectors produce a final good using a potentially general intermediate input whose quantity is still denoted by *Q*. We now add the size of the market demand for the specific good produced by each sector, denoted by *S*, and we assume that S is distributed across firms according to a density function *f*(*S*). Thus our sectors are exogenously of different sizes, though the size of their demand for the general input is endogenous.[14]

The gross surplus in the AS is (*S*, *Q*). Apart from the obvious assumptions that $\pi_s > 0$ and $_Q > 0$, we assume that the cross-partial of with respect to *S* and *Q* is positive, $\pi_{SQ} > 0$. A larger sector obtains larger gross benefits from *Q*. There are underlying price and quantity decisions about the final good in the background, but they can most safely remain there.[15] We focus attention on the demand for *Q* and the technology choice of the AS as a function of *S* and of the GS' offerings.

We are interested in two experiments. The first experiment consists of increasing the *number* of applications sectors, which is accomplished by replication, increasing the density *f*(*S*) proportionately at all *S*. The second

[14] We will refer to *S* as being the size of the market of each sector or more simply the size of the sector. To save notation, we will not use the subscript *n* for firms when referring to *S* and *Q*.

[15] Equilibrium in the AS maximizes $\pi(S, Q)$. Because of our theory's strong cooperative flavor, it is natural to interpret it $\pi(S, Q)$ either as the profits of the AS, if the AS is supplied by a monopoly firm, or as the sum of consumer plus producer surplus. See Bresnahan and Trajtenberg (1995) for more interpretation.

experiment consists of increasing the *size* of all applications sectors. This is accomplished by scaling the size of the sectors to be $S^* = \alpha S$, instead of just S. We assume that S is still distributed according to the density function $f(S)$, and that α is a proportionality factor.

Localized or GPT Choice for a Single AS

Once again, the input Q can be of two types. On the one hand, the AS can choose to utilize a "local" technology. That is, they develop a specialized technology at their own location, and this can be designed to suit in a fairly exact way their special needs. If the sector uses a local technology, they incur costs $K(Q)$ to generate it. The net surplus is

$$\Pi^L \equiv \pi(S, Q) - K(Q),$$

where the superscript L denotes the "localized" technology regime. We also define $Q^L(S)$ to be the optimized Q under this regime and $\quad ^L(S)$ to be the corresponding optimal profits. Our assumption on the cross-partial of $\quad$ implies that both $Q^L(S)$ and $\quad ^L(S)$ increase with S.

On the other hand, the sector can choose to buy the input Q from the GPT sector (if one exists). In this case they do not incur $K(\,)$, but they incur the d costs and a cost w per unit of Q, which is the price of the GPT. The total unit price of using the GPT will be $(w + d)$. The payoff to the AS in the "generalized" technology regime (superscript G) will then be

$$\Pi^G \equiv \pi(S, Q) - (w + d)Q,$$

where as above we define $Q^G(S, w + d)$ and $\quad ^G(S, w + d)$ to be the optimal demand and payoffs of the AS sector respectively. Both $Q^G(\,)$ and $\quad ^G(\,)$ increase with S and decrease with $w + d$.

"Marginal" S

At this point, we can define the marginal S, denoted by S^0, to be the size S such that the sector is indifferent between using the localized technology or the GPT. The marginal S is defined implicitly by the following expression

$$\Pi^L(S^0) = \Pi^G(S^0, w + d). \qquad (4)$$

We know from the analysis of the previous section that the larger S sectors will prefer localized over generalized choice. Thus the AS using the GPT will be all those with $S \quad S^0$, while all AS sectors with $S > S^0$ will use the localized technology. It is straightforward to see that the marginal S declines with $w + d$. Hence the set of firms that use the GPT increases as d declines or if scale economies in a GPT sector lead to lower w in equilibrium.

10.5.2— General Specialist GPT Sector

The general specialist sector sells the intermediate input Q to the applications sector, once again bearing costs that depend on the total demand. We now add an opportunity for the GS sector to invest in reducing the costs, d, of failures to localize. As the examples in our previous section suggest, these can really be thought of as investments in generalized knowledge, inventions, and technologies. By reducing d, the invention lowers the "economic" distance between all applications sectors and the GPT. In turn this reduces the penalty from using the GPT instead of a more specialized technology.

The GPT sector incurs a fixed cost for reducing d, $C(d, \theta)$, where θ is a parameter indexing the exogenous state of science or engineering. We assume that $C_d < 0$ and $C_{dd} > 0$. The cost of reducing d increases at an increasing rate, and it is natural to think that further generalizations of technologies or knowledge bases are harder to obtain.

The flavor of our analysis continues to be cooperative and cost minimizing, even though we now treat GS and AS as choosing separately. On the revenue side, the GS sector benefits from selling the GPT input Q, at a price w, to all industries or firms that are willing to buy it instead of resorting to their local technology. We assume, consistent with the cost-minimizing structure, that w is set to cover the (average) costs of the GS sector.

10.5.3— Equilibrium: The Game between GPT and AS sectors

To analyze the cost-minimizing solution to this system, we assume that the AS and GPT sector have payoffs consistent with cost minimization and play a static Nash equilibrium game. AS decide whether to use GPT or localized technology depending on their price-taking profit. The GPT sector picks d to maximize its own surplus but takes w as given. Since each AS sector chooses individually, we will focus our results on the size of the subset of AS sectors that uses the GPT good rather than localized technology.

Total demand for the GPT good is the demand by all sectors with S S^0. Once again we measure the number of sectors by N. That is, the total number of sectors of type S is equal to $N \cdot f(S)$, where $f(\cdot)$ is the density function introduced above. Then the sales of the GPT sector are

$$Q^{GS}(w+d) = \int_{S=0}^{S^0} NQ^{G}(S, w+d) f(S)\, dS. \tag{5}$$

The payoffs are

$$\Pi^{GS} \equiv wQ^{GS} - K(Q^{GS}) - C(d, \theta). \tag{6}$$

Now w is taken as given by the GS, but at a level greater than MC (by our average cost pricing rule). This is an incentive to lower d. S^0 is a function of d, and S^0 increases as d declines. Hence the generalist sector's willingness to reduce d arises from the desire to attract new applications' demand to GPT use.

Given the expressions for ^{G} and ^{GS}, we look for the Nash equilibria of the game among the GS sector and the ASs. The strategies of the latter are the level demands for the input Q, whereas the strategy of the former is the level of d.[16] The first-order conditions of the problem are

$$\Pi^{G}_{Q} \equiv \pi_Q - (w+d) = 0, \tag{7}$$

$$\Pi^{GS}_{d} \equiv N \cdot f(S^0) \cdot Q^{G}(S^0, w+d) S^0_d \cdot (w - K'(Q^{GS})) - C_d = 0. \tag{8}$$

The first-order condition (7) is the one for all the ASs. The first-order condition (8) is the one for the GPT sector. Note that the latter is evaluated at the marginal S, S^0. Thus the behavior of the generalist-specialist sector is determined by the conditions that characterize the "marginal" application. This last feature comes because our game structure has the GPT taking w and the Q of all the inframarginal ASs as given. We also want to note at this point that our structure is not meant to represent a realistic market game. In fact some readers may be confused by this lack of "realism." But our problem is simply to find the collectively coordinated, cost-minimizing equilibrium, and the concept of Nash equilibrium is a useful one for this purpose. Thus we do not think that this realism is crucial for our argument. At the same time, although we do not put any structure on how the cost-minimizing equilibrium is reached, we do want to note that a cost minimizing industry structure is a strong "attractor."[17]

---

[16] To avoid confusion, it is important to recall that the benefits associated with the GPT increase as d decreases, and vice versa.

[17] Moreover, if one is very keen about realism in the game, in the lemma below we use the concept of supermodular games developed by Milgrom and Roberts (1990). Milgrom and Roberts (1990) also show that the results of supermodular games can be obtained as the outcome of an "adaptive dynamic" process wherein the strategies of the players are played

(footnote continued on next page)

Lemma

The game between the ASs and the generalist sector is supermodular in (Q, d). Hence there exists a set of Nash equilibria, and these are ranked. The best Nash equilibrium is the one with the largest Q's and the smallest d.

Proof

Our game satisfies the conditions for supermodular games in Milgrom and Robert's (1990) theorem 4. First, our strategies are bounded.[18] Second, the cross-partial derivatives of the objectives functions, namely ■

10.6— Division of Inventive Labor as Limited by the Extent of the Market, II

We now once again see that changes in the total size of the market to which the GPT can be applied ($N \times S$) have opposite implications on the size of the GPT sector according to whether they occur through increases in N or S. The new result here is that the division of inventive labor also shifts distinctly as N and S change. Theorem 10.3 below shows that if N increases, with S constant, the GPT sector expands, in the sense that the optimal d declines and the number of firms using the GPT increases. Theorem 10.4 shows that a proportional increase in the size of all firms (or sectors), S, with N constant, leads to a decline of the GPT sector, with opposite effects on d and on the number of firms using the GPT.

Theorem 10.3

An increase in the number of AS in the economy (N) implies a larger GPT sector, in the sense that (i) d is smaller and (ii) a larger set of firms buys the GPT.

(footnote continued from previous page)

over time, and there is "learning" over time. They show that within an adaptive dynamic framework, in any finite strategy supermodular game there exists a date after which the strategies are bounded from above and below by the largest and smallest Nash equilibrium. (See Milgrom and Roberts 1990, thm. 8 and related corollaries.)

[18] Both d and Q are bounded if the demand curves for Q as a function of $w + d$ cut both axes.

Proof

From conditions (7) and (8) it is easy to see that ■

Thus theorem 10.3 says that increases in the size of the market are a powerful determinant of the rise and expansion of a generalist specialist sector, and correspondingly of generalized knowledge bases and technologies. A more extensive market means a larger number of distinct uses of a general purpose technology. With more distinct uses there is, as we noted in the simpler model, a cost-minimizing incentive for the existence of a GS. We see moreover that the generalist specialist industries have greater incentives to reduce *d* when their markets are more extensive. The technology process of a diverse market economy involves the creation of GPT industries and systematic attempts to make them more general despite their specialization.

But as theorem 10.4 below suggests that even in the broader framework it is still the case that larger individual uses lead to less not more vertical division of labor. The possibility of making the GPT more general introduced in this section is not sufficient to reverse that result.

Theorem 10.4

A proportional increase in the size of all firms in the economy (S) implies a smaller GPT sector, in the sense that (i) d is higher and (ii) a smaller set of firms buys the GPT.

Proof

Transform S into $S^* = \alpha S$, with > 0. The distribution of S^* evaluated at the marginal S^* is $G(S^0) = F(S^0/\alpha)$, where $F(\ )$ is the distribution of S. It is straightforward to see that as α increases the percentage of firms that buy the GPT decreases.

To show that d increases, the density of S^* evaluated at S^0 is $g(S^0) = (1/\ )f(S^0/\ )$. Replace this expression for $f(S^0)$ in condition (8). The marginal benefit of lowering d decreases with as long as $(1/\ )f(S^0/\ )$ decreases with . This expression decreases with unless the percentage of firms at the margin, namely $f(S^0/\ )$, is declining so rapidly that a reduction in $S^0/$ implies a significant increase in the number of firms of marginal size.[19] ■

Theorem 10.4 says that with proportional expansions in the size of all using sectors, there are fewer incentives to utilize generalized technologies.

[19] Technically the elasticity of $f(\ )$ evaluated at $S^0/$ must be greater than –1.

As in our simpler model of section 10.2, one can have an analogous result for increasing the size of any particular using sector.

Factors Affecting Expansion of GPT: Advances in Science, θ

Finally we examine the impact of changing the conditions of knowledge representation. We introduced a parameter θ into the cost function for lowering d. We now examine the likely form that θ will take if it is the level of scientific knowledge, and its likely impact on a commercial general specialty.

If anything, advances in science mean greater ability to comprehend a wider set of previously unrelated phenomena within common explanatory frameworks, and this facilitates efforts to reduce the distance among them. As a matter of fact, scientific advances have often created technological linkages among formerly distinct industries. For example, greater understanding of solid state physics during the 1950s led to the development of the transistor, thereby inducing greater commonality among the technological bases of industries such as telecommunications, office equipment, consumer products. Similarly advances in the theory of organic chemistry enabled the German chemical industry during the nineteenth century to link molecular structures to the properties of many different substances. Organic chemistry then became the common scientific and technological basis of sectors such as dyestuffs, pharmaceuticals, and explosives. In a very similar way, after World War II, theoretical advances in polymer chemistry provided the common framework to "design" the molecular structures of new plastics, fibers or rubberlike products.

A related interpretation is that θ measures the size and the quality of the professional bodies that in any given industry or society are dedicated to the production of ideas, and that are specialized in the creation of generalized technologies and knowledge bases. It has been suggested, for instance, that the United States provides a more "scientific" education for software programmers than Japan and that this accounts for some of the difficulties the Japanese software industry faces in producing more basic software templates instead of specific, and often highly customized, applications (Cusumano 1991; Nakahara 1993). Similarly Rosenberg (1997) argues that the systematic training provided by the U.S. universities, MIT in particular, in chemical engineering since the end of World War II has been critical for the diffusion of highly skilled professionals in the field. The training of U.S. chemical engineers involved a solid grasp of the scientific foundations of chemical process design. In turn this created a body of professional expertise that could be employed to design and engineer many different types of chemical plants or refineries.

We can then assume that in the problem of the generalist specialist sector the cost of reducing the distance among the application industries is $C(d, \theta)$, where along with the previous assumptions about C_d and C_{dd}, $C_\theta < 0$, and $C_{d\theta} > 0$. The latter two assumptions characterize the role played by the parameter θ. This parameter reduces the cost of a given d, and it reduces the marginal cost of additional reductions in d. It controls for the "ease" with which the generalist specialist sector can lessen the economic distance of the GPT from its applications. Using the framework of our model, we can state the following theorem:

Theorem 10.5

Advances in science (θ), or any other factor that lowers the marginal cost of reducing d, implies a larger GPT sector, in the sense that (i) d is smaller and (ii) a larger set of firms buys the GPT.

Proof

Follows that of theorem 10.3. From the first-order conditions, ■

Theorem 10.5 clarifies the importance of factors that increase the ability of the generalist sector to utilize basic knowledge to produce general technologies. Our argument corroborates the conjecture made by Arora and Gambardella (1994). There it was suggested that science is a powerful instrument to codify knowledge in ways that enable industries to link seemingly "distant" products and technologies. This led to a "division of innovative labor" because the fixed cost of producing a given piece of knowledge could be spread over a larger market.

One can also inquire about the comparative statics in d. What if all applications sectors of an economy grow more diverse, in the sense that all of their *d's* grow? We have examined this question, and it turns out that the answer is ambiguous. The reason for the ambiguity is that there are important offsetting effects. A more diverse (higher d) economy has a higher return to localization in a model like our first one. But it also has a higher return to creation of a science- or engineering-based technology (d-lowering). Either of these can dominate in equilibrium, depending on the size distribution of sectors, the shape of $K(\ )$, and so on.

Finally another natural extension is to the case in which apart from the GPT sector, each individual AS can make investments that help reduce their own d, and these investments are complementary to those made by the GPT sector. From the point of view of our theory, this is a trivial case.

The complementarity assumption about the GPT and AS investments implies that any factor (N, S, or θ) that shifts d in a certain direction would shift the individual d efforts of the AS in the same direction.[20] We wanted to mention this case, however, because in many high-tech industries today users often make such complementary investments. For instance, in the development of new information systems many software companies develop general "templates," and the customization of the template occurs through a "rapid prototyping" process in which the general system is passed on to the buyers who start using it and suggest ways to make it closer to their actual needs.[21] Here d is lowered because of efforts made by both parties. Our analysis suggests that users will be more willing to make complementary investments to customize "generic" products to their own needs in markets where there are many uses (N), high scientific skills (θ), or uses of proportionally smaller scale (S).

A Modicum of Evidence: N ≠ S in Two Interesting GPTs

Our theorems about the relationships between GPTs and the size of markets states that the division of "inventive" labor is associated with markets that feature a greater number of distinct uses of a basic technology. These larger markets expand the boundaries of the GPT by widening its breadth of applications and by encouraging investments that reduce the cost of using the GPT vis-à-vis more specific solutions. In this section we look at two industries that each permit examination of this point from two different perspectives.

First is the software industry. From the founding of the computer industry through the mid 1970s, there was no other notion of a commercial computer than large mainframes. The users were large organizations that could afford the high capital outlay and could justify these costs because of extensive utilization of the machines—such as the U.S. military and large corporations. Software was customized. Consultants, systems integrators, or the employees of a particular user would handcraft software applications for it. The rise of a "packaged" software industry, selling "standard" tools for many users, took place only in the 1970s and especially in the 1980s, and was associated to a good extent with the development of minicomputers and later on of the PC.[22] PCs and mini-

---

[20] This is also one of the main results in Bresnahan and Trajtenberg (1995).

[21] See, for example, Hofman and Rockart (1994).

[22] See OECD (1985) on the development of the software industry. Also see Mowery (1996) for a very interesting international perspective.

Table 10.1
Japan-U.S. Hardware and software comparison (1987)

	Japan	United States
Hardware shipments	21.0	45.6
Large systems	8.7	9.1
Medium systems	3.1	8.7
Small systems	5.0	8.2
Personal computers	4.2	19.6
Software-vendor revenues	13.0	24.8
Total packages[a]	1.4	13.1
Custom software and system integration	10.1	9.6
Facilities management and maintenance	1.4	2.1
Total market	34.1	70.4

Source: Cusumano (1991, p. 49)
Note: In 1987 billion U.S. dollars; $1 = ¥125.
a. These include systems/utilities, application tools, and application packages.

computers, with the implied reduction in the price of computational power, meant a bewildering diffusion of computers among many users employing them for very different purposes. Notably many of these users were small (in the sense of small S), and hence they could not afford the upfront cost that was required to purchase or develop customized software. But there were millions of them, definitely large N.

The time-series flavor of this example is incompletely convincing. There was a good deal of technical advance in computers and in software which may well have shifted out the supply curve of independent software vendors. But consider the following cross-sectional evidence: For many reasons the diffusion of minicomputers and PCs in Japan has been slower than in the United States (Cusumano 1991 or Nakahara 1993). As table 10.1 shows, the comparative value of large hardware systems in Japan visà-vis the United States is much bigger than in the case of smaller systems and PCs. The numbers are striking. While the sales of large systems in Japan are approximately as big as in the United States (8.7 vs. 9.1 billion dollars), the U.S. PC market is worth 19.6 billion dollars as opposed to 4.2 billion dollars in Japan. There the computer still appears to be the province of large users, who can afford and manage the large systems, and its diffusion among the vast population of smaller users is slow. At the same time table 10.1 shows that the U.S. market of packaged software dwarfs the corresponding Japanese market: 13.1 versus 1.4 billion dollars. By contrast, the figures about custom software are, again, of comparable size (9.6 vs. 10.1 billion dollars).

One can be skeptical of this evidence given the names of the countries in it. After all, market organization is generally more important in the United States, relationship organization for commerce generally more important in Japan. Another example shows that the inference is not country-specific. The history of the computer numerical control (CNC) machines during the 1980s is similar to that of software, but the positions of the United States and Japan are reversed. CNCs are machine tools whose automated tasks are controlled by a computer. The latter can be easily re-programmed to enable the machine to perform a variety of tasks. CNCs, which first emerged in the early 1980s, advanced the earlier technology of numerical controls (NCs) in which the automatic movements were controlled by computer punch tapes.

The United States pioneered this industry, and CNCs were actually invented in the United States. But while the United States had been the world leader in machine tools since the 1970s, Japanese producers expanded dramatically in the world market during the 1980s and increased considerably their exports into the United States. Particularly, the Japanese were able to enter with smaller, microprocessor-based job-shop CNCs, whereas the U.S. producers had remained with large mainframe-and minicomputer-based CNC machines.

The early NC machines were developed by the U.S. Air Force in the 1950s. Since then their diffusion in the United States occurred largely in two sectors, aerospace and automobile, and within the latter predominately among the Big Three—GM, Ford, and Chrysler. In the 1970s about one-third of the total U.S. market of NCs was in the aerospace sector, and the share of the automobile market was a little smaller (Rand 1994, vol. 1, p. 37). With the introduction of CNCs, the U.S. machine tool producers kept focusing on "large, sophisticated users in the automobile and aerospace industries with the available resources and complex requirements to enable adoption of large, expensive, difficult-to-use mainframe-and minicomputer-based CNC machines" (Rand 1994, vol. 2, p. 116); see also March 1988). This also meant that these machines were largely designed for the special purposes and requirements of these users and that the competencies of many machine tool makers were to a good extent sector-specific (Rand 1994; March 1988).

By contrast, Japanese CNC makers immediately focused on smaller, microprocessor-based CNCs for many more types of users. Fujitsu Automatic Numerical Control (FANUC) rapidly became the world leader, and the firm that set the world standards. What FANUC and other Japanese producers did was to develop machines with fairly standardized, commodity-type characteristics. This enabled them to reach the market of many smaller firms in quite distinct industries that were unable to develop large and expensive customized systems. As a matter of fact, even at the end of the 1980s the number of adoptions of NC or CNC machines by Japanese small-medium-sized firms was about 40 percent higher than of small-medium U.S. firms (Rand 1994, vol. 2, p. 112).

Moreover, because their market was composed of many different buyers with distinct features and needs, Japanese producers made significant investments in "modularizing" their production and design operations. This was a real revolution in the organization of their work and in the design of their products that enabled them to take advantage of economies of scale while still maintaining the ability to customize products to meet customer demands. For instance, they made considerable efforts to identify parts of machines for different uses that could be standardized at little loss in terms of specificity of the application (Rand 1994, vol. 2, p. 13). By mixing and matching standardized components, they could then package machines that were suited for different uses—the classic d-lowering strategy for a GPT.

10.8—
Conclusion

We have examined the endogenous emergence of GPTs as the creation of a general specialty where otherwise distinct and localized inputs or bodies of knowledge might have existed. This saves the Smith-Stigler conjecture from being stimulating but false. From a firm, market, and institutional microstructure perspective, the distinction between an extensive market (many customers) and a large individual customer seems quite useful in explaining some important technologies, both historically and in the present. From a long-run growth perspective, the results help explain why the scale economies inherent in GPT are growing more important even as economies get larger. Increased specialization of knowledge has been offset by increased generality in the representation of knowledge. With (nontrivial) commercialization efforts, and with increasingly diverse markets, new GPTs and new general specialties continue to contribute to growth, and growth continues to permit their invention. This counterintuitive but extremely important positive feedback loop would have been seen by Smith as one of God's better accomplishments and by Stigler as one of humankind's.

References

Arora, A., and A. Gambardella. 1994. The changing technology of technical change: General and abstract knowledge and the division of innovative labour. *Research Policy* 23: 523–32.

Arora, A., and A. Gambardella. 1997. Evolution of industry structure in the chemical Industry. In A. Arora, R. Landau, and N. Rosenberg, eds., *Dynamics of Long Run Growth in the Chemical Industry*. New York: Wiley.

Braun, E., and S. Macdonald, eds. 1982. *Revolution in Miniature*: *The History and Impact of Semiconductor Electronics*, 2d ed. Cambridge: Cambridge University Press.

Bresnahan, T., and R. Gordon, 1997. The economics of new goods: An introduction. In T. Bresnahan and R. Gordon, eds., *The Economics of New Goods*.

Bresnahan, T. and M. Trajtenberg. 1995. General purpose technologies: "Engines of growth"? *Journal of Econometrics* 65: 83–108.

Chandler, A., with T. Hikino, eds. 1990. *Scale and Scope*: *The Dynamics of Industrial Capitalism*. Cambridge, MA.: Belknap Press.

Cusumano, M., 1991. *Japan's Software Factories*: *A Challenge to U.S. Management*. New York: Oxford University Press.

Freeman, C. 1968. Chemical process plant: Innovation and the world market. *National Institute Economic Review* 45: 29–51.

Hofman, J. D., and Rockart, J. F. 1994. Application templates: Faster, better, and cheaper systems. *Sloan Management Review* (Fall) 36: 49–60.

Helpman, E., and M. Trajtenberg. 1994. A time to sow and a time to reap: Growth based on general purpose technologies. Tel Aviv Foerder Institute for Economic Research Working Paper 23.

Langlois, R., and P. Robertson, eds. 1995. *Firms*, *Markets and Economic Change*: *A Dynamic Theory of Business Institutions*. London: Routledge.

March, A. 1989. The US machine tool industry and its foreign competitors. In M. Dertouzos et al., eds., *Made in America*: *Regaining the Productive Edge*. MIT Commission on Industrial Productivity, Cambridge: MIT Press.

Milgrom, P., and J. Roberts. 1990. Rationalizability, learning, and equilibrium in games with strategic complementarities. *Econometrica* 58: 1255–77.

Mowery, D., ed. 1996. *The International Computer Software Industry*. Oxford: Oxford University Press.

Nakahara, T. 1993. The industrial organization and information structure of the software industry: A U.S.-Japan comparison. Unpublished manuscript. Center for Economic Policy Research, Stanford University.

Organization for Economic Cooperation and Development. 1985. *Software*: *An Emerging Industry*. Paris: OECD.

Perry, M. K. 1989. Vertical integration: Determinants and effects. In R. Schmalensee and R. D. Willig, eds., *Handbook of Industrial Organization*. Amsterdam: North Holland.

Finegold, D., ed 1994. *The Decline of the US Machine Tool Industry and Prospects for its Sustainable Recovery*. Santa Monica: Rand Corporation.

Remsberg, C., and H. Higdon. 1994. Ideas for rents: The UOP story, Universal Oil Corp. Des Plaines, IL.

Romer, P. 1986. Increasing returns and long-run growth. *Journal of Political Economy* 94: 1002–37.

Romer, P. 1990. Endogenous technological change. *Journal of Political Economy* 98: S71–102.

Romer, P. 1992. Growth based on increasing returns due to specialization. In K. D. Hoover, ed., *The New Classical Macroeconomics*. Aldershot, UK: Elgar.

Romer, P. Why, indeed, in America? Theory, history, and the orgins of modern economic growth. *American Economic Review* 86: 202–206.

Rosenberg, N. 1963. Technological change in the machine tool industry: 1840–1910. *Journal of Economic History*.

Smith, A. 1776. The Wealth of Nations. London.

Spitz, P. H. 1988. *Petrochemicals*: *The Rise of an Industry*. New York: Wiley.

Stigler, G. 1951. The division of labor is limited by the extent of the market. *Journal of Political Economy* 59: 185–93.

Von Hippel, E. 1995. "Sticky information" and the locus of problem solving: Implications for innovation. *Management Science* 40: 429–39.

Von Ungem-Sternberg, T. 1988. Monopolistic competition and general purpose products. *Review of Economic Studies* 55: 231–46.

Weitzman, M. 1994. Monopolistic competition with endogenous specialization. *Review of Economic Studies* 61: 45–56.