Sources, Procedures, and Microeconomic Effects of Innovation

By Giovanni Dosi
University of Sussex and University of Rome

Fabio Arcangeli, Paul David, Frank Engelman, Christopher Freeman, Massimo Moggi, Richard Nelson, Luigi Orsenigo, Nathan Rosenberg, Michele Salvati, G. N. von Tunzelman, two anonymous referees, and the participants at the meeting of the Committee on Distribution, Growth, and Technical Progress of the Italian National Research Council (CNR), Rome, November 16, 1985, have helped with various redraftings. A particularly grateful acknowledgment is for the insightful and patient help of Moses Abramovitz.

This work has been undertaken at the Science Policy Research Unit (SPRU), University of Sussex, as part of the research program of the Designated Research Centre, sponsored by the Economic and Social Research Council (ESRC). Earlier support to the research that led to this paper by the Italian National Research Council (CNR) is also gratefully acknowledged. The statistical research has been undertaken with the assistance of Stephano Brioschi, Ilaria Fornari, and Giovanni Prennushi.

I. Introduction

This essay concerns the determinants and effects of innovative activities in contemporary market economies. In the most general terms, private profit-seeking agents will plausibly allocate resources to the exploration and development of new products and new techniques of production if they know, or believe in, the existence of some sort of yet unexploited scientific and technical opportunities; if they expect that there will be a market for their new products and processes; and, finally, if they expect some economic benefit, net of the incurred costs, deriving from the innovations. In turn, the success of some agents in introducing or imitating new products and production processes changes their production costs, their market competitiveness and, ultimately, is part of the evolution of the industries affected by the innovations.

It is the purpose of this essay to analyze the processes leading from notional technological opportunities to actual innovative efforts and, finally, to changes in the
structures and performance of industries.

Thus, I shall discuss the sources of innovation opportunities, the role of markets in allocating resources to the exploration of these opportunities and in determining the rates and directions of technological advances, the characteristics of the processes of innovative search, and the nature of the incentives driving private agents to commit themselves to innovation.

It is not my purpose to review the whole body of innovation-related literature. Rather I limit my discussion to a selected group of (mostly empirical) contributions and focus on the microeconomic nature of innovative activities and the effects of innovation upon techniques of production, product characteristics, and patterns of change of industrial structures. The discussion will aim to identify (a) the main characteristics of the innovative process, (b) the factors that are conducive to or hinder the development of new processes of production and new products, and (c) the processes that determine the selection of particular innovations and their effects on industrial structures.

There are two major sets of issues here: first, the characterization, in general, of the innovative process, and, second, the interpretation of the factors that account for observed differences in the modes of innovative search and in the rates of innovation between different sectors and firms and over time.

Typically, the search, development, and adoption of new processes and products in noncentrally planned economies are the outcome of the interaction between (a) capabilities and stimuli generated within each firm and within industries and (b) broader causes external to the individual industries, such as the state of science in different branches; the facilities for the communication of knowledge; the supply of technical capabilities, skills, engineers, and so on; the conditions controlling occupational and geographical mobility and/or consumer promptness/resistance to change; market conditions, particularly in their bearing on interfirm competition and on demand growth; financial facilities and patterns and criteria of allocation of funds to the industrial firms; macroeconomic trends, especially in their effects on changes in relative prices of inputs and outputs; public policies (e.g., tax codes, patent laws, industrial policies, public procurement). It is impossible to consider here each of these factors in detail and the survey will focus upon the procedures, determinants, and effects of the innovative efforts of business firms; however, at each step of the analysis, I will try to show how those broader factors affect the opportunities, incentives, and capabilities of innovating in different firms and industries.

The empirical evidence rests on studies of several industries and technologies; however, particular attention is devoted to the characteristics and effects of microelectronics-based innovations. The obvious reason is the pervasiveness of these technologies and the scope of the transformations that they are inducing in the contemporary economic system.

Various forms of innovations affect all sectors of economic activity. The present discussion, however, concentrates on the production of goods (in primis, manufacturing) and it emphasizes the efforts concerned with the improvements of the techniques of production and the search for new products.

---

1 An extensive survey of the literature on innovation and technical change can be found in Freeman (1982). See also National Science Foundation (1983). A more specific survey on technical change and productivity growth is in Nelson (1981a). Other surveys of the economics of technological change, oriented more to the theoretical literature, include Charles Kennedy and Anthony Thirlwall (1981), and Paul Stoneman (1983).
### TABLE 1

**R & D Expenditure by Country and by Source of Finance, R & D Real Growth and R & D Employment**

<table>
<thead>
<tr>
<th>Country:</th>
<th>USA</th>
<th>Japan</th>
<th>West Germany</th>
<th>United Kingdom</th>
<th>France</th>
<th>Italy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yearly Percentage Growth of Total National (R &amp; D) Expenditures (at Constant Prices)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1969–75</td>
<td>-0.6</td>
<td>8.3</td>
<td>6.2</td>
<td>1.3</td>
<td>2.3</td>
<td>4.9</td>
</tr>
<tr>
<td>1975–81</td>
<td>4.2</td>
<td>7.9</td>
<td>4.7</td>
<td>3.1</td>
<td>4.2</td>
<td>4.6</td>
</tr>
<tr>
<td>1981–83</td>
<td>3.8</td>
<td>8.2</td>
<td>1.9</td>
<td>-0.7</td>
<td>4.7</td>
<td>4.9</td>
</tr>
<tr>
<td>Total R &amp; D as Percentage of GDP: 1983</td>
<td>2.7</td>
<td>2.8</td>
<td>2.8</td>
<td>2.8</td>
<td>2.5</td>
<td>1.6</td>
</tr>
<tr>
<td>Total R &amp; D Employment per Thousand of Total Labor Force: 1983</td>
<td>6.6</td>
<td>5.8</td>
<td>4.7*</td>
<td>3.6*</td>
<td>3.9</td>
<td>2.3</td>
</tr>
<tr>
<td>Business-financed R &amp; D as a Percentage of Total R &amp; D: 1983</td>
<td>49.0</td>
<td>65.3</td>
<td>58.1</td>
<td>42.1</td>
<td>42.0</td>
<td>45.5</td>
</tr>
<tr>
<td>Business-performed R &amp; D as a Percentage of Total R &amp; D: 1983</td>
<td>71.1</td>
<td>63.5</td>
<td>69.8</td>
<td>61.0</td>
<td>56.8</td>
<td>57.0</td>
</tr>
<tr>
<td>Military R &amp; D as Percentage of Total R &amp; D: 1983b</td>
<td>27.8</td>
<td>0.6</td>
<td></td>
<td></td>
<td></td>
<td>13.5c</td>
</tr>
</tbody>
</table>


Note: i) Unless otherwise specified, the data of rows 4 to 8 refer to 1983; ii) despite normalization efforts, stimulated in particular by the OECD, some discrepancies are still likely to appear among the various countries in coverage and definitions; iii) some caution should be used in comparing rows 4 and 5: the differences are the likely result of both statistical discrepancies and different relative wages of research workers to average workers in each country.

a 1981 (Source for R & D employment: National Science Foundation).
b Calculated by Patel and Pavitt (1986).
c All Western Europe.

In Part II I recall some stylized evidence on the allocation of resources to research and on the patterns of innovation across countries and sectors. The interpretation of these observed patterns will begin in Part III with an analysis of the characteristics of the search process aimed at the discovery and development of innovations. Part IV discusses the nature of the opportunities and knowledge on which innovations draw and the incentives leading profit-motivated actors to innovate and/or imitate other people's innovations. I argue that the suggested interpretation of the innovation process helps to explain why sectors differ in their modes and rates of innovation. Moreover, firms within each industry differ, too, in their propensity to innovate. Part V discusses this phenomenon. Finally, Part VI considers the relationship between innovative activities and the dynamics of industrial structures and performances.

### II. Searching for Innovations—The General Patterns

Modern industrial countries devote a significant share of their income and labor force to formalized activities of pure and applied research and technological development, within both nonprofit institutions (universities, government laboratories, etc.) and business enterprises. Table 1 provides an overview of employ-
TABLE 2
UNITED STATES R & D EXPENDITURES BY TYPE AND BY SOURCES OF FINANCE, VARIOUS YEARS (PERCENTAGES)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total R &amp; D</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Basic research</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Federal government</td>
<td>8.9 (100)</td>
<td>13.6 (100)</td>
<td>12.9 (100)</td>
<td>12.6 (100)</td>
</tr>
<tr>
<td>Industry</td>
<td>5.3 (59.7)</td>
<td>9.5 (70.1)</td>
<td>8.9 (68.8)</td>
<td>8.4 (66.4)</td>
</tr>
<tr>
<td>University and colleges</td>
<td>2.5 (28.6)</td>
<td>2.0 (14.9)</td>
<td>2.0 (15.7)</td>
<td>2.3 (18.4)</td>
</tr>
<tr>
<td>Other nonprofit institutes</td>
<td>0.5 (6.0)</td>
<td>1.3 (10.0)</td>
<td>1.3 (10.0)</td>
<td>1.3 (10.0)</td>
</tr>
<tr>
<td>Applied research</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Federal government</td>
<td>22.3 (100)</td>
<td>21.9 (100)</td>
<td>22.4 (100)</td>
<td>23.4 (100)</td>
</tr>
<tr>
<td>Industry</td>
<td>12.5 (55.9)</td>
<td>11.8 (53.8)</td>
<td>10.5 (47.0)</td>
<td>10.6 (45.4)</td>
</tr>
<tr>
<td>University and colleges</td>
<td>9.1 (40.6)</td>
<td>9.3 (42.4)</td>
<td>10.7 (47.7)</td>
<td>11.6 (49.6)</td>
</tr>
<tr>
<td>Other nonprofit institutes</td>
<td>0.5 (2.1)</td>
<td>0.4 (1.7)</td>
<td>0.7 (3.0)</td>
<td>0.7 (2.0)</td>
</tr>
<tr>
<td>Development</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Federal government</td>
<td>68.9 (100)</td>
<td>64.5 (100)</td>
<td>64.6 (100)</td>
<td>64.0 (100)</td>
</tr>
<tr>
<td>Industry</td>
<td>46.8 (68.1)</td>
<td>35.7 (55.3)</td>
<td>27.6 (42.7)</td>
<td>27.6 (43.1)</td>
</tr>
<tr>
<td>University and colleges</td>
<td>21.8 (31.7)</td>
<td>28.6 (44.4)</td>
<td>36.7 (56.7)</td>
<td>36.0 (56.3)</td>
</tr>
<tr>
<td>Other nonprofit institutes</td>
<td>0.01 (0.1)</td>
<td>0.0 (0.1)</td>
<td>0.1 (0.2)</td>
<td>0.02 (0.2)</td>
</tr>
</tbody>
</table>

Source: National Science Foundation (1986).

Note: i) Data in parentheses are percentages of each research category subtotal; ii) Subdivisions between "pure" research, "applied" research, and "development," are taken from NSF classifications.

* Based on preliminary estimates.

* Federally funded university-based research is included in the "federal government" source.

ment and expenditures on R & D by country, shares of business-performed research, and sources of finance.\(^2\)

As regards the composition of R & D expenditures (see Table 2 for evidence on the USA), about one-tenth is devoted to pure research, more than one-fourth to applied research, and the rest to development. Not surprisingly, pure research, with its character of relative publicness, is financed mainly by the federal government, universities, and other nonprofit institutions, while industry meets about one-half the cost of applied research and development; however, private industry also devotes roughly 20 percent of its total R & D expenditures to pure research.

Moreover, within the broad picture of national R & D investments, one observes marked intersectoral differences in the allocation of resources to research (see Table 3). As regards the sources of these investments and their institutional location, in contemporary market economies roughly half of the total investment in R & D is, as said, financed by business

\(^2\) In an effort to standardize definitions and data collection on research expenditures, the Organization of Economic Cooperation and Development (OECD) has proposed, in the so-called "Frascati Manual," that "Research and Experimental Development comprise creative work undertaken on a systematic basis in order to increase the stock of knowledge . . . and the use of this stock of knowledge to devise new applications" (OECD 1981, p. 25). Within that general definition, "pure" research broadly corresponds to activities aimed at knowledge growth, "applied" research involves the search for "applications," and "development" concerns the activities of design, implementation, and prototype manufacturing of the "new applications" themselves. Still the details of the activities actually surveyed in different countries—in terms of both expenditures and employment—are often not strictly homogeneous and some caution should be used in comparing the investment figures on R & D among different countries. For an in-depth discussion of R & D measurement problems, see Freeman (1982).
### TABLE 3
**Expenditures on Research and Development as a Percentage of Value Added by Sector and by Country and Sectoral Ratios of R & D Use to Expenditure**

<table>
<thead>
<tr>
<th>Sector</th>
<th>USA</th>
<th>Japan</th>
<th>West Germany</th>
<th>France</th>
<th>United Kingdom</th>
<th>Italy</th>
<th>Estimated USA Ratio of Use to Generation of R &amp; D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electric and electronics industries</td>
<td>12.7</td>
<td>8.5</td>
<td>8.8</td>
<td>13.7</td>
<td>16.2</td>
<td>5.7</td>
<td>0.34</td>
</tr>
<tr>
<td>Chemicals</td>
<td>6.5</td>
<td>7.7</td>
<td>5.8</td>
<td>7.0</td>
<td>6.8</td>
<td>5.5</td>
<td></td>
</tr>
<tr>
<td>Organic and inorganic chemicals</td>
<td>4.3</td>
<td>8.0</td>
<td>8.4</td>
<td>7.6</td>
<td>5.3</td>
<td>6.0</td>
<td>0.50</td>
</tr>
<tr>
<td>Drugs</td>
<td>12.1</td>
<td>10.0</td>
<td></td>
<td>7.6</td>
<td>17.8</td>
<td></td>
<td>0.17</td>
</tr>
<tr>
<td>Petroleum refineries</td>
<td>6.4</td>
<td>3.0</td>
<td>0.6</td>
<td>3.4</td>
<td>2.0</td>
<td>4.6</td>
<td>1.31</td>
</tr>
<tr>
<td>Instruments</td>
<td>20.5</td>
<td>(8.6)</td>
<td>8.3</td>
<td>(5.4)</td>
<td>8.5</td>
<td>(1.2)</td>
<td>0.14</td>
</tr>
<tr>
<td>Office machinery and computers</td>
<td>21.7</td>
<td>7.5</td>
<td>4.2</td>
<td>2.4</td>
<td>19.8</td>
<td>2.7</td>
<td>0.11</td>
</tr>
<tr>
<td>Industrial nonelectrical machinery</td>
<td>2.5</td>
<td>2.9</td>
<td></td>
<td>2.5</td>
<td></td>
<td></td>
<td>0.17</td>
</tr>
<tr>
<td>Aerospace</td>
<td>32.6</td>
<td></td>
<td>7.2</td>
<td>30.8</td>
<td>10.0</td>
<td>30.9</td>
<td>0.37</td>
</tr>
<tr>
<td>Transport equipment</td>
<td>10.0</td>
<td></td>
<td>5.5</td>
<td>3.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Motor vehicles</td>
<td>12.6</td>
<td>6.5</td>
<td>5.9</td>
<td>n.a.</td>
<td>4.2</td>
<td>n.a.</td>
<td>0.20</td>
</tr>
<tr>
<td>Ships</td>
<td>7.8</td>
<td>1.2</td>
<td>n.a.</td>
<td>0.8</td>
<td></td>
<td>n.a.</td>
<td></td>
</tr>
<tr>
<td>Other transport equipment</td>
<td>n.a.</td>
<td>n.a.</td>
<td>1.6</td>
<td>n.a.</td>
<td>0.0</td>
<td>n.a.</td>
<td>0.32</td>
</tr>
<tr>
<td>Food, drink, and tobacco</td>
<td>0.7</td>
<td>1.3</td>
<td>0.5</td>
<td>0.3</td>
<td>0.8</td>
<td>2.4</td>
<td>1.18</td>
</tr>
<tr>
<td>Textile and clothing</td>
<td>2.7</td>
<td>1.3</td>
<td>0.5</td>
<td>0.5</td>
<td>0.3</td>
<td>0.3</td>
<td>1.31</td>
</tr>
<tr>
<td>Rubber and plastic products</td>
<td>2.5</td>
<td>2.8</td>
<td>1.9</td>
<td>4.4</td>
<td>1.1</td>
<td>1.8</td>
<td>1.12</td>
</tr>
<tr>
<td>Ferrous metals</td>
<td>1.6</td>
<td>2.9</td>
<td>1.6</td>
<td>1.1</td>
<td>1.1</td>
<td>0.5</td>
<td>1.63</td>
</tr>
<tr>
<td>Nonferrous metals</td>
<td>2.4</td>
<td>4.3</td>
<td>1.8</td>
<td>2.4</td>
<td>2.1</td>
<td>3.2</td>
<td>1.06</td>
</tr>
<tr>
<td>Fabricated metal products</td>
<td>1.1</td>
<td>1.2</td>
<td>1.4</td>
<td>1.0</td>
<td>0.8</td>
<td>0.0</td>
<td>0.49</td>
</tr>
<tr>
<td>Lumber, wood products, and furniture</td>
<td>0.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.33</td>
</tr>
<tr>
<td>Paper and printing</td>
<td>0.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.31</td>
</tr>
<tr>
<td>Stone, clay, and glass</td>
<td>1.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.86</td>
</tr>
<tr>
<td>Total manufacturing</td>
<td>8.1</td>
<td>4.9</td>
<td>5.4</td>
<td>(4.6)</td>
<td>6.6</td>
<td>(2.9)</td>
<td>0.42</td>
</tr>
</tbody>
</table>

**Sources:** OECD (1986), National Science Foundation (1986), OECD, *Industrial Structures Statistics*, various years, and Scherer (1982); data of R & D expenditures and value added have been aggregated, whenever necessary, by the author for comparability purposes.

**Notes:** The sectoral R & D intensities are calculated as the ratio of business-performed R & D to sectoral value added.

Special caution must be taken in comparing the data along any one row: The coverage of value-added data differ among countries (e.g., for Italy, it includes only firms with more than 20 employees).

* Ratio of the total R & D used to the R & D performed by the sector as estimated by Frederick M. Scherer (1982).
* Professional instruments include photographic equipment.
* No comparable data available.
* Estimates based on the subset of manufacturing for which sectoral data are available.
* Based on aggregate OECD data on the Italian economy.

and roughly between half and two-thirds of R & D is carried out by business firms (cf. Table 1).

Of course, Tables 1 through 3 show only the commitment of resources to innovation that fund formalized research activities, typically in R & D laboratories; however, in addition to formalized R & D, and in many ways complementary to it, a significant amount of innovation and improvements is originated through design improvements, “learning by doing,”
and “learning by using” (see, for example, Kenneth Arrow 1962a; Rosenberg 1982; David 1975; Samuel Hollander 1965; Louis Yelle 1979). Such informal effort is generally embodied in people and organizations (primarily firms) (David Teece 1977, 1986; Keith Pavitt 1986a), and its cost is hard to trace. Again, sectors differ in the relative importance of the four basic modes of technological advance, namely (a) economically expensive and formalized processes of search whose costs are measured in the tables; (b) informal processes of diffusion of information and of technological capabilities (e.g., via publications, technical associations, watch-and-learn processes, personnel transfers); (c) those particular forms of “externalities,” internalized within each firm, associated with learning by doing and learning by using; and (d) the adoption of innovation developed by other industries and embodied in capital equipment and intermediate inputs (cf. Pavitt 1984).

In the interpretation of the evidence on innovative activities in contemporary economies, one faces, first, the question of the nature of the process leading from a perception of an economically exploitable opportunity to its actual development: That is, what do people actually do? How do they search? Why do sectors differ in their search procedures?

Second, one should account for the observed directions of technological change: To what extent do such observed patterns represent reactions to market signals? Are there other factors that influence the patterns of technological change?

Third, one should explain why sectors differ in their commitment of resources to search activities and in the rates at which they generate new products and processes of production. In short, I call “propensity to innovate” the empirical outcome of both sets of phenomena and try to disentangle its determinants.

In the following, I deal, in turn, with these questions.

III. Innovation: The Characteristics of the Search Process

Over the past 20 years, various analyses have been made of the process of innovation, concerned with both the relationship between inputs and outputs of innovative activities (that is, the relationship between the resources devoted to innovative search and rates of generation of innovations, however measured) and the nature of the innovation process itself. In this section I focus first on the second issue.

These analyses, which can be classified under the broad heading of “innovation studies” (Zvi Griliches 1984b), include those of William Abernathy and James Utterback (1975, 1978), E. W. Constant (1980), David (1975), Freeman (1982), Burton Klein (1977), Nelson and Sidney Winter (1977, and 1982), Rosenberg (1976, 1982), Devandra Sahal (1979, 1981, 1985), Pavitt (1979, 1984), Eric von Hippel (1979, and 1982), and Dosi (1982, 1984). The analytical aims of these studies are different and their contributions quite heterogeneous. Nonetheless, most of them point toward some common characteristics of innovation which, in my view, are of crucial importance in the economics of technological change.

A. Innovation as Problem-solving: Technological Paradigms

In very general terms, technological innovation involves the solution of problems—for example, on transformation of heat into movement, shaping materials in certain ways, producing compounds with certain properties—meeting at the same time some cost and marketability requirements. Typically, the problems are “ill structured,” in that the available
information (e.g., on the limits in the cutting speed of a certain machine, the physical reasons it breaks at higher speed) does not provide by itself a solution to the problem (relevant discussions of this class of problems are in Herbert Simon 1973, 1979; and Nelson and Winter 1982; see also Massimo Egidi 1986 and Dosi and Egidi 1987). In other words, an “innovative solution” to a certain problem involves “discovery” and “creation,” since no general algorithm can be derived from the information about the problem that generates its solution “automatically” (more on this in Dosi and Egidi 1987). Certainly, the “solution” of technological problems involves the use of information drawn from previous experience and formal knowledge (e.g., from the natural sciences); however, it also involves specific and uncodified capabilities on the part of the inventors. Following Nelson and Winter (1982) and Winter (1984), I use the term knowledge base for the set of information inputs, knowledge, and capabilities that inventors draw on when looking for innovative solutions. A first characterization that can be made of different technologies is in terms of the degrees of “publicness” and universality versus tacitness and specificity of their knowledge bases (Winter 1984). Following Michael Polanyi (1967), tacitness refers to those elements of knowledge, insight, and so on that individuals have which are ill defined, uncodified, unpublished, which they themselves cannot fully express and which differ from person to person, but which may to some significant degree be shared by collaborators and colleagues who have a common experience. Conversely, scientific inputs are typically universal and public. Nelson (1986) cites the results of the Yale questionnaire, showing that in 30 sectors out of 130, university research—especially in chemistry, materials science, computer science, and metallurgy—is considered to be very important for sectoral innovativeness; in the cases of biotechnologies François Chesnais (1986) analyzes a complex thread of joint ventures between university and industry. Also the knowledge base in several chemical sectors is directly linked to scientific knowledge on chemical/physical properties of complex organic molecules.

However, even in these rather science-based activities and, more so, in other technologies, public knowledge is complementary to more specific and tacit forms of knowledge generated within the innovating units (for evidence, see Freeman 1982; SPRU 1972; J. Langrish 1972; Michael Gibbons and Ron Johnston 1974; and Pavitt 1984). For example, in mechanical engineering (e.g., machine tools) an important part of the knowledge base consists of tacit knowledge about the performance of previous generations of machines, their typical conditions of use, the productive requirements of the users, and so on. In the case of microelectronics, one finds three major and complementary forms of knowledge, namely (a) advances in solid-state physics (e.g., electrical properties of semiconductors at the micron/submicron level) (b) knowledge related to the construction of semiconductor manufacturing and testing equipment, and (c) programming logics. As regards the applications of microelectronics, embodied in components and equipment, the fundamental forms of knowledge consist of (a) systems architectures and systems engineering; (b) programming logics (ranging from the logics embodied in the “firmware” of computers, to the proper applicative software), (c) the interfaces between information processing and the mechanical or chemical processes to which it is applied (e.g., the interfaces between an electronic control and the mechanical movements of a machine tool or the flows in a chemical
plant), and (d) the interacting devices (e.g., sensors).

The crucial point is that this (technology-specific and sector-specific) variety in the knowledge base of innovative search implies also different degrees of tacitness of the knowledge underlying innovative success and, as will be discussed below, also helps explain the differences across sectors in the typical organization of research activities. Whatever the knowledge base on which innovation draws, each problem-solving activity implies the development and refinement of “models” and specific procedures.

Elsewhere (Dosi 1982, 1984), I suggest a broad similarity, in terms of definition and procedures, between science and technology. More precisely, as modern philosophy of science suggests the existence of scientific paradigms (or scientific research programs), so there are technological paradigms. Both scientific and technological paradigms embody an outlook, a definition of the relevant problems, a pattern of enquiry. A “technological paradigm” defines contextually the needs that are meant to be fulfilled, the scientific principles utilized for the task, the material technology to be used. In other words, a technological paradigm can be defined as a “pattern” of solution of selected technoeconomic problems based on highly selected principles derived from the natural sciences, jointly with specific rules aimed to acquire new knowledge and safeguard it, whenever possible, against rapid diffusion to the competitors. Examples of such technological paradigms include the internal combustion engine, oil-based synthetic chemistry, and semiconductors. A closer look at the patterns of technical change, however, suggests the existence of “paradigms” with different levels of generality, in several industrial sectors.

A technological paradigm is both an exemplar—an artifact that is to be developed and improved (such as a car, an integrated circuit, a lathe, each with its particular technoeconomic characteristics)—and a set of heuristics (e.g., Where do we go from here? Where should we search? What sort of knowledge should we draw on?).

These aspects of technological change which relate to the improvement of some typical performance attributes of exemplars (e.g., four-wheeled internal-combustion cars, jet aircraft) underlie Sahal’s idea of “technological guide posts” (Sahal 1981, 1985), a guidepost being the basic artifact whose technoeconomic characteristics are progressively improved. Basic artifacts (such as car) are also functionally specified (e.g., a car’s locomotive attributes) in relation to some use in the socioeconomic system (a car is used jointly with human time for household mobility and also in market production activities).

(For an attempt to map characteristics of technological paradigms and socioeconomic uses or “needs,” see Paolo Saviotti and J. Stanley Metcalfe 1984.) In this respect, technological paradigms define “bundles” of characteristics of the various commodities. If, following Kevin Lancaster (1971), the latter are defined in terms of combination of hedonic attributes, technological paradigms restrict the actual combinations in a notional characteristics space to a certain number of prototypical bundles.

On the other hand, the development and improvement of these basic “exemplars” involve the development of specific competences and “rules.” Rosenberg (1976) highlights the importance of “focusing devices,” that is, typical problems, opportunities, and targets that tend to focus the search process in particular directions.

Of course, the procedures, competences, and heuristics involved in the search process are, to varying degrees, specific to each technology. In other
words, each technological paradigm involves a specific "technology of technical change." For example, in some sectors (such as organic chemicals), these procedures relate to the ability of coupling basic scientific knowledge with the development of new molecules that present the required characteristics. Thus, one often searches around the existing compounds, helped by the scientific knowledge of the relationship between chemical structures and physical properties, by previous experience, and by chance. In other sectors (such as microelectronics devices) the methods of innovative search involve scientific advances on submicron electrical flows in semiconductors, the development of more sophisticated hardware capable of "writing" the chips at the desired level of miniaturization, and advances in the programming logic to be built into the chips. In mechanical engineering, the search process is generally "focused" by trade-offs involved in the use of machines (e.g., between speed, flexibility to different uses, and cutting precision). The skills required by this search process typically involve also unwritten and relatively tacit experience in design and use of mechanical equipment, and more recently, in the interface between electronic controls and mechanical movements. Yet in other sectors (e.g., the top end of textile, clothing, leather, and shoemaking) fundamental "search skills" are the capabilities of understanding/anticipating/influencing the trends in tastes and fashion.

It quite often happens that prototypical problem-solving models, rules on how to search and on what targets to focus, and beliefs as to "what the market wants" become the shared view of the engineering community. A paradigm is economically exploited and reproduced over time also through the development of institutions that train the would-be practitioners in methods for the improvement of basic exemplars, and peers' judgments are also based on the success achieved in the refinement and use of these methods (in this respect, Noble's history of the development of American engineering schools and their relationship with industry and Hughes' history of electrification are vivid illustrations of the institutional process that goes together with the establishment of "technological paradigms;" see David Noble 1987, and Thomas Hughes 1982).

B. Technological Paradigms and Patterns of Innovation: Technological Trajectories

A crucial implication of the general paradigmatic form of technological knowledge is that innovative activities are strongly selective, finalized in quite precise directions, cumulative in the acquisition of problem-solving capabilities. This accounts also for the relatively ordered patterns of innovation that one tends to observe at the level of single technologies, as shown by several studies of "technological forecasting" (for a comprehensive review and discussion, see Joseph Martino 1976). Let us define as a technological trajectory (Nelson and Winter 1977; Sahal 1981, Dosi 1982, Theodore Gordon and Thomas Munson 1981; Savioiti and Metcalfe 1984) the activity of technological process along the economic and technological trade-offs defined by a paradigm.

Thus, for example, technological progress in aircraft technology has followed two quite precise trajectories (one civilian and one military) characterized by log-linear improvements in the trade-offs between horsepower, gross takeoff weight, cruise speed, wing loading, and cruise range (Sahal 1985 and an oral com-

---

3 This was also the title of an important conference, coordinated by R. Nelson at the Royal College of Arts, London, July 1985. See also Nelson (1981b).
munication of P. Saviotti on ongoing research at Manchester University). In microelectronics, technical change is accurately represented by an exponential trajectory of improvement in the relationship between density of the electronic chips, speed of computation, and cost per bit of information (Dosi 1984). More generally, there is growing evidence that specific “innovation avenues” are a widespread feature of the observed patterns of technical change (Sahal 1985).

Of course, there is no a priori economic reason why one should observe limited clusters of technological characteristics at any one time and ordered trajectories over time. Indeed, given consumers with different preferences and equipment users with different technical requirements, if technology had the malleable attributes of information and if the innovative search were a purely random process, one would tend to observe sorts of “technological indifference curves” at any one time, and, over time, random search all over the n-dimension characteristics space. Of course, “how different” are consumers and users of goods, pieces of equipment, intermediate components, is, in principle, an empirical question. However, relatively wide differences (given the high dimensionality of the space of characteristics/technical requirements demanded by consumers/users of commodities) cannot be ruled out by either casual empiricism or general theoretical arguments. Moreover, for whatever distribution of characteristics at any arbitrary time t, one should expect that income growth and division of labor among different productive activities would increase such diversity of micro demands. Were technologies simply pieces of information (or “recipes”) that could be added, convexly combined, etc., one would also tend to observe an increasingly dispersed variety of technical and performance combinations in actual products and production inputs. Over time, this would lead toward the exploration of the entire characteristics space of final products, machine tools, components, etc. Indeed, the evidence surveyed suggests that one still observes “explorations” limited to some, much smaller, subsets of the notional characteristics space. It is precisely the paradigmatic cumulative nature of technological knowledge that accounts for the relatively ordered nature of the observed patterns of technological change.

Engineers typically try to improve the desirable characteristics that are specific to a certain product, tool, or device, keeping in mind the trade-offs among them. Relatedly, historical evidence strongly suggests that a major impulse to innovation has derived from imbalances between the technical dimensions that characterize a “trajectory” (or “avenue”) e.g., between cutting speed and tool resistance in machine tools or shuttle speed in eighteenth century looms and spinning speed in spindles. For a discussion of several examples of this process of solution of technical imbalances, which Hughes (1987) calls “adverse salients” and “critical problems,” see Rosenberg (1976, especially chapter 6). Arguments broadening the scope of “imbalances” to the relationships between technical change and social roles and behaviors of different groups of workers are in William Lazonick (1979, 1987), and von Tunzelmann (1982). Other examples can be drawn from David Landes (1969).

Conversely, a change in the paradigm generally implies a change in the trajectories: Together with different knowledge bases and different prototypes of artifacts, the technoeconomic dimensions of innovation also vary. Some characteristics may become easier to achieve, new desirable characteristics may emerge, some others may lose importance. Relatedly, the engineers’ vision of future tech-
nological advances changes, together with a changing emphasis on the various trade-offs that characterize the new artifacts. Thus, for example, the technological trajectory in active electrical components based on thermionic valves had, as fundamental dimensions, heat-loss parameters, miniaturization, and reliability over time. With the appearance of solid-state components, heat loss became much less relevant, while miniaturization increased enormously in importance and also the rates at which progress could be achieved shot up. More generally, it has also been suggested that major clusters of prevailing technological paradigms (e.g., those related to oil-based synthetic chemistry, to electromechanical production, or, more recently, to microelectronics) involve the intensive utilization of some crucial input abundantly available at low cost (e.g., energy in the former two examples, and information-processing in the latter; Carlota Perez 1987).

C. Technology: Freely Available Information or Specific Knowledge?

The view of technology just presented is very different from the concept of technology as information that is generally applicable, and easy to reproduce and reuse (Arrow 1962b), one where firms can produce and use innovations by dipping freely into a general “stock” or “pool” of technological knowledge. It implies that firms produce things in ways that are differentiated technically from the products and methods of other firms and that they make innovations largely on the basis of in-house technology, but with some contributions from other firms, and from public knowledge. In such circumstances, the search process of industrial firms to improve their technological capabilities is not likely to be one where they survey the whole stock of notional technological knowledge before making their technical choices (see Nelson and Winter 1982). Given its highly differentiated nature, firms will instead seek to improve and to diversify their technology by searching in zones that enable them to use and to build on their existing technological base and also on their existing markets, distribution arrangements, and so on (Teece 1982, 1986). In other words, technological search processes in each firm are cumulative processes too. What the firm can hope to do technologically in the future is narrowly constrained by what it has been capable of doing in the past.

The distinction between technology and information—with the latter being only a subset of the former—entails important analytical consequences for the theory of production. To illustrate that distinction let us take a scientific analogy (note also that science is somewhat closer to information in that the ethos of the scientific community is to disclose results, while in privately generated technology it is to withhold and appropriate them, see Partha Dasgupta and David 1985). Certainly, a good part of “science” can be embodied in “information.” There are freely available journals, textbooks, and university lectures that disseminate this information. Moreover, there are market conditions of access to it; for example, there is a market for textbooks and economic conditions of access to higher education (e.g., the level of registration fees, the availability or scarcity of grants for students unable to support themselves); however, in any proper sense of the word, getting a PhD is not simply acquiring information, and it is even less true to say that there is a market for PhDs. In this analogy, “information” stands vis-à-vis innovative technological capabilities as a subscription to the American Economic Review stands vis-à-vis

---

4 This paragraph is partly based on Dosi, Pavitt, and Soete (1988), which in turn draws from Pavitt (1984d).
winning the Nobel Prize in economics. In both cases there is an irreducible element that is not information and cannot be bought and sold, but rather depends on cumulatively augmented abilities and skills. In each technology there are elements of *tacit and specific* knowledge that are not and *cannot* be written down in a “blueprint” form, and cannot, therefore, be entirely diffused either in the form of public or proprietary information (see Polanyi 1967 and the discussion of this same issue in Nelson and Winter 1982). Of course, this does not imply that such skills and forms of tacit knowledge are entirely immobile: People can be hired away from one firm to another or can start their own firms (and sometimes supply goods and knowledge to competitors of their own original firm), learning procedures of one firm may be imitated by other firms, and so on. It still holds, however, that the innovative activities present—to different degrees—firm-specific, local, and cumulative features. This is borne out by empirical studies.

It has been found that *information* about what other firms are doing spreads quite quickly (Edwin Mansfield 1985); however, the ability to produce or replicate innovative results is much more sticky. Successful innovations are more closely related to firms’ existing ranges of technological and marketing skills than unsuccessful ones (Robert Cooper 1983; Modesto Maidique 1983); they tend to occur in product fields proximate to firms’ current fields; the activities that firms undertake entail initial learning costs that are recovered later as a consequence of cumulative improvements in product performance and in wider market applications (John Enos 1962; David 1975; Rosenberg 1976, 1982; Sahal 1981; Morris Teubal 1982; Paul Gardiner 1984; Roy Rothwell and Gardiner 1984).

Once the cumulative and firm-specific nature of technology is recognized, its development over time ceases to be random, but is constrained to zones closely related technologically and economically (e.g., related markets and distribution networks) to existing activities. If those zones can be identified, measured, and explained, it also is possible to predict likely future patterns of innovative activities in firms, industries, and countries (see David 1975; Sahal 1981, 1985; Pavitt 1984; Dosi, Pavitt, and Luc Soete 1988).

Each technological paradigm, I suggest, entails a specific balance between exogenous determinants of innovation (e.g., university-based advances in pure science) and determinants that are endogenous to the process of competition and technological accumulation of particular firms and industries. Moreover, each paradigm involves specific *search modes, knowledge bases, and combinations between proprietary and public forms of technological knowledge.*

Given these features of technology and technological innovation, how are search processes organized? Who are the actors that undertake them? How do they relate to the rest of the economic system?

D. How Organizations Build Knowledge Bases

The increasing complexity of technologies and research activities in this cen-
tury militates in favor of formal organizations (R & D laboratories of big firms, government and university labs, etc.) as opposed to individual innovators, as the most conducive environments to the production of innovations. This is also shown by the secular growth in the share of corporate as opposed to individual patents registered in the USA as well as other western economies.

David Mowery (1980, 1983) has reconstructed the growth of research and development activities in American industry from the beginning of this century. Notably, he finds that industry-performed R & D—which grows at a much higher rate than industrial output or employment—also tends to be internalized within manufacturing companies. In other words, contrary to Stigler’s hypothesis (George Stigler 1956), R & D growth has not led to a comparable process of market-based division of labor and the emergence of specialized “innovation suppliers.” Inhouse R & D is the dominant form of organization for corporate technological search (on this point, see also Leonard Reich 1985; Rosenberg 1985; and Nelson 1986). As Richard Nelson puts it, “the modern industrial R & D laboratory, linked within the firm with production and often marketing, had a number of advantages over reliance on outside research and development laboratories, particularly when aspects of the relevant technologies were somewhat idiosyncratic and tacit, and R & D needed to be tailored to those idiosyncracies and to particular firm strategies. In addition to the general advantage of integration in such circumstances stressed by Oliver Williamson (1985), here, as Mowery has stressed, integration had the additional advantages of facilitating better information flow from the R & D laboratory to those who would have to implement the new technology, and from the latter to the former. It also served to limit cross-organization information leaks” (Nelson 1986, p. 10). Of course, one often observes also market transfers of innovations and technical competences—such as licensing and consultancy deals; however, “the predominant mode of industrial research in the private sector, at least in the United States, is the integrated research organization, part of a business enterprise which engages in at least one other activity vertically related to research and development such as manufacturing, marketing, distribution, sales and service” (Teece 1986, p. 1). Moreover, even when licensing and other forms of interfirm transfer of technology occur, they do not stand as an all-or-nothing substitute for in-house search: One needs to have substantial in-house capacity in order to recognize, evaluate, negotiate, and finally adapt the technology potentially available from others.

Williamson’s analysis (1975, 1985) of the costs of transactions involving informational asymmetries, monitoring problems, and possibilities of opportunistic behavior is clearly part of the interpretation of this phenomenon: Market transactions involving research activities generally imply (a) incomplete specifications of contracts, given the uncertainty about the research outcomes; (b) lack of adequate protection of proprietary information; (c) possibilities of “lock-in” phenomena with research suppliers, who can subsequently earn rents from that asymmetric advantage; (d) weak incentives to least cost performance; (e) monitoring costs (on all these points, see Teece 1988).

In addition (and complementary) to these transaction-related factors, however, the foregoing discussion of the nature of technology and innovative search suggests another set of factors related to the characteristics of knowledge and problem solving. Indeed, the heuristics
on "how to do things" and "how to improve them" are often embodied in organizational routines, which, through practice, repetition, and more or less incremental improvements make certain firms "good" at exploring certain technical opportunities and translating them into specific marketable products. In such matters, there is a significant amount of organizational indivisibility, because organizational learning may well not be additive in the learning of individuals or groups who compose the organization: indeed, it was Adam Smith who first emphasized the possible dichotomy between "system learning" (e.g., the beneficial effects on economic efficiency of the division of labor), on the one hand, and the degrading brutality which repetitive and mindless tasks could imply for some groups of workers, on the other. Intrafirm processes of specialization and division of labor are good examples of this possibility. Individuals and groups may well decrease the scope of knowledge and competences that they are required to put into production or innovative search (in a sense, they may be required to "forget"), while at the same time these same individuals and groups become linked through routines that increase organizational efficiency (on whatever criterion the latter is evaluated). For example, the emergence of the modern factory has also implied "deskilling" of particular categories of craftsmen; the abilities of several groups of artisan-like workers became redundant, the skills of making particular machines became increasingly separated from the skills involved in using them; the introduction of electromechanical techniques of automated mass production in big plants has further reduced the knowledge required of significant portions of the work force. These same processes, however, have been associated with major increases in the abilities of (more and more complex) business organi-

ciations to learn, that is to "store" and develop internally, procedures for a growing production efficiency.

The exploration of the characteristics of the organizational competences with specific reference to research an innovation is still at an early stage (see Pavitt 1986a; Teece 1986, 1988; Winter 1987a, 1987b; Neil Kay 1979, 1982); however, in my view, they are a fundamental ingredient (together with transaction costs and monitoring factors emphasized by Williamson 1985) of the explanation of both the integration of research within production/marketing units and, more generally, of the boundaries of the firms in contemporary market economies. More precisely, Teece (1986) and Pavitt, Mike Robson and Joe Townsend (1987) have independently put forward the conjecture that these boundaries are defined by the scope of their "core competences," that is, loosely speaking, by the scope of what "they are good at" and the relevance of this specific knowledge to the activities of innovation, production, and marketing of a certain commodity. This—it is suggested—affects also the scope of efficient vertical integration and diversification of any one firm (more on this in Dosi, Teece, and Winter 1987).

Organizational routines and higher-level procedures to alter them in response to environmental changes and/or to failures in performance embody a continuous tension between efforts to improve the capabilities of doing existing things, monitor existing contracts, allocate given resources, on the one hand, and the development of capabilities for doing new things or old things in new ways. This tension is complicated by the intrinsically uncertain nature of innovative activities, notwithstanding their increasing institutionalization within business firms. The technical (and, even more so, the commercial) outcome of research activities can hardly be known ex
ante (for empirical evidence on individual research projects, see Mansfield 1968 and Mansfield et al. 1977). In general, the uncertainty associated with innovative activities is much stronger than that with which familiar economic model deals. It involves not only lack of knowledge of the precise cost and outcomes of different alternatives, but often also lack of knowledge of what the alternatives are (see Freeman 1982; Nelson 1981a; Nelson and Winter 1982). In fact, let us distinguish between (a) the notion of uncertainty familiar to economic analysis defined in terms of imperfect information about the occurrence of a known list of events and (b) what we could call strong uncertainty whereby the list of possible events is unknown and one does not know either the consequences of particular actions for any given event (more on this in Dosi and Egidi 1987). I suggest that, in general, innovative search is characterized by strong uncertainty. This applies, in primis to those phases of technical change that could be called preparadigmatic. During these highly exploratory periods one faces a double uncertainty regarding both the practical outcomes of the innovative search and also the scientific and technological principles and the problem-solving procedures on which technological advances could be based. When a technological paradigm is established, it brings with it a reduction of uncertainty, in the sense that it focuses the directions of search and forms the grounds for forming technological and market expectations more surely. (In this respect, technological trajectories are not only the ex post description of the patterns of technical change, but also, as mentioned, the basis of heuristics asking "where do we go from here?"). However, even in the case of "normal" technical search (as opposed to the "extraordinary" exploration associated with the quest for new paradigms) strong uncertainty is present. Even when the fundamental knowledge base and the expected directions of advance are fairly well known, it is still often the case that one must first engage in exploratory research, development, and design before knowing what the outcome will be (what the properties of a new chemical compound will be, what an effective design will look like, etc.) and what some manageable results will cost, or, indeed, whether very useful results will emerge (Mansfield et al. 1977).

As a result, firms tend to work with relatively general and event-independent routines (with rules of the kind "... spend x% of sales on R & D;" "... distribute your research activity between basic research, risky projects, incremental innovations according to some routine shares . . . " and sometimes metarules of the kind "with high interest rates or low profits cut basic research," etc.). This finding is corroborated by ample managerial evidence and also by recent more rigorous econometric tests; see Griliches and Ariel Pakes (1986) who find that "the pattern of R & D investment within a firm is essentially a random walk with a relatively low error variance" (pp. 10–11). In this sense, Schumpeter's hypothesis about the routinization of innovation (Joseph Schumpeter 1942) and the persistence of innovation-related uncertainty must not be in conflict but may well complement each other. As suggested by the "late" Schumpeter, one may conjecture that large-scale corporate research has become the prevailing form of organization of innovation because it is most effective in exploiting and internalizing the tacit and cumulative features of technological knowledge (Mowery 1980; Pavitt 1986). Moreover, companies tend to adopt steady policies (rules), because they face complex and unpredictable environments where they cannot forecast future states of the world, or
even “map” notional events into actions, and outcomes (Dosi and Orsenigo 1986; Heiner 1983, 1988). Internalized corporate search exploits the cumulative nature and complexity of technological knowledge. Together with steady rules, firms try to reduce the uncertainty of innovative search, without, however, eliminating it.

Internalization and routinization in the face of the uncertainty and complexity of the innovative process also point to the importance of particular organizational arrangements for the success or failure of individual innovative attempts. This is what was found by the SAPHO Project (cf. Science Policy Research Unit 1972 and Rothwell et al. 1974), possibly the most extensive investigation of the sources of commercial success or failure of innovations: Institutional traits, both internal to the firm—such as the nature of the organizational arrangements between technical and commercial people, or the hierarchical authority within the innovating firm—and between a firm and its external environment—such as good communication channels with users, universities, and so on—turn out to be very important. Moreover, it has been argued (Pavitt 1986; Robert Wilson, Peter Ashton, and P. Thomas Egan 1984) that, for given incentives and innovative opportunities, the various forms of internal corporate organization (U form versus M form, centralized versus decentralized, etc.) affect innovation and commercial success positively or negatively, according to the particular nature of each technological paradigm and its stage of development.

In general, each organizational arrangement of a firm embodies procedures for resource allocation to particular activities (in our case, innovative activities), and for the efficient use of these resources in the search for new products, new processes, and procedures for improvements in existing routines; however, the specific nature of these procedures differs across firms and sectors. For example, the typical degrees of commitment of resources vary by industry and so do the rates at which learning occurs. I now turn to the interpretation of these phenomena.

IV. Opportunities, Incentives, and the Intersectoral Patterns of Innovation

Clearly, the commitment of resources by profit-motivated agents must involve both the perception of some sort of opportunity and an effective set of incentives. Are the observed intersectoral differences in innovative investment the outcome of different incentive structures, different opportunities, or both? Jacob Schmookler, in his classic work, argued that the serendipity and universality of modern science provide a wide and intersectorally indifferent pool of opportunities that are exploited to different degrees in each economic activity according to differential economic incentives, and, in particular, to different patterns of demand growth (Schmookler 1966). (In fact, he was not so much concerned with innovative investments as with innovative outputs, which he measured by patents. However, the same argument applies: For identical opportunities, the elasticity of innovative outputs to R & D inputs should be the same.) Schmookler’s thesis has been criticized on both theoretical and empirical grounds (see Rosenberg 1976, chapter 15, and Freeman 1982). The foregoing analysis of the innovation process supports these criticisms and helps to clarify the merits and limitations of Schmookler’s hypothesis.

A. Technological Opportunities:

Exogenous Science and Specific Learning

I first discuss the role of science-related opportunities for innovation and,
then, the importance of other sources of opportunities.

Scientific knowledge plays a crucial role in opening up new possibilities of major technological advances. In this century, the emergence of major new technological paradigms has frequently been directly dependent and directly linked with major scientific breakthroughs; see, for example, the origin of synthetic chemistry (John Beer 1959; Freeman 1982), the transistor (Nelson 1962; H. S. Kleiman 1977; Dosi 1984), and bioengineering (Orsenigo 1988). Certainly, in western civilization there is a long history of linkages between science and technology, hinting at rather close feedbacks, at least since Leonardo da Vinci and Galileo. What is new and increasingly important in this century is that the generation and utilization of part of the scientific knowledge is internal to, and often a necessary condition of, the development of new technological paradigms. Until the end of the nineteenth century, technological innovations were typically introduced by imaginative craftsmen; for example, engines were developed by practical-minded inventors well before the works of Carnot on thermodynamics. In this century, as far as major innovations are concerned, one moves somewhat closer to the “transistor archetype,” whereby the discovery of certain quantum mechanics properties of semiconductors, yielding a Nobel Prize for physics, and the technological development of the first microelectronics device have been one and the same thing (Nelson 1962; Ernest Braun and Stuart MacDonald 1978; Dosi 1984).

Prima facie, the increasing role of scientific inputs in the innovative process can be taken as evidence of the importance of factors exogenous to competitive processes among private economically motivated actors. This is true, subject, however, to two qualifications.

First, the link between science and technology runs also from the latter to the former. It has been noted, for example, that the development of scientific instruments has exerted a major impact on subsequent scientific progress. In general, however, the scope, timing, and channels of influence of technological advances on science have a different nature from the more direct influence of scientific discoveries on technological opportunities. A detailed discussion is beyond the limits of this survey. (On these topics, see John Bernal 1939; Rosenberg 1982; and Derek de Solla Price 1984.) Second, scientific advances play a major direct role, especially at an early phase of development of new technological paradigms. It is often the case that the establishment of a major new paradigm involves also the solution of problems of a theoretical nature and/or the development of devices, compounds, molecules, and so on which are themselves challenging tests for the scientists (the transistor, polypropylene, and genetic engineering are obvious examples).

In a sense, progress in general scientific knowledge yields a widening pool of potential technological paradigms. In another work (Dosi 1984), I analyze the specific mechanisms through which a few of these potential paradigms are actually developed, economically applied, and often become dominant. Here, suffice it to say that this process of selection depends, in general, on (a) the nature and the interests of the “bridging institutions” (Freeman 1982) between pure research and economic applications (these institutions, which can be private establishments, such as Bell Labs, or public organizations, are instrumental in applying theoretical advances to the development of practical devices even under remote or nonexistent direct economic incentives); (b) quite often, especially in this century, strictly institutional factors,
such as public agencies (e.g., the military); (c) the trial-and-error mechanisms of exploration of the new technologies, often associated with "Schumpeterian" entrepreneurship; (d) the selection criteria of markets and/or the technoeconomic requirements of early users (e.g., the technical specification of NASA and the Pentagon in the early days of integrated circuits, FDA requirements in the case of bioengineering, and the technical needs of the American navy in the case of nuclear reactors).

Once a technological paradigm becomes established, the objectives and heuristics of technological search often tend to diverge from those of scientific inquiry. This is partly due to the different ethos of the technological and scientific communities. (For example, the development of a first transistor had a deep scientific interest; a "better" transistor might have had a great interest for the engineer, but very little for the scientist); however, particular scientific activities (especially of an applied nature) often become part of the technological search along the "trajectories" defined by a particular paradigm. In other words, part of the scientific activity becomes "endogenized" within the activities of technological accumulation and search of profit-motivated firms (consider, for example, the applied scientific research of drug and chemical companies; for an analysis of the relationship between "endogenous" and "exogenous" research with regard to this case, see Chesnais 1986 and Orsenigo 1988).

All this has to do with the science-related opportunities of innovative activity; however, it has already been mentioned that, even in technologies that draw more directly on scientific advances, the knowledge base underlying innovative search also includes more specific forms of technical knowledge. A fortiori, this applies to technologies less directly dependent on science. These considerations have important implications for technological opportunities.

First, the specificity, cumulativeness, and tacitness of part of the technological knowledge imply that both the realized opportunities of innovation and the capabilities for pursuing them are to a good extent local and firm-specific. Second, the opportunity for technological advances in any one economic activity (and, thus, also the "innovative productivity"—were we able to measure it—of a dollar investment in R & D) can also be expected to be specific to and constrained by the characteristics of each technological paradigm and its degree of maturity. Moreover, the innovative opportunities in each economic sector will be influenced by the degree to which it can draw from the knowledge base and the technological advances of its suppliers and customers. The sectoral specificity of technological opportunities is also consistent with Scherer's findings that in econometric cross-firm, cross-industry estimates of rates of innovation—approximated by patenting—42.5 percent of the total variance must be attributed to the interindustry component: Frederick Scherer suggests that a good part of such variance is likely to relate to interindustry differences in opportunity (admittedly, despite the lack of any quantitative measure of it; Scherer 1986, ch. 9) other analyses confirm such sectoral specificities (e.g., Pakes and Mark Schankerman 1984).

In many respects, the idea that technological opportunities are paradigm-bound is also consistent with the historical evidence and interpretive conjectures put forward earlier by Simon Kuznets (1930) and Arthur Burns (1934) about a "secular retardation" in the growth of output and productivity, by commodity and industry, stemming—in the terminology suggested here—from the gradual
exhaustion of technological opportunities along particular trajectories.

New paradigms reshape the patterns of opportunities of technical progress in terms of both the scope of potential innovations and the ease with which they are achieved. Moreover, they generally spread their effect well beyond their sector of origin and provide new sources of opportunity, via input-output flows and technological complementaries, to otherwise stagnant activities. The emergence of new paradigms and the diffusion of their effects throughout the economy are possibly the main reasons why in modern economies we have not seen an approach to a “stationary state.” More precisely, one tends to observe two broad phenomena which reinforce each other. First, new technological paradigms have continuously brought forward new opportunities for product developments and productivity increases. Second, a rather uniform characteristic of the observed technological trajectories is their wide scope for mechanization, specialization, and division of labor within and among plants and industries (Nelson and Winter 1977). Contrary to the pessimistic expectations of classical economists and contrary also to many prevailing contemporary formalizations of problems of allocation of resources in decentralized markets, decreasing returns historically did not emerge even in those activities involving a given and “natural” factor such as agriculture or mining: Mechanization, chemical fertilizers and pesticides, new breeds of plants and animals and improved techniques of mineral extraction and purification prevented “scarcity” from becoming the dominant functional feature of these activities. A fortiori, this applies to manufacturing. Similarly, new technological paradigms, directly and indirectly—via their effects on “old” ones—generally prevent the establishment of decreasing returns in the search process for innovations. Think of the effects of biotechnology on the search efficiency for new drugs or the effects of electronics controls and computers on the innovative opportunities in machine building.

Contemporary studies of technological effort and progress, indeed support the conjecture that (a) at any point in time, technological opportunities vary according to the sectors and the degrees of development of the various paradigms under which they work, and (b) this is an important part of the explanation of why the commitment to innovative investment varies across sectors. (These hypotheses are supported on both empirical and theoretical grounds by Michael Gort and Richard Wall 1986.) Another—complementary—reason for interindustry differences in R & D investment relates to the different modes of innovative search that each paradigm entails. For example, in some technologies (e.g., electronics, organic chemicals, drugs, aerospace) innovation involves laboratory research and/or complex development and testing of prototypes. In other technologies (e.g., several kinds of nonelectrical machinery) innovation is much more “informal,” often embodied in incremental improvement in design, and as such neither recorded nor, often, perceived as the result of an “investment” in R & D.

As argued by Rosenberg (1976, pp. 277–79), differentiated scientific and technological opportunities determine different cost structures of technological advance (for example, the cost of a x percent improvement in the trade-offs implied by a particular technological trajectory). The cross-sectoral distribution of technological opportunities is far from homogeneous (Scherer 1982, republished in 1986; Pavitt 1984; Louise Dulude 1983). The appearance of new paradigms is unevenly distributed across sec-
tors and so are (a) the degrees of technical
difficulties in advancing production efficiency and product performance, and (b) the technological competence to innovate, embodied in people and firms. These distributions of opportunities and competence, in turn are not random, but depend on (a) the nature of the sectoral production activities, (b) their technological distance from the “revolutionary core” where new paradigms are originated, and (c) the knowledge base that underpins innovation in any one sector. As regards the effects of demand levels and changes upon sectoral rates of innovation (Schmookler’s “demand pull” hypothesis, whose discussion introduced this section), all the foregoing considerations need not conflict with the hypothesis that, other things being equal, market size and market growth may exert a positive influence on the propensity to innovate. However, the ceteris paribus clause is indeed a crucial one, since—it has been argued in this section—technological opportunities may vary widely across sectors and also over the history of individual technologies.

Given any one level of notional opportunities for innovation, the incentive to commit resources to their discovery and development will depend, of course, on the incentives that interest-motivated agents perceive in terms of expected economic returns. Let us consider the nature of these incentives.

B. Appropriability of Technological Innovations

As suggested by the classical and—even more so—Schumpeterian traditions, varying degrees of private appropriation of the benefits of innovation are both the incentive to and the outcome of the innovative process. To put it another way, each technology embodies a specific balance between public-good aspects and private (i.e., economically appropriable) features (see Arrow 1962b; Nelson 1984; and for an empirical analysis, Richard Levin et al. 1984 and Chesnais 1986). Call appropriability those properties of technological knowledge and technical artifacts, of markets, and of the legal environment that permit innovations and protect them, to varying degrees, as rent-yielding assets against competitors’ imitation.

Appropriability conditions differ among industries and among technologies: Levin et al. (1984) study the varying empirical significance as appropriability devices of (a) patents, (b) secrecy, (c) lead times, (d) costs and time required for duplication, (e) learning-curve effects, (f) superior sales and service efforts. To these one should add more obvious forms of appropriation of differential technical efficiency related to scale economies. In an extreme synthesis, Levin et al. (1984) find that for most industries, “lead times and learning curve advantages, combined with complementary marketing efforts, appear to be the principle mechanisms of appropriating returns for product innovations” (p. 33). Learning curves, secrecy and lead times are also the major appropriation mechanisms for process innovations. Patenting often appears to be a complementary mechanism which, however, does not seem to be the central one, with some exceptions (e.g., chemicals and pharmaceutical products). Moreover, by comparing the protection of processes and products, one tends to observe that lead times and learning curves are relatively more effective ways of protecting process innovations, while patents are a relatively better protection for product innovations. Finally, there appears to be quite significant interindustrial variance in the importance of the various ways of protecting innovations and in the overall degrees of appropriability: Some three-quarters of the industries surveyed by the study
reported the existence of at least one effective means of protecting process innovation, and more than 90 percent of the industries reported the same regarding product innovations (Levin et al. 1984, p. 20).\footnote{For detailed discussions of appropriability mechanisms, see also Christopher Taylor and Aubrey Silverston (1973), von Hippel (1978, 1980, 1982) and Terje Christian Buer (1982). The relative costs of innovation versus imitation—which is clearly a good proxy for appropriability—are studied by Levin et al. (1984) and Edwin Mansfield (1984). A detailed company-level study of patenting strategies is presented in Sally Wyatt and Gille Bertin (1985).}

Take, as an example, the case of microelectronics. Here one should distinguish between patterns of appropriability in the “core” technologies (semiconductors, computers, telecommunications, industrial controls) and in the technologies where it is applied (e.g., machine tools, consumer durables, cars). In the former, appropriability is a function of cumulative R & D (Franco Momigliano 1985); lead times; quite often, economies of scale in production (e.g., semiconductors and computers) and in R & D (minimum thresholds are sometimes very high, as in telecommunications); marketing and servicing networks (as in mainframe computers). Conversely, in the sectors where microelectronics is introduced as part of processes and products, the patterns of appropriability continue to correspond broadly to the “traditional” sector-specific features (see below for a more detailed taxonomy). An additional source of appropriability, however, relates to the capability of internalizing and/or efficiently exploiting the interfaces and synergies between microelectronics and applicative processes, for example, the capability of mastering both innovation in electronic equipment and the design of mechanical machinery. In fact, the latter is an example of a more general phenomenon, discussed by Teece (1986), whereby the control of complementary technologies becomes a rent-earning firm-specific asset.

In general, it must be noticed that the partly tacit nature of innovative knowledge and its characteristics of partial private appropriability makes imitation, as well as innovation, a creative process, which involves search, which is not wholly distinct from the search for “new” development, and which is economically expensive—sometimes even more expensive than the original innovation (for evidence on the cost of imitation relative to innovation, see Mansfield, Mark Schwartz, and Samuel Wagner 1981; Mansfield 1984; and Levin et al. 1984). This applies to both patented and non-patented innovations.

C. The Driving Forces of Technical Change

I have argued that opportunities—stemming partly from “exogenous” scientific advances and partly from the knowledge endogenously accumulated by the firms—and appropriability conditions account for the varying degrees of commitment of business enterprises to innovation. It is important to remark that what has just been said does not imply that market-determined inducement mechanisms are irrelevant to the propensity to search for new products and new techniques. The levels and changes in demand (market size and growth, income elasticities of the various products), and the levels and changes in relative prices, in particular the price of labor to the price of machines\footnote{On this point compare Paolo Sylos Labini (1984).} and also to the price of energy are influential factors. Indeed, they are likely to be fundamental ones, influencing (a) the rate and also the direction of technical progress, particularly within the boundaries defined by the nature of each technological paradigm, and (b) the selection of potential paradigms...
for exploration and thus for eventual appearance and dominance. My general point, however, is that the observed sectoral patterns of technical change are the result of the interplay between various sorts of market-inducements, on the one hand, and opportunity and appropriability combinations, on the other.

As an illustration of these points take, first, the case of automobiles. Throughout the seventies there was a clear inducement to produce energy-saving cars. Moreover, the more general demand conditions appear to be supportive (a very large market, although not growing very fast in advanced countries). Finally the appropriability conditions seemed favorable (relatively few producers with extensive distribution networks, marketing a complex product that is not so easy to copy). However, despite these favoring conditions, and leaving aside a significant change in the composition of output and demand (from big to smaller cars) progress in energy saving was rather modest: The technical opportunities on the internal combustion engine trajectory were certainly the major limiting factor. (Energy saving in U.S.-manufactured cars was, indeed, quite substantial, but this was due to the fact that American products were behind the “best-practice frontier” already reached by European and Japanese producers.)

Conversely, one can cite examples of a relatively low commitment to research and innovation, despite significant scientific and technological opportunities, due to the lack of satisfactory appropriability conditions. A case in point is part of agricultural research (until the advent of bioengineering) (Nelson 1986). There, the atomistic structure of production did not provide any incentive to research on seed variety, and so on, for individual farmers and lack of sufficient appropriability hindered industry-based research. Thus, most of the research in these fields has been publicly sponsored (e.g., in the USA, by the Department of Agriculture): the exceptions are hybrid sterile varieties, in addition, of course, to most industrial inputs to agriculture—pesticides, fertilizers, machinery—whose appropriability conditions have been broadly similar to the rest of manufacturing industry.

Finally, one can find examples of industries where both opportunities and notionally adequate appropriability conditions are there, but the firms generally lack the appropriate skills and technical competence to undertake research and innovation (to my knowledge, this is, for example, the case of Italian ceramic producers with respect to advanced ceramic materials or, more generally, of most firms in developing countries).

The conceptualization of technology and technical change based on “paradigms,” “guideposts” or whatever name is chosen, also helps in resolving the long debate in the innovation literature about the relative importance of “market pull” (cf., again, Schmookler 1966) versus “technology push” (for critical review, see Mowery and Rosenberg 1979). As known, in the former approach innovation is represented as a choice/allocation process on some sort of metaproduction function (the innovation possibility frontier) driven by market signals. In the latter, innovation drops from an exogenous domain (typically, it is a freely available by-product of scientific advances) and thus can be treated parametrically; however, the evidence from diverse technologies, such as aircraft (Constant 1980; Sahal 1981), agricultural technology and farm equipment (David 1975; Sahal 1981), synthetic chemicals (Freeman 1982), and semiconductors (Dosi 1984) is at odds with both accounts.

It is often the case that environment-related factors (such as demand and relative prices) are instrumental in shaping (a) the selection criteria among new po-
tential technological paradigms; (b) the rates of technical progress; (c) the precise trajectory of advance, within the set allowed by any given paradigm. However, it is useful to distinguish between what I call “normal” technical progress (i.e., those processes of innovation within the bounds of a given technological paradigm) and “extraordinary” technical progress (associated with the development of new paradigms). As regards the former, I suggest, unlike market-pull accounts, the set of possible trajectories is quite limited, bounded by the rules, technical imperatives, and specific scope of advance of each technology (Mowery and Rosenberg 1979)—which in the short term are to a good extent invariant to market conditions.

On a generally broader time horizon, market conditions exert a powerful influence on the conduct of technological search, but they do so primarily by stimulating, hindering, and focusing the search for new technological paradigms. When established, however, each paradigm—even when at the origin of its selection there were direct market stimuli—remains quite “sticky” in its basic technical imperatives, rules of search, and input combinations. For example, the number of ways of making polymers from fossil fuels is far from unlimited and so are the input intensities, irrespective of input prices. Even the substitution among different fuels (e.g., oil versus coal) often present major technical problems. Certainly, market changes may stimulate the search for new products and new “ways of doing things.” I suggest, however, that environmental factors are going to succeed in radically changing the directions and procedures of technical progress only if and when they are able to foster the emergence of new paradigms (for example, in the earlier example, new materials that substitute for plastics, bioengineering pro-

cesses to produce inputs that are alternative to fossil hydrocarbons).

Moreover, unlike both market-pull and “exogenous” accounts of technical progress, it appears misleading to consider innovation simply as a reactive process (to relative prices and demand, in one case, to new exogenous opportunities, in the other). On the contrary, technical progress is largely endogenously driven by a competitive process whereby firms continuously try to improve on their basic technologies and artifacts. Whether market signals change or not, firms try to perfect their products and processes, by trial-and-error mechanisms of search and imitations of the results already achieved by other firms, motivated by the competitive edge that innovations are expected to offer. Thus, according to this interpretation, each body of knowledge, expertise, selected physical and chemical principles, and so on (that is, each paradigm) constrains both the opportunities of technical progress and the boundaries within which “inducement effects” can be exerted by the market, while appropriability conditions motivate the economic agents to explore these technological opportunities as a rent-yielding competitive device. Finally, the evolution of the economic environment, in the longer term, is instrumental in the selection of new technological paradigms, and, thus, in the long-term selection of the fundamental directions and procedures of innovative search.

D. Inducement Factors, Patterns of Technical Change, and Irreversibility

Whatever the nature of the stimuli to change products and production processes exerted by an economic environment on microeconomic agents, “...they are naturally led to search the technological horizon... within the framework of [their] current activities and to
attack the most restrictive constraint . . . “ (Rosenberg 1976, p. 11). “Most mechanical productive processes throw off signals of a sort which are both compelling and fairly obvious; indeed, these processes when sufficiently complex and interdependent, involve an almost compulsive formulation of problems” (Rosenberg 1976, p. 11). The foregoing discussion on the general “paradigm-bound” nature of technical change allows the extension of Rosenberg’s thesis to most contemporary innovative processes and also reconciles it with those historical interpretations of different national/sectoral patterns of innovation that trace a cause of different rates of technical progress down to different environmental inducements, especially relative prices, availability or scarcity of natural resources (a locus classicus is the debate on the relative degrees of mechanization in the United States and England in the nineteenth century; see Erwin Rothbarth 1946; Hrthgar Habakkuk 1962; Peter Temin 1966; David 1975; Rosenberg 1976, especially chapters 3, 4, and 6).

As known, if one sticks to a general equilibrium framework and a representation of technology based on well-behaved production functions or convex production possibility sets, it is very difficult and often logically incoherent to attribute any observed bias in the rates and direction of technical change to particular biases in relative input prices (see David 1975 for a critical overview of a long debate). In the last instance, “economic incentives to reduce costs always exist in business operations, and precisely because such incentives are so diffuse and general, they do not explain very much in terms of the particular sequence and timing of innovative activity” (Rosenberg 1976, p. 110); however, specific incentives, coupled with the paradigm-bound, cumulative, and local nature of technological learning can explain particular rates and directions of technological advance (David 1975, 1986a and 1986b; Nelson and Winter 1982; Anthony Atkinson and Joseph Stiglitz 1969; W. Brian Arthur 1983, 1988).

To illustrate this point, consider the following story. Suppose that, once upon a time, when an imaginary technological history began, there were production possibility sets with all the right properties of continuity, convexity, and so on. Then, people started learning in a particular direction (to make it easy, suppose that this particular direction was triggered by an exogenous relative-price shock). With the help of some cumulativeness in technological knowledge and in search skills, local technological capabilities (that is, capabilities associated with the neighborhood of particular input combinations and output characteristics) developed more than proportionally to the “general” growth of knowledge on other notional portions of the production possibility set. Thus, other things being equal, technological progress became easier in this direction than in others. Then, with or without further shocks, people proceeded in this direction of search, which, in turn, further increased specific knowledge and skills. It is easy to see the moral of the story: One ends up with dynamic increasing returns along specific trajectories that channel also the response to particular environmental incentives to innovate. (A formal equivalent of this story is told in Arthur 1983, 1988).

A fundamental implication of this view is that, even when technical change is “triggered,” say, by relative price changes, the new techniques developed as a result are likely to be or become superior to the old ones irrespective of relative prices—immediately, as in the case of several microelectronics-based innovations (Soete and Dosi 1983), or after some learning time as in agricultural ma-
chinery (David 1975). In other words, if they had existed before, they would also have often been adopted at the “old” relative prices. That is to say, technical progress generally exhibits strong irreversibility features.  

Let us consider in greater detail the example of microelectronics. As discussed at greater length in Freeman and Soete (1985), Morniglano (1985), Soete and Dosi (1983), and Benjamin Coriat (1983, 1984), electronics-based production technologies are (a) labor-saving; (b) fixed-capital saving (i.e., they often induce a fall in the capital/output ratio; for sectoral evidence in the U.K., see Soete and Dosi 1983); (c) circulating-capital saving (i.e., the optimization of production flows allows a fall in the stocks of intermediate inputs per unit of output); (d) quality improving (i.e., they increase the accuracy of production processes, allow quality testing, etc.); (e) energy saving (in so far as the energy use generally is also a function of mechanical movements of the various machineries, the substitution of information-processing equipment for electromechanical parts reduces the use of energy). Taking all these characteristics together, it is clear that electronics-based production techniques are generally unequivocally superior to electromechanical ones irrespective of relative prices. That is, the new wage/profit frontiers associated with the new techniques do not generally intersect the “old” ones for any positive value (see Dosi, Pavitt, and Soete 1988). Remarkably, this example illustrates also the complex intersectoral linkages in the innovative process and their bearing on the “exogeneity versus endogeneity” issue in technical change. In the example of electronics technologies, unequivocally “superior” techniques and pieces of equipment appear, for several users’ sectors, as “dropping from an exogenous domain” (see Section IV C concerning “technology-push” accounts of technical change). In actual fact, they are generated through processes of exploration of technological opportunities endogenous to some other industrial sectors (in the example here, semiconductors, computers, industrial controls, etc.). Moreover, even in these cases, the full and efficient utilization of these potentially superior technologies (e.g., electronics-based automation as compared to electromechanical automation) relies on a painstaking process of learning on the side of the users, which is favored/hindered by the technological capabilities of the users themselves and the market conditions in which they operate. (It is an issue that relates also to the economics of innovation diffusion and that is impossible to discuss length in this work: For more historical evidence see Rosenberg 1975, 1982; a highly stylized attempt to model these learning processes is in Dosi, Orsenigo and Gerald Silverberg 1986.)

In other cases, the irreversibility properties of innovation emerge more slowly. At the start, the process of development/diffusion of new technologies may indeed involve choice-of-technique issues (see David 1975 on agricultural machinery). In the long term, the outcome of the rivalry between old and new technologies clearly depends also on the “latent opportunities” in the background, implicit in the two alternative paradigms; however, the degrees to which these opportunities are perceived, exploited, and expanded is likely to show path-dependent, cumulative, and irreversible features (for discussions and examples, see again David 1975, 1986b; Arthur 1983). Learning-by-

---

8 For microeconomic accounts of the local and irreversible features of technological learning, see David (1975, 1986b). On more general grounds, the study by Anne Carter (1970) on the technological coefficients of the American economy shows (a) the unequivocal superiority of the 1958 coefficients with respect to 1947 coefficients and (b) the dominance of a labor-saving trend on other variations of input coefficients. My informed guess is that this continues to remain true today.
doing and by-using, incremental improvements on the new technologies, and economies of scale in their production tend to improve their performance and lower their cost. Moreover, if adopted, a new product or process then attracts R & D efforts to itself, which, in turn, tend to improve costs and performance further. As a result, whenever the new technological trajectory is established, it is likely to dominate the old one (in the sense that, to repeat, it is economically superior, irrespective of relative prices).

Whatever the case, it is important to distinguish between the factors that induce, stimulate, or constrain technical change from the outcomes of the changes themselves. As analyzed in Dosi, Pavitt, and Soete (1988), drawing upon Rosenberg (1976), inducement mechanisms relate to a broad set of factors, including (a) technological bottlenecks in interrelated activities; (b) scarcities of critical inputs; or, conversely, (c) abundance of particular inputs (e.g., energy, raw materials); (d) major shocks in prices/supplies; (e) composition, changes, and rate of growth of demands; (f) levels and changes in relative prices (first of all, as mentioned, the relative price of machines to labor); (g) patterns of industrial conflict. Where the critical stimuli come from depends on the nature of the technologies and on the economic and institutional context of each country. One can find plenty of evidence on the role of each of these factors (for evidence and references on different technologies and countries, see Rosenberg 1976, 1982; Dosi, Pavitt, and Soete 1988; Ergas 1984). However, irrespective of the immediate triggering factor, I suggest that the patterns of innovation tend to follow rather irreversible "trajectories" defined by specific sets of knowledge and expertise. Moreover, irreversibility in the technological advances also means that, using neoclassical language, the changes of the production possibility sets are likely to dominate changes within any given set. More precisely, at any given time, instead of a well-behaved set we are likely to observe only one (or very few) points corresponding to the best-practice techniques, while, over time, the dominant process of change will imply improvements in these (very few) best-practice techniques, rather than processes of "static" interfactoral substitution. Admittedly, this interpretative conjecture is going to require more evidence and tests (which will not be easy) in order to corroborate its levels of empirical generality. (And there are also more subtle but normatively crucial issues: For example, in the historical evidence, how irreversible and local are learning processes? How powerful are "dynamic increasing return" phenomena? How can we measure, for normative purposes, the likely emergence of non-convexities, despite the obvious impossibility of making counterfactual historical experiments?) In any case, my assessment of the state of the art in innovation studies suggests, at the very least, that significant path dependencies, nonlinearities, and processes of specific, cumulative learning should be taken very seriously also at the level of the general, theoretical representations.

Finally, the irreversibility features of technical progress tend to be reinforced by the likely emergence of various sorts of externalities and specific infrastructures and institutions associated with the generation and/or exploitation of specific skills. I will now consider these latter aspects of innovation.

E. The Externalities of the Innovation Process

It has already been mentioned above that technology typically involves "public" aspects and "private" ones. The appropriability of the economic returns from innovation clearly relates to the lat-
ter. Conversely, the “public” aspects essentially take two forms.

First, there are certainly “free-good” elements, in technological progress, essentially stemming from the free flow of information, readily available publications, and so on. As mentioned earlier, economic theory tends to assume this to be the dominant feature of technology (save for institutionally granted rights to appropriation, such as patent rights). Of course, I do not suggest that models such as Arrow’s (Arrow 1962a) hinted at such a narrow equivalence of technology and information; however, it is fair to say that it provided some legitimacy for several contemporary formulations that have assumed it as least as a “workable hypothesis”. The foregoing survey of the characteristics of technology and technical progress implies a rejection of such a view as, at best, incomplete.

Moreover, the “public” characteristics of technology relates to the information flows and the untraded interdependencies among sectors, technologies, and firms and takes the form of technological complementarities, “synergies,” flows of stimuli, and constraints that do not entirely correspond to commodity flows. For example, knowledge and expertise about continuous chemical processing may allow technological innovations in food processing even when the latter do not involve any chemical inputs; “arms-length” relationships between producers and users of industrial equipment (such as informal exchanges of information, exchanges of technical specifications, and manpower mobility) are often a fundamental element in the innovation process even if sometimes no economic transaction is involved; at its origins the production of bicycles drew technological knowledge from the production of shotguns, even if obviously neither product is an output or an input in the other activity. All these phenomena represent a structured set of technological externalities that can sometimes be a collective asset of groups of firms/industries within countries or regions (see Bengt-Ake Lundvall 1984 and 1988) or else, tend to be internalized within individual companies (e.g., Teece 1982 and Pavitt 1986a). By a “structured set” of externalities, I mean some sort of consistent, and sometimes hierarchical, pattern linking different industries and technologies (such as different kinds of machinery production and users and producers of particular types of equipment). In other words, technological bottlenecks and opportunities (Rosenberg 1976) and experiences and skills embodied in people and organizations, capabilities, and “memories” overflowing from one economic activity to another tend to organize context conditions that (a) are country-specific, region-specific, or even company-specific and (b) as such, determine different incentives/stimuli/constraints to innovation, for any given set of strictly economic signals.

Relatedly, technological progress along any one trajectory is linked with (a) the development of specific infrastructures; (b) system scale economies; (c) complementary technologies; and (d) particular technical standards that positively feed upon specific patterns of innovation David (1986) and Hughes (1982), for example, discuss the interrelation between the development of the electricity-grid infrastructure and what, in the terminology suggested here, are electricity-based technological trajectories. Other obvious examples of “infrastructure technologies” that perform as an externality to a wide range of innovative activities are transport systems and telecommunications. Arthur (1983) and David (1985) illustrate the latter two points in the case of the development of the QWERTY keyboard in typewriters (QWERTY refers to the first top letters in the standard American
Although it was designed to meet problems which subsequent developments have overcome, and though it is no longer the optimal keyboard, the QWERTY standard has remained dominant as a result of cumulative development on an early lead: The specific skill of typists in QWERTY fostered QWERTY standards on the manufacturing side, which in turn increased the incentive to acquire QWERTY typing skills. Moreover, the interrelatedness between different technologies that compose a technological system or a complex product helps us understand why companies and countries may be "locked" into technologies—see the classic work by incumbent Marvin Frankel (1956) linking this point to development topics. Finally, Teece (1982, 1986) discusses a somewhat similar issue from the point of view of firms’ structures and strategic management, identifying the crucial role of internalized complementary technologies in firms’ competitive performances.

Untraded interdependencies and context conditions are, to different degrees, the unintentional outcome of decentralized, but irreversible, processes of environmental organization (one obvious example is Silicon Valley) and/or the result of explicit strategies of public and private institutions.

The evolution over time and the spatial differences in these untraded interdependencies also represent an important link between innovation studies and the regional economics of technical change (see Edward Malecki 1983 and Morgan Thomas 1985). Whenever these technological externalities—in the form of specific infrastructures, skill availability, competences embodied in local firms, easier information about new production inputs—reproduce through time as a sort of dynamic increasing returns (Arthur 1986), they also help explain the differentiation in the technological capabilities, rates of innovation, and rates of diffusion among regions and countries (see Alfred Thwaites and Ray Oakey 1985). In this field, an original tradition of studies has developed particularly in France: the analysis of "filieres" (literally "webs"), linking groups of industries and technologies via input-output flows and technological complementarities, is a promising way of relating the microeconomic process of innovation with the evolution of the wider economic environments (on "filieres," see Joel Toledano 1978; Alexis Jacquemin and Michael Rainelli 1984; Richard Arena, Michael Rainelli, and Andre Torre 1984; Ehud Zuscovitch 1984; Patrick Cohendet, Regis Larue de Tournemine, and Zuscovitch 1982; Jean-Louis Truel 1980).9

F. Determinants and Patterns of Investment in Innovation: Toward a Sectoral Taxonomy

Let me summarize the discussion of the foregoing five subsections. As analyzed in detail by Nelson (1986, 1988), the process of innovation in Western economies embodies complex and varying balances between public and proprietary forms of knowledge, and different combinations between notional opportunities of innovation, firm-based capabilities to reap these opportunities, and economic incentives to do so (related to appropriability mechanisms, market conditions, relative prices, broader socio-economic conditions such as industrial relations). Moreover, the specific opportunities that are seized, the appropriability mechanisms that are developed, and the actual capabilities that are used tend to grow with each other. Phenomena of hysteresis are likely to emerge:

9 The concept of "filieres" partly overlaps with Hirschman’s insights on "backward" and "forward" linkages (Albert Hirschman 1958).
The exploration of particular technologies and the development of particular problem-solving methods increase the capabilities of firms and industries in these specific directions and thus increase the incentive to do so also in the future. These technology-specific forms of dynamic increasing returns tend to “lock in” the processes of technical change into particular trajectories, entailing a mutual reinforcement (a positive feedback) between a certain pattern of learning and a pattern of allocation of resources into innovative activities where learning has already occurred in the past (for general discussions of these path-dependent processes, see David 1975, 1986a; and Arthur 1983, 1988).

In line with Nelson (1986), I suggest that the different combinations among these factors explain the “rich and variegated institutional structures supporting technical advance that have grown up in capitalist countries” (p. 1). They also constitute the rather complicated constellation of factors by which a significant group of contemporary economists seeks to explain the pace and characteristics of technological progress, and its international, interindustry, and intertemporal changes. Certainly, while this outlook appears to be the most promising approach we have, and while it appears to be consistent with some blocks of experience, far more empirical and historical work will be needed to establish its validity and the manner in which the different elements of the new outlook operate. A first step is to generalize on some common empirical characteristics of technologies and sectors (which I do below) and, then, try tentatively to “map” these characteristics into the features of technologies and innovative processes discussed so far (see Section VI).

Scherer, as mentioned, has recently developed an intersectoral matrix of the origin and use of R & D in the U.S. economy, based on the intersectoral generation and use of a large sample of patents (Scherer 1982). The sectoral ratios of (direct plus indirect) use of R & D to performed R & D are shown in the last column of Table 3. On the grounds of a data base on innovations in the U.K. from 1945 to 1979 collected at the Science Policy Research Unit of the University of Sussex, Pavitt (1984) has developed a sectoral taxonomy of sectors of production/use of innovation. The two data sets appear to be in many ways complementary and provide interesting insights into the “anatomy” of contemporary economic systems and their major inner loci of innovation generation (the Yale questionnaire—partly discussed in Levin et al. (1984) and summarized in Nelson (1986)—adds further, and broadly consistent, evidence). Pavitt (1984) identified four major groups of manufacturing industries, namely:

1. “Supplier-dominated” Sectors. Innovations are mainly process innovations, embodied in capital equipment and intermediate inputs and originated by firms whose principal activity is outside these sectors proper. Supplier-dominated industries include agriculture, textiles, clothing, leather, printing and publishing, wood products, and the simplest metal products. In these sectors the process of innovation is primarily a process of diffusion of best-practice capital goods and of innovative intermediate inputs (such as synthetic fibers) while endogenously generated opportunities are rather limited and so are R & D expenditures. The knowledge base of these technologies tends to relate to incremental improvements in the equipment produced elsewhere and/or to its efficient use, and to organizational innovations. Cumulativeness and appropriability of technological capabilities are relatively restricted and firms are typically not very big (with some exceptions in those activ
ties characterized by some significant economies of scale in production, such as part of textiles, or in marketing and distribution networks, such as in clothing).

2. "Specialized Suppliers." Innovative activities relate primarily to product innovations that enter most other sectors as capital inputs. Firms tend to be relatively small, operate in close contact with their users and embody a specialized and partly tacit knowledge in design and equipment building. Typically, this group includes mechanical and instruments engineering. Opportunities for innovation are generally abundant, but are often exploited through "informal" activities of design improvement (thus, formal R & D is often rather low). Idiosyncratic and cumulative skills make for a relatively high appropriability of innovation (think of the secular advantage of German machine-tool makers).

3. "Scale-intensive" Sectors. Innovation relates to both processes and products, and production activities generally involve mastering complex systems (and, often, manufacturing complex products); economies of scale of various sorts (in production and/or design, R & D, distribution networks) are significant; firms tend to be big, produce a relatively high proportion of their own process technologies, often devote a relatively high proportion of resources to innovation, and tend to integrate vertically into manufacturing their own equipment. This group includes transport equipment, several electric consumer durables, metal manufacturing, food products, glass, and cement.

4. "Science-based" Sectors. Innovation is directly linked to new technological paradigms made possible by scientific advances; technological opportunity is very high; innovative activities are formalized in R & D laboratories; investments in innovative search are quite high; a high proportion of their product innovation enters a wide number of sectors as capital or intermediate inputs; firms tend to be big (with the exception of new "Schumpetarian" ventures, and highly specialized producers). This group includes the electronics industries, most of the organic chemical industries, drugs, and bioengineering. (Aerospace and some military-related activities share with science-based sectors the importance of inputs from scientific advances and of formalized research, while sharing with the production-intensive sectors the importance of economies of scale and of an efficient organization of complex production systems.)

Taxonomic exercises on the intersectoral differences in the sources, procedures, and intensity of innovative search are rather new and a lot of comparative work is still to be done; however, let me briefly mention the importance of these analyses on both positive and normative grounds. As regards the former, the fact that innovations are located at different places within the "capitalist engine" (Nelson 1986, p. 20) demands a better understanding of the factors that tend to concentrate innovative opportunities and investments in some activities more than in others. Taxonomic efforts help in this understanding and also in answering puzzling comparative questions such as, How did Germany and Sweden become so good in mechanical engineering? How does this relate to their productive structure and international competitiveness? Why is the United States relatively strong in science-based industries? (some tentative answers to these comparative questions are attempted in Pavitt 1988; Pavitt, Dosi, and Soete 1988; Ergas 1984). On normative grounds, a more detailed understanding of the intersectoral patterns of innovation directs attention to questions of importance for industrial and development policies. Given the
objective of an acceleration of the rate of technical progress, would R & D incentives be well suited for science-based industries but not for "supplier-dominated" industries? Would the development of a large internal market be important for scale-intensive industries but not so much for "specialized suppliers"?

G. Some Conclusions

In this section, I have focused on the broad differences in opportunities, incentives, R & D investments, and innovative procedures among industries. The thrust of the argument has been that these differences exist, are important, and help explain the internal structure of the complex engine which in modern noncentrally planned economies continuously generates new products and production processes. Moreover, intertechnological differences in opportunity, appropriability conditions, knowledge bases, and search procedures help explain what Nelson (1986) calls the "problem of institutional assignment," that is, the location within the socioeconomic system of specific search and development activities to particular actors, for example, why certain activities are undertaken by nonprofit institutions and others by business firms, why some sectors produce their own process innovations and others buy them from the market (more on the latter in Williamson 1985; Buer 1982), and why some economic activities contribute a disproportionate share of innovations while others are mainly recipients. Relatedly, the input-output structure of the economy, together with information flows and intersectoral flows of knowledge embodied in people and organizations, diffuses through the system the economic effects of particular innovations, thus amplifying the opportunities for productivity growth and new product development.

There is a yet finer level of analysis, however. After all, industry-specific characteristics are averages of cross-firm distributions. The fact that these averages show recognizable patterns that are relatively stable over time and across countries entails the relative stability of the industry-specific and technology-specific factors analyzed above. Still, one must explain also the intra-industry variance in innovative investments and degrees of innovative success. Moreover, innovation and imitation continuously modify the relative performance and competitiveness of firms and thus affect also the dynamics of industrial structures. Parts V and VI will discuss these topics.

V. Intrasectoral Differences in Innovativeness and Economic Performance

One of the most common features of industrial case studies is the description of significant differences among firms not only in terms of size, but also in terms of technological capabilities, product-market strategies, degrees of innovativeness and competitive success, costs of productions, and profitability. Putting it another way, nothing similar to the "representative firm" stylized by economic theory seems to emerge from the empirical accounts (to see this, consult a random sample of articles in, say, the Harvard Business Review and the California Management Review, or, for more detail of industrial histories, Alan Altshuler et al. 1984 on automobiles, Dosi 1984 and Franco Malerba 1985 on semiconductors, and Enos 1962 on petroleum refining). With reference to the foregoing discussion, one is led to ask what are the relationships between the characteristics of innovation analyzed earlier, on the one hand, and the intrasectoral differences in firm structures (e.g., in size) and per-
formance (e.g., rates of innovation and production costs), on the other. In this section I focus mainly on the general features of interfirm, intrasectoral differences in innovativeness, and, more generally, economic performance, leaving to Section VI a more detailed account of the processes that generate them. The empirical reference upon which this section is based is rather commonsensical, for example, the fact that firms can be generally found to be widely different, in terms of various performance indicators and also of behaviors, structures, and strategies. However, these simple ideas might be usefully conceptualized in the sense that meaningful classifications might help, first, in providing empirically sound hypotheses for theoretical modeling and, second, in moderating the innocent acceptance—widespread in the economic literature—of “representative firms,” “equilibrium conditions of production,” “technological identity of producers,” and so on. Thus, in the following, I attempt a classification of the factors that account for intrasectoral differences in both structures (e.g., size) and performances (e.g., degrees of innovativeness).

I start from the intrasectoral, interfirm differences in innovative investment as shown by their R & D expenditures.

A. Interfirm Differences in R & D

A long debate has taken place in industrial economics on the relationship between size of firm and innovation (both R & D investments and innovative output, typically patents). I shall not enter the details of the discussion that concern both the meaning of particular measures (it has been argued that patenting underestimates the innovation output of big firms—which appear to show a lower propensity to patent; the R & D expenditures are likely to underestimate innovative contributions of small firms—which sometimes innovate on an “informal” basis, etc.) and the degrees of empirical corroboration of the improperly termed Schumpeterian hypothesis (i.e., that bigness is relatively more conducive to innovation, that concentration and market power affect the propensity to innovate). For reviews and (partly conflicting) findings, see Scherer (1988), Soete (1979), Griliches (1984a), Griliches and Pakes (1986), Morton Kamien and Nancy Schwartz (1982), Wesley Cohen and Levin (1988), and Pavitt, Robson, and Townsend (1987). For the purposes of the present work, it is enough to mention three major regularities that come out of empirical studies.

First, there appears to be a roughly log-linear relation within industries between firm size and R & D expenditures (or patenting). This is, however, a rather crude approximation. On closer inspection, subject to industry differences and different measures of innovativeness, one finds better fits of quadratic and cubic relationships between size (i.e., sales or employment) and innovativeness (R & D expenditure, R & D employment, number of patents, or number of innovations); however, irrespective of the form of the econometric model, the estimates show roughly nondecreasing returns of innovative proxies to firm size (Scherer 1986 argues in favor of a lower degree of innovativeness for the greatest size classes; Soete 1979, using partly different data, shows the opposite for about a third of his industry sample).

Second, the size distribution of innovating firms within sectors depends on the technological characteristics of the sectors themselves. Pavitt, Robson, and Townsend (1987), using the SPRU sample of innovation mentioned earlier, conclude that in sectors with high technological opportunities (chemicals, electrical/electronics) the innovating firms “can be found heavily represented among those
that are very large and those that are small” (p. 16). Conversely, in machinery and mechanical engineering (approximately those “specialized suppliers” with the characteristics identified earlier) a relatively greater proportion of innovation is undertaken by small firms (which, however, are “small” in comparison with the size of distribution of the manufacturing firms’ universe, but not necessarily in relation to the specific national or world market in which they operate).

Third, irrespective of the statistical proxy for innovativeness (and in particular, irrespective of the choice between an investment measure or an output measure), after allowing for the effect of firm size, one still generally observes a substantial unexplained interfirm, intrasectoral variance, in terms of both R & D investments and, even more so, innovative output. (Moreover, note that a significant proportion of firms in each industry do not patent and do not produce significant innovations; for evidence, see John Bound et al. 1984.)

There are three obvious caveats for the interpretation of these results. The first relates to the fact that the statistical proxies for innovativeness cannot capture those aspects of technical change, discussed above, based on “informal” learning (thus, independent of measured R & D investments) and/or yielding incremental innovations (hence, unrecorded in patents or discrete innovation counts). The second is that some (generally unidentified) part of the intrasectoral variance in innovative performance must be attributed to differences in the actual lines of business (and, thus, in opportunity, appropriability) which are, nevertheless, statistically classified within the “same” industry. Third, some firms may not patent or innovate but still engage in substantial R & D which is simply devoted to keeping up and adapting to what other competitors are doing.

Despite such limitations, however, these empirical regularities tell a story that, in my view, is consistent with the characteristics of the innovative process discussed earlier. More precisely, the intra- and intersectoral differences in the size distribution of firms in general and in particular of innovating firms are linked with the characteristics of different technological paradigms and the ways innovative capabilities develop and can be competitively exploited by individual firms. After all, any particular distribution of firms’ characteristics (e.g., size, R & D propensities, unit costs) at any one time is itself the result of processes of corporate learning and market competition whereby certain corporate features turned out to yield a competitive advantage. The general interpretations suggested here are that (a) the sectoral distributions of characteristics such as firms’ sizes are affected by the specific characteristics of the technological paradigms on which the production of that sector is based, in terms of appropriability, technological opportunities, scope for automation, and economies of scale; however, (b) any observed bias in the size distribution of firms in a particular sector is not sufficient evidence to make inferences on the “true” effect of size upon innovativeness. For example, an industry may show a relative bias toward “bigness” even if the latter does not confer particular innovative advantages (or disadvantages); it may be due simply to technical requirements on the production side (such as economies of scale in production and marketing). Alternatively, size may actually be conducive to innovation (because of indivisibilities of R & D projects, high R & D minimum thresholds) or detrimental to it (e.g., if it induces organizational inflexibilities). Finally, there may be cases in which the correlation between size and innovativeness reflects a causal process in the re-
verse direction: Big firms happened to grow big because they innovated successfully in the past and continue to do so in the present without, however, finding a differential advantage in "bigness" as such. In general, the relationship between industrial structures and degrees of innovativeness runs both ways and the understanding of particular intrasectoral distributions of firms' structural and performance characteristics implies the understanding of the (technology-specific) effects of innovation on firms' economic performance and competitiveness. Some of these effects obviously relate to the scope for economies of scale that each technological paradigm entails. Others relate to the impact that the differential innovative capabilities of certain firms exert on their ability to acquire a lead in efficiency and/or product quality vis-à-vis other firms. Let me start with the former.

B. Flexibility and Economies of Scale

Most technological trajectories since the Industrial Revolution involved increasing mechanization of production and increasing exploitation of economies of scale (see Nelson and Winter 1977 and the works cited therein); however, each technological paradigm is characterized by different trade-offs between flexibility (with respect to production runs and variety of outputs, for a given equipment) and economies of scale. Thus, a first determinant of any observed sectoral distribution of firms (and/or plants), by size, relates to the degrees to which individual firms have explored and possibly improved along a particular technological trajectory. Take the contemporary example of the transition from electromechanical patterns of automation to electronics-based ones. As compared to "classical" (electromechanical) automation of mass production, numerically controlled machine tools, flexible manufacturing systems, and robots allow a much greater flexibility of production in terms of (a) acceptable variance of throughputs (defined in terms of number of cost-effectively produced homogeneous items per unit of time), (b) acceptable variances in output varieties, and (c) minimum scale of production (see Coriat 1983, 1984; Michael Piore and Charles Sabel 1984).

This has two consequences. First, it increases the efficiency of small-scale productions. Second, it increases the importance of plant-related economies of scale that were one of the main sources of both productivity growth and production rigidity in "classical" Fordist automation.

Within the electromechanical paradigm, higher efficiency of production (stemming from standardization, economies of scale, etc.), generally associated with "Taylorist" and "Fordist" principles of organization and production, is also correlated with very high degrees of inflexibility—in terms of acceptable variance in production runs and mixes. Figure 1 illustrates such a case. Suppose that, in the "old" technology, the line AA represents the technical relationship between "normal" average total unit costs (c) and rates of throughput (q), while the line FF represents the corresponding relationship between unit costs and degrees of flexibility (F), say, approximated by the standard deviation in the rate and mixes of throughput that does not significantly increase "normal" unit costs. Fundamental dimensions of technical progress along the old technological trajectory are the increasing exploitation of economies of scale and economies of standardization. Thus, any increase of the flexibility requirements (due, for example, to increasing uncertainty about the levels and composition of demand) indirectly represents a retardation factor of technological innovation/diffusion within the electromechanical paradigm, insofar as technical advances are also
scale biased; however, different technological paradigms embody different trade-offs between flexibility and scale. Suppose for example that, in Figure 1, the line $A^*A^*$ represents the relationship costs/quantities for a new electronics-based paradigm, while the line $F^*F^*$ is the corresponding relationship flexibility costs. Thus, the trade-off quantity/flexibility is $TT$ for the old technology and $T^*T^*$ for the new one.\(^{10}\) Now, consider again an increase in the desired flexibility. Remarkably, this is likely to have two effects: First, it is likely to hinder “normal” technical progress/diffusion along the “old” technological trajectory while, second, fostering innovation/diffusion in the new technological paradigm. Suppose we start from production runs equal to $q_0$, normal total costs at $c_0$ and a degree of flexibility $f_0$. Now, say, an increasing instability of economic growth, an increasing uncertainty about consumers’ demand and rivals’ strategies increases the required flexibility of production from $f_0$ to $f_1$. On the grounds of the “old” technological paradigm, this would mean very short production runs ($q_1$) and very high costs ($c_1$). The new paradigm (e.g., electronics-based automation) changes the nature of the trade-offs, allowing, for example, the required flexibility to be achieved at throughput $q_2$ and unit costs $c_2$. Moreover, I suggest, the higher technological opportunities of the new paradigm (with its scope for learning, decreasing costs of capital equipment, etc.) in the long term will shift the technoeconomic relation between costs and quantities, say, down to $A^{**}A^{**}$.

In contemporary economies one often observes a fall in plant sizes (see, for example, Fabrizio Barca 1984 on Italy),

\(^{10}\) This example owes a lot to the discussions with B. Coriat on automation in general and, in particular, on the car industry.
somewhat analogous to the change in the scale of production from $q_0$ to $q_2$ in Figure 1; however, this empirical observation per se does not allow any conclusion on either the "optimal" or "equilibrium" relationship between characteristics of the technology and size, or on the long-term trends in technical opportunities for economies of scale. As illustrated above, the actual change in production scale is the joint outcome of (a) differences in the scope for economies of scale of the two technological paradigms, (b) (sector-specific) differences in the trade-offs between flexibility and scale economies that they imply, and (c) different degrees of technical progress along the two trajectories defined by the two paradigms.

One can see here a first link going from the characteristics of each technology to industrial structure (and its changes). The observed intra-industry distributions of firms by size are obviously affected by the sector-specific opportunities for various kinds of economies of scale and the trade-offs between the latter and production flexibility. If different firms position themselves differently on the notional trade-offs between flexibility and scale economies and/or explore and exploit the opportunities of automation at different paces, one should observe a distribution of varying plant sizes and firm sizes even when the propensity to innovate is neutral with respect to size. As an historical example, I suggest we are currently observing, at least in the industrialized countries, a process of change in the size distribution of plants and firms that is significantly influenced by (a) the new flexibility-scale trade-offs associated with electronic production technologies, and (b) the painstaking attempts to learn how to use them efficiently and slowly explore the (still largely unknown) potential for economies of scale that they entail (for some highly preliminary and impressionistic evidence, see Mehmet Gonenc 1984; Fabio Arcangeli, Dosi and Moggi 1986 and Giancarlo Caimarca, Massimo Colombo, and Sergio Mariotti 1987); and, possibly (c) an increasing variety of demanded characteristics of products, greater refinement and tolerance for (or desire for) novelty, associated with market segmentation and higher consumers' income. As general outcome of all these factors, in line with Pavitt (1986b), I conjecture that changes in the size distribution (by plant and especially by firm) in the "specialized supplier" industries—machine making, etc.—will tend to be biased toward the bigger size classes, because of R & D indivisibilities, economies of scope based on electronics flexible manufacturing systems, and so on. Conversely, in mass-production industries the higher flexibility of the new forms of automation is likely to allow the efficient survival also of relatively smaller firms (as compared to the past).

More generally, on the grounds of sectoral, and still scattered, evidence, it is plausible to conjecture that, at any one time, there may be a certain number of (technology-specific) size distributions (by plant and by firm), that represent, so to speak, notional "evolutionary equilibria," in the sense that a variety of firms and plants coexist at roughly similar levels of economic performance, by exploiting more economies of scale with less flexibility, more economies of scope and lesser economies of scale, and so on.

This is not the only source of difference among firms. Other differentiating mechanisms relate even more directly to innovation and innovative strategies.

C. Innovation, Variety, and Asymmetries Among Firms

A major implication of the characteristics of cumulativeness, tacitness, and partial appropriability of innovation is the permanent existence of asymmetries among firms, in terms of their process technologies and quality of output. That is, firms can be ranked as "better" or
“worse” according to their distance from the technological frontier. As discussed in another work (Dosi, Pavitt, and Soete 1988), one can see here an interesting convergence between the findings from international trade analyses suggesting widespread technology gaps between countries (see Freeman 1963; Freeman, C. J. Harlow, and J. K. Fuller 1965; Gary C. Hufbauer 1966; Dosi and Soete 1983; OECD 1968; Soete 1981; and Mario Cimoli 1988) and the evidence from industrial economics within each country on similarly wide gaps in technology among firms as measured by their costs of production (see, for example, Tsung-Yuen Shen 1968; Nelson 1968, 1981a; Bela Gold 1969; and Dosi 1984. This confirms earlier findings by the U.S. Bureau of Labor Statistics, cited in Nelson 1981a). Moreover, some recent studies in industrial economics have begun to explore the existence and intertemporal persistence of above- or below-average profitabilities of individual firms (see Paul Geroski and Alexis Jacquemin 1986; Bruno Contini 1986; Dennis Mueller 1977; and Yiyoyoki Odogiri and Hideki Yamawaki 1986); incidentally, note that interfirm profitability differences are likely to underestimate the differences in production efficiency and product technology insofar as their “quasi-rents” are distributed as differential wages and salaries.

Call degrees of asymmetry of an industry its dispersions of (a) input efficiencies for a given (homogeneous) output and (b) price-weighted performance characteristics of firms’ (differentiated) outputs—were we able to measure them with precision. Certainly, part of these interfirm asymmetries in production efficiency are due to (a) economies of scale in production (see Cliff Pratten 1971; Aubrey Silberston 1972; and Donald Hay and Derek Morris 1979), and (b) different vintage distributions of the equipment of each firm (Wilfred Salter 1969); however—perhaps more important—these asymmetries are also the effect of different innovative capabilities, that is, different degrees of technological accumulation and different efficiencies in the innovative search processes. Other things being equal, one would expect that the higher the potential that a technological paradigm entails for creating asymmetries in product quality and production efficiency (that is, the higher are, jointly, technological opportunities and appropriability of innovative advantages), the higher the scope for the “best” firms to enjoy a competitive advantage and grow bigger, irrespective of any possible bias in the “returns” of innovativeness to size (I come back to this issue in Section VI). Of course, any observed pattern of asymmetry among firms depends also on many other features of the markets in which firms operate. For example, varying degrees of demand elasticity affect the degrees of protection that any firm enjoys against the greater efficiency of rivals, or conversely, the ease with which technological leaders can grow at the expense of less efficient rivals. In fact, there is here an obvious complementarity between the findings and conceptualizations emerging from innovation studies, on the one hand, and the analyses of entry and mobility barriers in industrial economics, on the other (see, for example, Richard Caves and Michael Porter 1977 and 1978; Scherer 1980; Sylos Labini 1967).

If such asymmetries are a factor of diversity among firms that correspond, in a loose biological analogy, to different degrees of “fitness,” there is yet another source of diversity which, in the same analogy, corresponds to roughly equal fitness and “polymorphism.” Take, for example, two firms that show identical unit costs and produce the same good. Thus, they do not show any asymmetry, in the sense defined above; however, they
might still show differences in their input combinations, which are the particular result of firm-specific histories of technological accumulation (Nelson and Winter 1982; Nelson 1985; Metcalfe 1985; Gibbons and Metcalfe 1986; Dosi, Orsenigo, and Silverberg 1986). Similarly, firms might well search for their product innovations in different product spaces, embodying different characteristics and aimed at different corners of the markets. Call this second set of sources of diversity technological variety, to mean all those technological differences that do not correspond to unequivocal hierarchies (i.e., “better” and “worse” technologies and products).

Finally, empirical studies often show the coexistence, within the same industry and for identical environmental incentives, of widely different strategies related to innovation, pricing, R & D, investment, and so on. Specifically with regard to innovation one notices a range of strategies concerning whether or not to undertake R & D; being an inventor or an early imitator, or “wait and see”; the amount of investment in R & D; the choice between “incremental” and risky projects, and so on (see Charles Carter and Bruce Williams 1957; Freeman 1982 and the bibliography cited therein). Call these differences behavioral diversity.

I suggest that technological asymmetries, varieties, and behavioral diversities manifest themselves also in the “unexplained” variances in R & D, patenting, and a number of discrete innovations cited earlier (Section V A).

To summarize: Each production activity is characterized by a particular distribution of firms according to their R & D investments, innovative output, size, degrees of asymmetries in product quality, and production efficiency. However, the picture of an industry that emerges at any time is itself the result of a competitive process which selected survivors within the technological variety and behavioral diversity of firms, put a premium or a penalty on early innovators and allowed varying degrees of technological imitation and diffusion. Thus, a satisfactory understanding of the relationship between innovation and distribution of firms’ structural and performance characteristics also implies an analysis of the learning and competitive process through which an industry changes. I turn now to these topics.

VI. Innovation and Industrial Change: Learning and Selection

A. The Innovative Process and Industrial Structures

Over time, as innovation proceeds, new products are introduced and later imitated by other firms, better methods of production are developed or adopted in the form of new types of capital equipment, and, relatedly, some firms are able to obtain below-average costs of production and/or a monopolistic/oligopolistic position, in the manufacturing of some new products. In turn, they can exploit these differential advantages by increasing their rates of profit, their market shares, or, of course, a combination of the two. Conversely, some firms find themselves with above-average costs and/or lower-quality products and, through various strategies of imitation, search, and attempts to “leapfrog,” must try to catch up in order to improve their profitability and market competitiveness. One version or another of this basic process is what determines the sectoral “snapshots” discussed in the previous section and is revealed, over time, also by the changes in the averages and distributions of firms’ inputs, productivity, unit costs, product performances, profit rates, and sizes. Putting it another way, industrial performance and industrial structures are endogenous to the process
of innovation, imitation, and competition.

Nelson and Winter (1982), Winter (1971), Katsushito Iwai (1981), Gunnar Eliasson (1986a), Gerald Silverberg (1987), Dosi, Orsenigo, and Silverberg (1986), Gibbons and Metcalfe (1986) and Öve Granstrand (1986) have tried to formalize this process in an evolutionary perspective: “Market structure and technological performance are endogenously generated by three underlying sets of determinants: the structure of demand, the nature and strength of opportunities for technological advance and the ability of firms to appropriate the returns from private investment in research and development” (Levin et al. 1984, p. 1). (Treatments of the endogeneity of market structures have been recently developed also within an “equilibrium” framework; see Dasgupta and Stiglitz 1980a, 1980b.) Case studies of individual industries confirm both the endogenous nature of market structures and the causal link going from technological success to changes in firm size and degree of industrial concentration; in addition to Gort and Steven Klepper (1982) and Gort and Akira Konakayama (1982) who provide comparative intertechnological evidence, see for example, some cross-sectoral evidence in Levin, Cohen and Mowery (1985); and the more qualitative sectoral evidence in Almarin Phillips (1971) on aircraft; Barbara G. Katz and Phillips (1982) on data processing; Wilson, Ashton, and Egan (1980), Dosi (1984), John Tilton (1971), Ed Sciberras (1977), and Malerba (1985) on semiconductors; Altschuler et al. (1984) on automobiles; Chesnais (1986) on drugs and bioengineering; and Momigliano (1983) for a cross-company international econometric analysis of the relationship between levels and changes in various indicators of innovativeness and changes in performance of computer firms.

Broadly speaking, the growing (but still largely inadequate) evidence on the dynamics of industries and technologies highlights complex and varied learning processes whereby firms explore specific domains of perceived technological opportunity, improve their search procedures, and refine their skills in developing and manufacturing new products, drawing partly on their internal accumulated knowledge, partly on artifacts and knowledge developed elsewhere, and partly by copying their competitors. In turn, market interactions select, to different degrees, particular directions of technological development, allowing some firms to grow bigger and penalize others. Note also that in this dynamics, technological asymmetries and technological and behavioral variety are both the outcome and a driving force of technological and organizational change. That they are the outcome of innovation is straightforward from the earlier discussion: Firms generally learn at different rates, and with modes and behavioral rules specific to their history, internal organization, and institutional context. These interfirm differences are also a major driving force of the process of change in that they underlie the competitive incentive (for the “winners”) and the competitive threat (for the “losers”) to innovate/imitate products, processes, and organizational arrangements.

Each observed industrial history is, in an essential sense, the outcome of a particular form of this general process; however, in order to account for the specific differences in the patterns shown by individual industries, one should move one step further and, so to speak, “map” the varied characteristics of innovation, as discussed in Sections II to IV, into empirically recognizable classes of evolutionary processes. So, for example, one should be able to link the characteristics of opportunity, appropriability, and so on, of
each technological paradigm and the patterns of change in, for example, firms' sizes, market concentration, and degrees of asymmetries. Here the evidence is still highly unsatisfactory and some conjectures can be related only to single case studies and to the plausibility of simulation results; however, the issue is worth pursuing for its analytical (and also normative) relevance.

B. Characteristics of Innovation and Patterns of Industrial Change

In general, the observed changes in industrial structures and the observed dynamics of industrial performance (e.g., rates of introduction of new products and rates of change in sectoral productivities) are the outcome of (a) innovative learning by single firms (together with that contributed by universities, government agencies, and so on); (b) diffusion of innovative knowledge and innovative products and processes, and (c) selection among firms. Relatedly, my general interpretative conjectures are the following. First, the empirical variety in the patterns of industrial change is explained by different combinations of selection, learning, and diffusion and different learning mechanisms (e.g., "informal" learning by doing, learning through formalized R & D, and experience in marketing). Second, the nature of each technological paradigm, with its innovative opportunities, appropriability conditions, and so on (jointly with other economic and institutional factors) helps explain the observed intersectoral differences in the relative importance of the three processes. (Some further discussion of these conjectures can be found in Dosi, Orsenigo, and Silverberg 1986 and Dosi, Winter, and Teece 1987.) I will proceed by making some broad remarks on the nature of the three processes and then highlight the ways in which they are affected by the characteristics of innovation, by means of some "ideal examples," simulation results, and case studies.

To begin, note that each successful innovation—whether related to process technology, products, or organizational arrangements—entails ceteris paribus, an asymmetry-creating effect, which allows one or some firm(s) to enjoy some improvement in its competitive position (e.g., lower prices or better products). Of course, changes in the asymmetries among individual firms do not necessarily correspond to changes in the overall degrees of asymmetry in any one industry. For example, a firm that was previously inefficient, or relatively unsuccessful because its product line was unattractive, now succeeds in devising better processes of products. Other things being equal, this might well reduce the dispersion in the general distribution of the industry; however, it still generates an asymmetry between the considered firm and its laggard competitors. Certainly, the possibility of imitation holds out a greater potential for gain (in productivity, etc.) to firms that are relatively laggard than to firms that are relatively advanced. Ceteris paribus, therefore, there is reason to think that the process of imitation and diffusion makes for convergence. But asymmetries in the capabilities of firms impose limits on this tendency and its strength remains to be determined. (I am obviously unable to deal here with the vast literature on diffusion which would require a work of its own and I will mention only a few results relevant to the present discussion.) In turn, the higher the asymmetries among firms, the higher also is the possibility for the technological leaders (or, in any case, the most efficient producers) to modify the industrial structure in their favor, and also improve aggregate industrial performance, by eliminating the laggard producers. Vice versa, the lower the
degrees of interfirn asymmetries, the more the improvements of whatever indicator of industrial performance have to rely on widespread learning and diffusion processes.

Moreover, note that the concepts of appropriability, cumulativeness, and tacitness of technological capabilities—introduced earlier—bear a direct link with those concepts developed in industrial economics such as entry and mobility barriers in that the former entails forms of competitive differentiation both between incumbents and potential entrants and among incumbents. In this respect, the "degree of asymmetry" of an industry is a synthetic representation of both sets of phenomena.

With these remarks in mind, let me consider in more detail the relationship between features of the innovative process and patterns of industrial dynamics.

Consider, first, differences in technological opportunity, holding other characteristics of innovation (such as appropriability) constant. It is straightforward that, ceteris paribus, one would expect the rates of performance improvement over time (e.g., productivity growth) to be positively correlated with the levels of technological opportunities; however, what can one say about the characteristics of the underlying evolutionary process driven by high technological opportunities? Of course, one would expect that, the higher the opportunities are, the higher also will be the innovative learning by some producers and the selective pressures against laggard firms. That is, the higher the opportunities are, the higher also is the probability that some firms will "learn a lot," a lot more than other competitors, and that—on the grounds of their vastly superior performance—they will, so to speak, drive forward the industry by eliminating backward producers. The simulation exercises in Nelson and Winter (1982) broadly corroborate these conjectures about the relationship between degrees of technological opportunities, possibilities of differential innovative learning, and selection, leading, ceteris paribus, to rather concentrated industrial structures. One must stress, recalling the discussion in Section IV, that "opportunity" is only a necessary but not sufficient condition for its actual exploitation, while the speed of sectoral performance improvements (in productivity and product characteristics) depend on the latter. Given any notional technological opportunity, its effective exploitation by business firms will depend, as mentioned, on factors such as appropriability conditions, and, also on market variables such as the size of the market, the elasticity of demand to price and quality changes, and the degrees of industrial concentration. The rates at which opportunities are actually exploited (in terms of new/better products and more efficient processes of production) by, at least, some firms, and the rates at which these new products and processes diffuse to other firms obviously affect also the rates of change of industrial performance over time—e.g., the rates of productivity growth or the changes in output prices. In this respect, the case studies cited in paragraph V A also present some evidence on the impressive record of productivity growth and real price fall in sectors characterized by promising new paradigms and high technological opportunities, e.g., computers and semiconductors. Moreover, as regards cross-industry analyses, the evidence that one often finds on the statistical link between sectoral R & D intensity and sectoral innovative performance (e.g., sectoral productivity growth) should, in the light of this discussion, be considered as evidence that relatively high technological opportunities tend to be associated with a formalized, R & D–based mode of technological learning (Nelson 1981a).
Second, consider the impact of cumulativeness of innovative capabilities. Here, the implications are straightforward. The more technical progress is cumulative at a firm level, the more success breeds success. Firm-level cumulativeness of technological capabilities entails a nonrandom distribution of probabilities of innovative advance and path dependence: Firms that achieve higher levels of innovativeness (competitiveness) increase also their probability of maintaining or increasing their levels of competitiveness (innovativeness). Technological variety and diffusion are then likely to play only minor roles in industrial dynamics, while the rates of innovative learning of the technological leader(s) directly determine rates of change in the aggregate performance of (often highly concentrated) industries. The converse is the case where cumulativeness is relatively low, as has often been true in “supplier-dominated” sectors. Innovations are mainly embodied in equipment and components bought from other sectors, and while technological opportunities might be significant, they are mainly generated exogenously to these industrial activities. In fact, they are the result of opportunities of developing, for example, new seeds, fungicides, pesticides, tractors, and textile machinery that can be efficiently adopted in agriculture, textile, clothing, and so on. Under these circumstances, one would expect diffusion of new vintages of equipment to be the main source of industrial dynamics, while selection processes leading to market concentration are likely to be relatively weak. The evidence on the structure of these sectors (for a discussion, see for example Pavitt 1984) is in line with this hypothesis. Suppliers of new types of machinery, components, seeds, and so on have an interest in the most rapid possible diffusion of their outputs, and thus the rates of change in average performance (productivity, etc.) in the user sectors depends jointly on (a) the pace of innovation in the supplier sectors and (b) the variant conditions governing adoption. The former, clearly, puts the upper ceiling to such rates of performance improvement. The latter are especially important in explaining the average gaps between faster- and slower-moving countries (again, for an historical illustration of American versus English agriculture in the nineteenth century, see David 1975).

Third, consider the role of appropriability conditions. Of course, the “ease” of imitation of a certain innovation (and, thus, of its diffusion into the output or the production process of other firms) or the ease with which rival firms may succeed in introducing a competitive product stands, ceteris paribus, in an inverse relationship with its appropriability.

In general, the overall degrees of appropriability, the relative effectiveness of the various sources of appropriability (e.g., patenting, innovative leads, and firm-specific learning by doing), the technological opportunities and their sources (e.g., internally generated by incumbents versus external to the industry and “public”), the size and rates of growth of the market all change significantly over the trajectory of development of a technology (its “life cycle”). These factors, jointly with the conditions governing market competition (e.g., various other sorts of entry barriers, necessary minimum scale, difficulties of breaking into or enlarging markets—both domestic and foreign, price and quality elasticity of demand) govern the evolution of both industrial performances and structures.

Certainly, from an empirical standpoint, the concepts emphasized in this work (such as opportunity and appropriability) do not have obvious and objective counterparts, because they are not directly measurable and empirical studies
are still difficult and uncertain; however, in my view, statistical difficulties do not detract from their crucial interpretative importance. And, despite all practical drawbacks, it seems to me that the empirical studies now available are quite consistent with this framework of analysis. For example, it has been shown that, along what has been defined here as paradigm-specific “technological trajectories,” the net rate of entry of new firms changes. Gort and Klepper (1982) find rather robust cross-innovations evidence (based on 46 innovations) of a five-stage cycle as depicted in Figure 2. It might be worth summarizing their major conclusions: “... there is no equilibrium number of firms in an industry”; “... [the] ultimate number of producers... and the number at each preceding point in time depends upon the sequence of events to that point”; “... technical change (innovations) plays a critical role in determining both entry rates and the eventual number of firms in the market”; “... the number of firms in product markets technologically adjacent to those of a new product—that is, the number of potential entrants—influence the entry rates”; “... the onset of Stage III [flattening net entry rates] and the ensuing net exit in Stage IV is not associated with the maturity of the market as measured by market size or the growth rate in demand”; “rather it corresponds to a decrease in the rate of innovations external to the industry, a compression of profit rates, and the accumulation of valuable experience by incumbent producers” (p. 634). Of course, there is no general necessity for an innovation to pass through all five stages. Whenever radically new (and competing) innovations emerge, one may observe “truncated cycles” (on semiconductors, see Dosi 1984). Still, the major conclusion holds: The net rate of entry and, more generally, the structure of production of any one innovation (the number and size of firms, degrees of industrial concentration, entry barriers, etc.) are endogenous to the technological dynamics, and depend also on rates and modes of innovative learning and on the extent to which this learning is appropriated and internalized within firms as a rent-earning asset.

Empirical and theoretical research on the properties of different modes of industrial evolution is still at a rather early stage; however, the foregoing conjectures and findings highlight the very promising link between the studies of the microeconomic features of the processes through which people and firms search for new products and processes—the domains of the “economics of innovation”—on the one hand, and the analyses of the competitive process, and of the structures, performance, and change of industries—a typical domain of industrial economics—on the other. Innovative learning, of course, is an important competitive weapon. Moreover, the ways economic actors learn also influence the degrees to which they can exploit such a weapon competitively and ultimately change the environment in which they operate. Such a process is inherently “evolutionary,” at least in the sense that
various economic actors are forced to search for technological changes whose success will only be determined ex post by market selection. Thus, there is inevitably a distribution of "mutations" of which at least some are destined to be "mistakes." In that model, markets select both among firms and among specific technological advances (more on this in Gibbons and Metcalfe 1986); however, such an evolutionary process, unlike a strict biological analogy, is not driven by any purely random change-generating mechanism. Agents learn—from the environment, from their rivals, from their own successes and mistakes—in ways that are specific to the body of knowledge that characterizes each technology, that is, each "technological paradigm." As a consequence, the features of the evolution of each industry are, so to speak, "ordered" by the patterns of learning, and by the ways the latter influence the competitive process; The understanding of the variety of observable industrial structures, performances, and their changes implies, I suggest, a sort of "microfoundation" in the underlying modes by which economic agents accumulate knowledge and competencies on how to solve technological and organizational problems. This, possibly, remains as one of the major fields of analysis of the variegated structures and dynamics by which noncentrally planned economies search for, generate, and select technological innovations. Still, I believe, a fundamental ingredient of the explanation of such variegated industrial structures (and of why in some sectors innovating firms are small and in others large, why some innovating firms diversify beyond the boundaries of their original activities and others do not, why some firms continue to innovate and others slow down, etc.) derives from the equally variegated nature of the evolutionary processes that generated them. Finally, in this perspective, one might also try to explain (a) the levels and changes of economic performance of different countries as the joint outcome of movements of the "technological frontiers" (that is, unequivocal improvements in best-practice production techniques and production inputs), (b) processes of learning/diffusion to more "backward"/imitative firms (and countries), and (c) processes of selection associated with higher competitiveness or higher international market shares of the most successful innovators (or imitators). Clearly, in considering international differences, factors that go beyond those that differentiate firms within a common national environment need to be considered (e.g., education, financial facilities, legal institutions, cultural traits, forms of social organization). However, it will be interesting to see, first, how far the factors that emerge from the present survey can take us, and, second, how such factors interact with those broader national characteristics which I have just mentioned.

VII. Some Conclusions

The number, variety, and scope of the studies that have been reviewed (albeit a subset of the recent literature on technological change) reveal the progress that has been made over the last 20 years in the conceptualization and, to some extent, in the empirical analysis of the process of generation of innovations and their effects. Some of the themes can be considered as developments on insights and hypotheses already present in the writings of classical economists, and, after them, Schumpeter. Other elements of analysis add novel understanding of the characteristics of technical progress. Certainly, the empirical analysis of the innovative process within and across industries and countries has made a promising beginning and is pursued with vigor. Progress in this area is often con-
strained by scarcity of the relevant data, but possibly also by the "vision" and approach to empirical analysis of economists who are generally trained to consider technology among the preanalytical data of their models.

In the new view, appropriability; partial tacitness; specificity; uncertainty; variety of knowledge bases, search procedures, and opportunities; cumulative-ness; and irreversibility (all concepts defined in Sections II and III) have been recognized as general features of technological progress. Relatedly, the endogenous nature of market structures associated with the dynamics of innovation, the asymmetries among firms in technological capabilities, various phenomena of nonconvexity, history dependence, dynamic increasing returns, and the evolutionary nature of innovation/diffusion processes are some of the main elements of the process of technological change.

My impression is that there is a significant gap between the wealth of findings by economic historians, students of technology, applied industrial economists, on the one hand, and the (more limited) conceptualization of these findings in economic theory, on the other. Clearly, there will always be a difference between the "empirical stories" and the "analytical stories" of the theoreticians. The former tend inevitably to focus on the uniqueness of the details while the latter are bound to involve varying degrees of simplification and abstraction. However, the core hypotheses made by the theory should not openly conflict with empirical phenomena that show enough persistence over time and/or across eco-

11 This is not the place to engage in a discussion of economic methodology. A rich exchange of views on "theory versus history" which directly touches on many issues involving technical change is in William Parker (1986). A pertinent methodological argument, which, again, is impossible to tackle here, despite several suggestive points and disagreements, is in John Elster (1983).

nomic environments. If one believes this, then some questions immediately come to mind: For example, how does one translate the features of the innovative process, outlined above, into theory-level propositions on microeconomic behaviors, production theory, adjustment processes, and so on? Are these propositions consistent with the corresponding assumptions that one generally makes in economic analysis? Or, putting it another way, can we build incrementally upon standard assumptions in order to account for the foregoing properties of the innovative process? Whatever one's answer to these questions, the field of innovation is, in my view, particularly fascinating and challenging. Innovation and technical change have been a privileged focus of attention also for those who have been trying to model economic dynamics in unorthodox fashion, based on "evolutionary" assumptions, a much looser concept of "equilibria," a characterization of behaviors that leave great room for institutional traits and a big emphasis on competition as a selective mechanism (notably, Nelson and Winter 1982; see also Dosi et al. 1988). However, the challenge that innovative processes pose to this approach is equally formidable. It must show that assumptions which, with little doubt, are empirically more plausible, can also generate models with levels of generality somewhat comparable with those based on a more conventional approach; explore the robustness of results that so far have been obtained mainly through simulations; achieve, despite an admittedly higher complexity, that threshold of elegance that makes models appealing to the professional community.

Almost certainly, competing theories in social sciences are somewhat like competing phenotypes in complex evolutionary environments: There is no way of telling ex ante which one will be "fitter."
It is hard to doubt, however, that the domain of innovation, with the characteristics discussed in this review, are a major—and still largely unexplored—frontier of economic analysis.

References


Coïnedet, Patrick; Tournebise, Regis Larue de and Zucovitch, Ehud. "Progrès Technique et
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.


GARDINER, PAUL. "Design Trajectories for Airplanes and Automobiles During the Past Fifty Years," and "Robust and Lean Designs," in FREEMAN 1984, pp. 121–68.


(SPRU) SCIENCE POLICY RESEARCH UNIT. Success and failure in industrial innovation. London: Centre for the Study of Industrial Innovation, 1972.


STONEMAN, PAUL. Technological diffusion and the


TRUEL, JEAN-LOUIS. "L'Industrie Mondiale de Semi-