Entrepreneurship and Innovation

Zoltan J. Acs and David B. Audretsch

Prepared for the 2005 Summer Institute of the Entrepreneurship, Growth & Public Policy Division of the Max Planck Institute of Economics

May 2005
1. Introduction

Just as the economy has been besieged by a wave of technological change that has left virtually no sector of the economy untouched, scientific understanding of the innovative process -- that is, the manner by which firms innovate, and the impact such technological change has in turn on enterprises and markets -- has also undergone a revolution, which, if somewhat quieter, has been no less fundamental. Well into the 1970s, a conventional wisdom about the nature of technological change generally pervaded. This conventional wisdom had been shaped largely by scholars such as Alfred Chandler (1977), Joseph Schumpeter (1942) and John Kenneth Galbraith (1962) who had convinced a generation of scholars and policy makers that innovation and technological change lie in the domain of large corporations and that small business would fade away as the victim of its own inefficiencies.

At the heart of this conventional wisdom was the belief that monolithic enterprises exploiting market power were the driving engine of innovative activity. Schumpeter had declared the debate closed, with his proclamation in 1942 (p. 106) that, "What we have got to accept is that (the large-scale establishment) has come to be the most powerful engine of progress." Galbraith (1956, p. 86) echoed Schumpeter's sentiment, "There is no more pleasant fiction than that technological change is the product of the matchless ingenuity of the small man forced by competition to employ his wits to better his neighbor. Unhappily, it is a fiction."

At the same time, the conventional wisdom about small and new firms was that they were burdened with a size inherent handicap in terms of innovative activity. Because
they had a deficit of resources required to generate and commercialize ideas, this conventional wisdom viewed small enterprises as being largely outside of the domain of innovative activity and technological change. Thus, even after David Birch (1982) revealed the startling findings from his study that small firms provided the engine of job creation for in the U.S., most scholars still assumed that, while small businesses may create the bulk of new jobs, innovation and technological change remained beyond their sphere.

While this conventional wisdom about the singular role played by large enterprises with market power prevailed during the first three decades subsequent to the close of the Second World War II, more recently a wave of new studies has challenged this conventional wisdom. Most importantly, these studies have identified a much wider spectrum of enterprises contributing to innovative activity, and that, in particular, small entrepreneurial firms as well as large established incumbents play an important role in the innovation and process of technological change.

Taken together, these studies comprise a new understanding about the about the links between entrepreneurship, firm size and innovative activity. The purpose of this chapter is to identify this new understanding about the role that entrepreneurship and small firms play with respect to technological change and innovation and to contrast it with the previous conventional wisdom. This chapter begins with the most prevalent theory about innovation and technological change – the model of the knowledge production function. Just as the conventional wisdom was shaped largely by the available empirical data and analyses, so it is with the newer view. Thus, in the following section of this chapter, issues arising when trying to measure innovative activity are discussed.
The debate and the evidence regarding the relationship between innovative activity and firm size is examined in the third section. In the fourth section, the impact that the external industry environment exerts on technological change is identified. The role that knowledge spillovers and geographic location plays in innovative activity is explained in the fifth section. This leads to a re-interpretation of the knowledge production function when entrepreneurial activity is considered in the sixth section.

Finally, a summary and conclusions are provided in the last section. A key finding is that the conventional wisdom regarding the process of innovation technological change is generally inconsistent with the new understanding about the role of entrepreneurship in innovative activity. The empirical evidence strongly suggests that small entrepreneurial firms play a key role in generating innovations, at least in certain industries. While the conventional wisdom is derived from the Schumpeterian Hypothesis and assumption that scale economies exist in R&D effort, for which there is considerable empirical evidence, more recent evidence suggests that scale economies bestowed through the geographic proximity facilitated by spatial clusters seems to be more important than those for large enterprises in producing innovative output.

1. The Knowledge Production Function

For example, in the most prevalent model found in the literature of technological change, the model of the *knowledge production function*, formalized by Zvi Griliches (1979), firms exist exogenously and then engage in the pursuit of new economic knowledge as an input into the process of generating innovative activity.

The most decisive input in the knowledge production function is new economic knowledge. As Cohen and Klepper conclude, the greatest source generating new economic knowledge is generally considered to be R&D (Cohen and Klepper, 1991 and 1992).

When it came to empirical estimation of the knowledge production function, it became clear that measurement issues played a major role. The state of knowledge regarding innovation and technological change has generally been shaped by the nature of the data which were available to scholars for analyses. Such data have always been incomplete and, at best, represented only a proxy measure reflecting some aspect of the process of technological change. Simon Kuznets observed in 1962 that the greatest obstacle to understanding the economic role of technological change was a clear inability of scholars to measure it. More recently, Cohen and Levin (1989) warned, "A fundamental problem in the study of innovation and technical change in industry is the absence of satisfactory measures of new knowledge and its contribution to technological progress. There exists no measure of innovation that permits readily interpretable cross-industry comparisons."

Measures of technological change have typically involved one of the three major aspects of the innovative process: (1) a measure of the inputs into the innovative process,
such as R&D expenditures, or else the share of the labor force accounted for by employees involved in R&D activities; (2) an intermediate output, such as the number of inventions which have been patented; or (3) a direct measure of innovative output.

These three levels of measuring technological change have not been developed and analyzed simultaneously, but have evolved over time, roughly in the order of their presentation. That is, the first attempts to quantify technological change at all generally involved measuring some aspects of inputs into the innovative process (Scherer, 1965a; 1965b; 1967; Grabowski, 1968; Mueller, 1967; and Mansfield, 1968). Measures of R&D inputs -- first in terms of employment and later in terms of expenditures -- were only introduced on a meaningful basis enabling inter-industry and inter-firm comparisons in the late 1950s and early 1960s.

A clear limitation in using R&D activity as a proxy measure for technological change is that R&D reflects only the resources devoted to producing innovative output, but not the amount of innovative activity actually realized. That is, R&D is an input and not an output in the innovation process. In addition, Kleinknecht (1987 and 1989), Kleinknecht and Verspagen (1989), and Kleinknecht et al. (1991) have systematically shown that R&D measures incorporate only efforts made to generate innovative activity that are undertaken within formal R&D budgets and within formal R&D laboratories. They find that the extent of informal R&D is considerable, particularly in smaller enterprises.¹ And, as Mansfield (1984) points out, not all efforts within a formal R&D laboratory are directed towards generating innovative output in any case. Rather, other

---

¹ Similar results emphasizing the importance of informal R&D have been found by Santarelli and Sterlachinni (1990).
types of output, such as imitation and technology transfer, are also common goals in R&D laboratories.

As systematic data measuring the number of inventions patented were made publicly available in the mid-1960s, many scholars interpreted this new measure not only as being superior to R&D but also as reflecting innovative output. In fact, the use of patented inventions is not a measure of innovative output, but is rather a type of intermediate output measure. A patent reflects new technical knowledge, but it does not indicate whether this knowledge has a positive economic value. Only those inventions which have been successfully introduced in the market can claim that they are innovations as well. While innovations and inventions are related, they are not identical. The distinction is that an innovation is "...a process that begins with an invention, proceeds with the development of the invention, and results in the introduction of a new product, process or service to the marketplace" (Edwards and Gordon, 1984, p. 1).

Besides the fact that many, if not most, patented inventions do not result in an innovation, a second important limitation of patent measures as an indicator of innovative activity is that they do not capture all of the innovations actually made. In fact, many inventions which result in innovations are not patented. The tendency of patented inventions to result in innovations and of innovations to be the result of inventions which were patented combine into what F.M. Scherer (1983a) has termed as the propensity to patent. It is the uncertainty about the stability of the propensity to patent across enterprises and across industries that casts doubt upon the reliability of patent measures.²

---
² For example, Shepherd (1979, p. 40) has concluded that, "Patents are a notoriously weak measure. Most of the eighty thousand patents issued each year are worthless and
According to Scherer (1983, pp. 107-108), "The quantity and quality of industry patenting may depend upon chance, how readily a technology lends itself to patent protection, and business decision-makers' varying perceptions of how much advantage they will derive from patent rights. Not much of a systematic nature is known about these phenomena, which can be characterized as differences in the propensity to patent."

Mansfield (1984, p. 462) has explained why the propensity to patent may vary so much across markets: "The value and cost of individual patents vary enormously within and across industries ... Many inventions are not patented. And in some industries, like electronics, there is considerable speculation that the patent system is being bypassed to a greater extent than in the past. Some types of technologies are more likely to be patented than others." The implications are that comparisons between enterprises and across industries may be misleading. According to Cohen and Levin (1989), "There are significant problems with patent counts as a measure of innovation, some of which affect both within-industry and between-industry comparisons."

Thus, even as new and superior sources of patent data have been introduced, such as the new measure of patented inventions from the computerization by the U.S. Patent Office (Hall et al., 1986; Jaffe, 1986; Pakes and Griliches, 1980 and 1984) as well as in Europe (Schwalbach and Zimmermann, 1991; Greif, 1989; and Greif and Potkowik, 1990), the reliability of these data as measures of innovative activity has been severely challenged. For example, Pakes and Griliches (1980, p. 378) warn that "patents are a flawed measure (of innovative output); particularly since not all new innovations are

are never used. Still others have negative social value. They are used as "blocking" patents to stop innovation, or they simply are developed to keep competition out."
patented and since patents differ greatly in their economic impact." And in addressing the question, "Patents as indicators of what?", Griliches (1990, p. 1669) concludes that, "Ideally, we might hope that patent statistics would provide a measure of the (innovative) output ... The reality, however, is very far from it. The dream of getting hold of an output indicator of inventive activity is one of the strong motivating forces for economic research in this area."³

It was not before well into the 1970s that systematic attempts were made to provide a direct measure of the innovative output. Thus, it should be emphasized that the conventional wisdom regarding innovation and technological change was based primarily upon the evidence derived from analysing R&D data, which essentially measure inputs into the process of technological change, and patented inventions, which are a measure of intermediate output at best.

The first serious attempt to directly measure innovative output was by the Gellman Research Associates (1976) for the National Science Foundation. Gellman identified 500 major innovations that were introduced into the market between 1953 and 1973 in the United States, the United Kingdom, Japan, West Germany, France, and Canada. The data base was compiled by an international panel of experts, who identified those innovations representing the "most significant new industrial products and

³ Chakrabarti and Halperin (1990) use a fairly standard source of data for U.S. patents issued by the U.S. Office of Patents and Trademarks, the BRS/PATSEARCH online database, to identify the number of inventions patented by over 470 enterprises between 1975 and 1986. Of particular interest is their comparison between the propensity of firms to patent and company R&D expenditures, and a measure not often found in the economics literature the number of published papers and publications contributed by employees of each firm. Not only do they bring together data from a number of rich sources, but they compare how the relationships between the various measures of innovative activity vary across firm size.
processes, in terms of their technological importance and economic and social impact" (National Science Board, 1975, p. 100).

A second and comparable data base once again involved the Gellman Research Associates (1982), this time for the U.S. Small Business Administration. In their second study, Gellman compiled a total of 635 U.S. innovations, including 45 from the earlier study for the National Science Foundation. The additional 590 innovations were selected from fourteen industry trade journals for the period 1970-1979. About 43 percent of the sample was selected from the award winning innovations described in the Industrial Research & Development magazine.

The third data source that has attempted to directly measure innovation activity was compiled at the Science Policy Research Unit (SPRU) at the University of Sussex in the United Kingdom. The SPRU data consist of a survey of 4,378 innovations that were identified over a period of fifteen years. The survey was compiled by writing to experts in each industry and requesting them to identify "significant technical innovations that had been successfully commercialized in the United Kingdom since 1945, and to name the firm responsible" (Pavitt et al., 1987, p. 299).

The most recent and most ambitious major data base providing a direct measure of innovative activity is the U.S. Small Business Administration's Innovation Data Base (SBIDB). The data base consists of 8,074 innovations commercially introduced in the U.S. in 1982. A private firm, The Futures Group, compiled the data and performed quality-control analyses for the U.S. Small Business Administration by examining over

---

4 The SPRU innovation data are explained in considerable detail in Pavitt et al. (1987), Townsend et al. (1981), Robson and Townsend (1984), and Rothwell (1989).
one hundred technology, engineering, and trade journals, spanning every industry in manufacturing. From the sections in each trade journal listing innovations and new products, a data base consisting of the innovations by four-digit standard industrial classification (SIC) industries was formed. These data were implemented by Acs and Audretsch (1987, 1988b, and 1990) to analyze the relationships between firm size and technological change and market structure and technological change, where a direct rather than indirect measure of innovative activity is used.

In their 1990 study (chapter two), Acs and Audretsch directly compare these four data bases directly measuring innovative activity and find that they generally provide similar qualitative results. For example, while the Gellman data base identified small firms as contributing 2.45 times more innovations per employee than do large firms, the U.S. Small Business Administration's Innovation Data Base finds that small firms introduce 2.38 more innovations per employee than do their larger counterparts. In general, these four data bases reveal similar patterns with respect to the distribution of innovations across manufacturing industries and between large and small enterprises. These similarities emerge, despite the obviously different methods used to compile the data, especially in terms of sampling and standard of significance.

Just as for the more traditional measures of technological change, there are also certain limitations associated with the direct measure of innovative activity. In fact, one of the main qualifications is common among all three measures -- the implicit assumption of homogeneity of units. That is, just as it is implicitly assumed that each dollar of R&D

5 A detailed description of the U.S. Small Business Administration’s Innovation Data Base can be found in chapter two of Acs and Audretsch (1990)
makes the same contribution to technological change, and that each invention which is patented is equally valuable, the output measure implicitly assumes that innovations are of equal importance.\(^6\) As Cohen and Levin (1989) observe, "In most studies, process innovation is not distinguished from product innovation; basic and applied research are not distinguished from development." Thus, the increase in the firm's market value resulting from each innovation, dollar expended on R&D, and patent, is implicitly assumed to be homogeneous -- an assumption which clearly violates real world observation.

In order to at least approximate the market value associated with innovative activity, FitzRoy and Kraft (1990 and 1991) follow the example of Pakes (1985), Connolly et al. (1986), and Connolly and Hirschey (1984). Based on data for 57 West German firms in the metalworking sector, FitzRoy and Kraft (1990 and 1991) measure innovation as the "proportion of sales consisting of products introduced within the last five years." Presumably the greater the market value of a given product innovation, the higher would be the proportion of sales accounted for by new products.

Similarly, Graf von der Schulenburg and Wagner (1991 and 1992) are able to provide one of the first applications of a direct measure of innovative activity in West Germany. Their measure is from the IFO Institute and is defined as the "percentage of shipments of those products which were introduced recently into the market and are still

---

\(^6\) It should be emphasized, however, that Acs and Audretsch (1990, chapter two) perform a careful analysis of the significance of the innovations based on four broad categories ranking the importance of each innovation.
in the entry phase.\textsuperscript{7} Like the measure of innovative activity used by FitzRoy and Kraft (1990 and 1991), the Graf von der Schulenburg and Wagner measure reflects the market value of the innovation and therefore attempts to overcome one of the major weaknesses in most of the other direct and indirect measures of innovative activity.

The knowledge production function has been found to hold most strongly at broader levels of aggregation. The most innovative countries are those with the greatest investments to R&D. Little innovative output is associated with less developed countries, which are characterized by a paucity of production of new economic knowledge. Similarly, the most innovative industries, also tend to be characterized by considerable investments in R&D and new economic knowledge. Not only are industries such as computers, pharmaceuticals and instruments high in R&D inputs that generate new economic knowledge, but also in terms of innovative outputs (Audretsch, 1995). By contrast, industries with little R&D, such as wood products, textiles and paper, also tend to produce only a negligible amount of innovative output. Thus, the knowledge production model linking knowledge generating inputs to outputs certainly holds at the more aggregated levels of economic activity.

Where the relationship becomes less compelling is at the disaggregated microeconomic level of the enterprise, establishment, or even line of business. For example, While Acs and Audretsch (1990) found that the simple correlation between R&D inputs and innovative output was 0.84 for four-digit standard industrial

\textsuperscript{7} The data based used by Graf von der Schulenburg and Wagner (1991) is the IFO-Innovations-Test and is explained in greater detail in Oppenlander (1990), and Konig and Zimmermann (1986).
classification (SIC) manufacturing industries in the United States, it was only about half, 0.40 among the largest U.S. corporations.

The model of the knowledge production function becomes even less compelling in view of the recent wave of studies revealing that small enterprises serve as the engine of innovative activity in certain industries. These results are startling, because as Scherer (1991) observes, the bulk of industrial R&D is undertaken in the largest corporations; small enterprises account only for a minor share of R&D inputs.

2. The Role of Firm Size

At the heart of the conventional wisdom has been the belief that large enterprises able to exploit at least some market power are the engine of technological change. This view dates back at least to Schumpeter, who in *Capitalism, Socialism and Democracy* (1942, p. 101) argued that, "The monopolist firm will generate a larger supply of innovations because there are advantages which, though not strictly unattainable on the competitive level of enterprise, are as a matter of fact secured only on the monopoly level." The Schumpeterian thesis, then, is that large enterprises are uniquely endowed to exploit innovative opportunities. That is, market dominance is a prerequisite to undertaking the risks and uncertainties associated with innovation. It is the possibility of acquiring quasi-rents that serves as the catalyst for large-firm innovation.

Five factors favoring the innovative advantage of large enterprises have been identified in the literature. First is the argument that innovative activity requires a high fixed cost. As Comanor (1967) observes, R&D typically involves a "lumpy" process that
yields scale economies. Similarly, Galbraith (1956, p. 87) argues, "Because development is costly, it follows that it can be carried on only by a firm that has the resources which are associated with considerable size."

Second, only firms that are large enough to attain at least temporary market power will choose innovation as a means for maximization (Kamien and Schwartz, 1975). This is because the ability of firms to appropriate the economic returns accruing from R&D and other knowledge-generating investments is directly related to the extent of that enterprise's market power (Cohen and Klepper, 1990 and 1991; Levin et al., 1985 and 1987; and Cohen et al., 1987). Third, R&D is a risky investment; small firms engaging in R&D make themselves vulnerable by investing a large proportion of their resources in a single project. However, their larger counterparts can reduce the risk accompanying innovation through diversification into simultaneous research projects. The larger firm is also more likely to find an economic application of the uncertain outcomes resulting from innovative activity (Nelson, 1959).

Fourth, scale economies in production may also provide scope economies for R&D. Scherer (1991) notes that economies of scale in promotion and in distribution facilitate the penetration of new products, thus enabling larger firms to enjoy a greater profit potential from innovation. Finally, an innovation yielding cost reductions of a given percentage results in higher profit margins for larger firms than for smaller firms.

There is also substantial evidence that technological change -- or rather, one aspect of technological change reflected by one of the three measures discussed in the
previous section, R&D -- is, in fact, positively related to firm size. The plethora of empirical studies relating R&D to firm size is most thoroughly reviewed in Acs and Audretsch (1990, chapter three), Baldwin and Scott (1987), and Cohen and Levin (1989). The empirical evidence generally seems to confirm Scherer's (1982, pp. 234-235) conclusion that the results "tilt on the side of supporting the Schumpeterian Hypothesis that size is conducive to vigorous conduct of R&D".

In one of the most important studies, Scherer (1984) used the U.S. Federal Trade Commission's Line of Business Data to estimate the elasticity of R&D spending with respect to firm sales for 196 industries. He found evidence of increasing returns to scale (an elasticity exceeding unity) for about twenty percent of the industries, constant returns to scale for a little less than three-quarters of the industries, and diminishing returns (an elasticity less than unity) in less than ten percent of the industries. These results were consistent with the findings of Soete (1979) that R&D intensity increases along with firm size, at least for a sample of the largest U.S. corporations.

While the Scherer (1984) and Soete (1979) studies were restricted to relatively large enterprises, Bound et al. (1984) included a much wider spectrum of firm sizes in their sample of 1,492 firms from the 1976 COMPSTAT data. They found that R&D increases more than proportionately along with firm size for the smaller firms, but that a fairly linear relationship exists for larger firms. Despite the somewhat more ambiguous

8 Fisher and Temin (1973) demonstrated that the Schumpeterian Hypothesis could not be substantiated unless it was established that the elasticity of innovative output with respect to firm size exceeds one. They pointed out that if scale economies in R&D do exist, a firm's size may grow faster than its R&D activities. Kohn and Scott (1982) later showed that if the elasticity of R&D input with respect to firm size is greater than unity, then the elasticity of R&D output with respect to firm size must also be greater than one.
findings in still other studies (Comanor, 1967; Mansfield, 1981 and 1983; and Mansfield et al., 1982), the empirical evidence seems to generally support the Schumpeterian hypothesis that research effort is positively associated with firm size.

The studies relating patents to firm size are considerably less ambiguous. Here the findings unequivocally suggest that "the evidence leans weakly against the Schumpeterian conjecture that the largest sellers are especially fecund sources of patented inventions" (Scherer, 1982, p. 235). In one of the most important studies, Scherer (1965b) used the Fortune annual survey of the 500 largest U.S. industrial corporations. He related the 1955 firm sales to the number of patents in 1959 for 448 firms. Scherer found that the number of patented inventions increases less than proportionately along with firm size. Scherer's results were later confirmed by Bound et al. (1984) in the study mentioned above. Basing their study on 2,852 companies and 4,553 patenting entities, they determined that the small firms (with less than $10 million in sales) accounted for 4.3 percent of the sales from the entire sample, but 5.7 percent of the patents.

Such results are not limited to the U.S. Schwalbach and Zimmermann (1991) find that the propensity to patent is less for the largest firms in West Germany than for the medium-sized enterprises included in their sample.

A number of explanations have emerged why smaller enterprises may, in fact, tend to have an innovative advantage, at least in certain industries. Rothwell (1989) suggests that the factors yielding small firms with the innovative advantage generally emanate from the difference in management structures between large and small firms.
For example, Scherer (1991) argues that the bureaucratic organization of large firms is not conducive to undertaking risky R&D. The decision to innovate must survive layers of bureaucratic resistance, where an inertia regarding risk results in a bias against undertaking new projects. However, in the small firm the decision to innovate is made by relatively few people.

Second, innovative activity may flourish the most in environments free of bureaucratic constraints (Link and Bozeman, 1991). That is, a number of small-firm ventures have benefited from the exodus of researchers who felt thwarted by the managerial restraints in a larger firm. Finally, it has been argued that while the larger firms reward the best researchers by promoting them out of research to management positions, the smaller firms place innovative activity at the center of their competitive strategy (Scherer, 1991).

Scherer (1988, pp. 4-5) has summarized the advantages small firms may have in innovative activity: "Smaller enterprises make their impressive contributions to innovation because of several advantages they possess compared to large-size corporations. One important strength is that they are less bureaucratic, without layers of "abominable no-men" who block daring ventures in a more highly structured organization. Second, and something that is often overlooked, many advances in technology accumulate upon a myriad of detailed inventions involving individual components, materials, and fabrication techniques. The sales possibilities for making such narrow, detailed advances are often too modest to interest giant corporations. An individual entrepreneur's juices will flow over a new product or process with sales prospects in the millions of dollars per year, whereas few large corporations can work up
much excitement over such small fish, nor can they accommodate small ventures easily into their organizational structures. Third, it is easier to sustain a fever pitch of excitement in small organization, where the links between challenges, staff, and potential rewards are tight. "All-nighters" through which tough technical problems are solved expeditiously are common."

Two other ways that small enterprises can compensate for their lack of R&D is through spillovers and spin-offs. Typically an employee from an established large corporation, often a scientist or engineer working in a research laboratory, will have an idea for an invention and ultimately for an innovation. Accompanying this potential innovation is an expected net return from the new product. The inventor would expect to be compensated for his/her potential innovation accordingly. If the company has a different, presumably lower, valuation of the potential innovation, it may decide either not to pursue its development, or that it merits a lower level of compensation than that expected by the employee.

In either case, the employee will weigh the alternative of starting his/her own firm. If the gap in the expected return accruing from the potential innovation between the inventor and the corporate decision maker is sufficiently large, and if the cost of starting a new firm is sufficiently low, the employee may decide to leave the large corporation and establish a new enterprise. Since the knowledge was generated in the established corporation, the new start-up is considered to be a spin-off from the existing firm. Such start-ups typically do not have direct access to a large R&D laboratory. Rather, these small firms succeed in exploiting the knowledge and experience accrued from the R&D laboratories with their previous employers.
The research laboratories of universities provide a source of innovation-generating knowledge that is available to private enterprises for commercial exploitation. Jaffe (1989) and Acs, Audretsch, and Feldman (1992), for example, found that the knowledge created in university laboratories "spills over" to contribute to the generation of commercial innovations by private enterprises. Acs, Audretsch, and Feldman (1994) found persuasive evidence that spillovers from university research contribute more to the innovative activity of small firms than to the innovative activity of large corporations. Similarly, Link and Rees (1990) surveyed 209 innovating firms to examine the relationship between firm size and university research. They found that, in fact, large firms are more active in university-based research. However, small- and medium-sized enterprises apparently are better able to exploit their university-based associations and generate innovations. Link and Rees (1990) conclude that, contrary to the conventional wisdom, diseconomies of scale in producing innovations exist in large firms. They attribute these diseconomies of scale to the "inherent bureaucratization process which inhibits both innovative activity and the speed with which new inventions move through the corporate system towards the market" (Link and Rees, 1990, p. 25).

Thus, just as there are persuasive theories defending the original Schumpeterian Hypothesis that large corporations are a prerequisite for technological change, there are also substantial theories predicting that small enterprises should have the innovative advantage, at least in certain industries. As described above, the empirical evidence based on the input measure of technological change, R&D, tilts decidedly in favor of the Schumpeterian Hypothesis. However, as also described above, the empirical results are somewhat more ambiguous for the measure of intermediate output -- the number of
patented inventions. It was not until direct measures of innovative output became available that the full picture of the process of technological change could be obtained.

Using this new measure of innovative output from the U.S. Small Business Administration's Innovation Data Base, Acs and Audretsch (1990) shows that, in fact, the most innovative U.S. firms are large corporations. Further, the most innovative American corporations also tended to have large R&D laboratories and be R&D intensive. At first glance, these findings based on direct measures of innovative activity seems to confirm the conventional wisdom. However, in the most innovative four-digit standard industrial classification (SIC) industries, large firms, defined as enterprises with at least 500 employees, contributed more innovations in some instances, while in other industries small firms produced more innovations. For example, in computers and process control instruments small firms contributed the bulk of the innovations. By contrast in the pharmaceutical preparation and aircraft industries the large firms were much more innovative.

Probably their best measure of innovative activity is the total innovation rate, which is defined as the total number of innovations per one thousand employees in each industry. The large-firm innovation rate is defined as the number of innovations made by firms with at least 500 employees, divided by the number of employees (thousands) in large firms. The small-firm innovation rate is analogously defined as the number of innovations contributed by firms with fewer than 500 employees, divided by the number of employees (thousands) in small firms.
The innovation rates, or the number of innovations per thousand employees, have the advantage in that they measure large- and small-firm innovative activity relative to the presence of large and small firms in any given industry. That is, in making a direct comparison between large- and small-firm innovative activity, the absolute number of innovations contributed by large firms and small enterprises is somewhat misleading, since these measures are not standardized by the relative presence of large and small firms in each industry. When a direct comparison is made between the innovative activity of large and small firms, the innovation rates are presumably a more reliable measure of innovative intensity because they are weighted by the relative presence of small and large enterprises in any given industry. Thus, while large firms in manufacturing introduced 2,445 innovations in 1982, and small firms contributed slightly fewer, 1,954, small-firm employment was only half as great as large-firm employment, yielding an average small-firm innovation rate in manufacturing of 0.309, compared to a large-firm innovation rate of 0.202 (Acs and Audretsch, 1988 and 1990).

The most important and careful study to date documenting the role of German SMEs (enterprises with fewer than 500 employees) in innovative activity was undertaken by a team of researchers at the Zentrum fuer Europaeische Wirtschaftsforschung (ZEW) led by Dietmar Harhoff and Georg Licht (1996). They analyzed the findings made possible by the Mannheim Innovation Data Base. This data base measures the extent of innovative activity in German firms between 1990 and 1992. Harhoff and Licht (1996) use the data base to identify that 12 percent of the research and development expenditures in (West) German firms comes from SMEs (defined as having fewer than 500 employees).
Harhoff and Licht show that the likelihood of a firm not innovating decreases with firm size. For example, 52 percent of firms with fewer than 50 employees were not innovative. By contrast, only 15 percent of the firms with at least 1,000 employees were not innovative. More striking is that the smallest firms that do innovate have a greater propensity to be innovative without undertaking formal research and development. While only 3 percent of the largest corporations in Germany are innovative without undertaking formal R&D, one-quarter of the innovative firms with fewer than 50 employees are innovative without formal R&D.

The study also shows that even fewer SMEs in the five new German Laender are innovative than is the case in West Germany. Over two-thirds of the smallest SMEs in East Germany are not innovative, and they are less than half as likely to undertake R&D as are their Western counterparts.

Systematic empirical evidence also suggests that the German *Mittelstand* is confronted by considerable barriers to innovative activity. Beise and Licht (1996) analyzed the *Mannheimer Innovationspanel* consisting of 43,300 innovating firms to identify the main barriers to innovative activity confronting German small- and medium sized enterprises. The major barrier to innovation listed in both 1992 and 1994 was too high of a gestation period required for innovative activity. In 1994 nearly 60 percent of German SMEs reported that too long of a high gestation period required to innovate was a very important barrier to innovative activity. Other major barriers to innovative activity include legal restrictions & restrictive government policies, too long of duration required to obtain government approval for a new product, a shortage of finance capital, a lack of competent employees, and too high of a risk.
Thus, there is considerable evidence suggesting that, in contrast to the findings for R&D inputs and patented inventions, small enterprises apparently play an important role in generating innovative activity, at least in certain industries. By relating the innovative output of each firm to its size, it is also possible to shed new light on the Schumpeterian Hypothesis. In their 1991a study, Acs and Audretsch find that there is no evidence that increasing returns to R&D expenditures exist in producing innovative output. In fact, with just several exceptions, diminishing returns to R&D are the rule. This study made it possible to resolve the apparent paradox in the literature that R&D inputs increase at more than a proportional rate along with firm size, while the generation of patented inventions does not. That is, while larger firms are observed to undertake a greater effort towards R&D, each additional dollar of R&D is found to yield less in terms of innovative output.

3. The Industry Context

In comparison to the number of studies investigating the relationship between firm size and technological change, those examining the relationship between innovation and the external industry structure or environment are what Baldwin and Scott (1987, p. 89) term "miniscule" in number. In fact, the most comprehensive and insightful evidence has been made possible by utilizing the Federal Trade Commission's Line of Business Data. Using 236 manufacturing industry categories, which are defined at both the three- and four-digit SIC level, Scherer (1983a) found that 1974 company R&D expenditures divided by sales was positively related to the 1974 four-firm concentration ratio. Scherer (1983a, p. 225) concluded that, "although one cannot be certain, it appears that the
advantages a high market share confers in appropriating R&D benefits provide the most likely explanation of the observed R&D-concentrator associations."

Scott (1984) also used the FTC Line of Business Survey Data and found the U-shaped relationship between market concentration and R&D. However, when he controlled for the fixed effects for two-digit SIC industries, no significant relationship could be found between concentration and R&D. These results are consistent with a series of studies by Levin et al. (1985 and 1987), Levin and Reiss (1984), and Cohen et al. (1987). Using data from a survey of R&D executives in 130 industries, which were matched with FTC Line of Business Industry Groups, Cohen et al. (1987) and Levin et al. (1987) found little support for the contention that industrial concentration is a significant and systematic determinant of R&D effort.

While it has been hypothesized that firms in concentrated industries are better able to capture the rents accruing from an innovation, and therefore have a greater incentive to undertake innovative activity, there are other market structure variables that also influence the ease with which economic rents can be appropriated. For example, Comanor (1967) argued and found that, based on a measure of minimum efficient scale, there is less R&D effort (average number of research personnel divided by total employment) in industries with very low scale economies. However, he also found that in industries with a high minimum efficient scale, R&D effort was also relatively low. Comanor interpreted his results to suggest that, where entry barriers are relatively low, there is little incentive to innovate, since the entry subsequent to innovation would quickly erode any economic rents. At the same time, in industries with high entry barriers, the absence of potential entry may reduce the incentives to innovate.
Because many studies have generally found positive relationships between market concentration and R&D, and between the extent of barriers to entry and R&D, it would seem that the conventional wisdom built around the Schumpeterian Hypothesis has been confirmed. However, when the direct measure of innovative output is related to market concentration, Acs and Audretsch (1988b and 1990) find a pointedly different relationship emerges. In fact, there appears to be unequivocal evidence that concentration exerts a negative influence on the number of innovations being made in an industry.

Acs and Audretsch (1987, 1988b, and 1990) found that not only does market structure influence the total amount of innovative activity, but also the relative innovative advantage between large and small enterprises. The differences between the innovation rates of large and small firms examined in the previous section can generally be explained by (1) the degree of capital intensity, (2) the extent to which an industry is concentrated, (3) the total innovative intensity, and (4) the extent to which an industry is comprised of small firms. In particular, the relative innovative advantage of large firms tends to be promoted in industries that are capital-intensive, advertising intensive, concentrated, and highly unionized. By contrast, in industries that are highly innovative and composed predominantly of large firms, the relative innovative advantage is held by small enterprises.

4. The Geographic Context
The evidence revealing small enterprises to be the engine of innovative activity in certain industries, despite an obvious lack of formal R&D activities, raises the question about the source of knowledge inputs for small enterprises. The answer emerging from a series of studies (Jaffe, 1990) is from other, third-party, firms or research institutions,
such as universities. Economic knowledge may *spill over* from the firm or research institution creating it for application by other firms.

That knowledge spills over is barely dispute. However, the geographic range of such knowledge spillovers is greatly contested. In disputing the importance of knowledge externalities in explaining the geographic concentration of economic activity, Krugman (1991) and others do not question the existence or importance of such knowledge spillovers. In fact, they argue that such knowledge externalities are so important and forceful that there is no compelling reason for a geographic boundary to limit the spatial extent of the spillover. According to this line of thinking, the concern is not that knowledge does not spill over but that it should stop spilling over just because it hits a geographic border, such as a city limit, state line, or national boundary.

A recent body of empirical evidence clearly suggests that R&D and other sources of knowledge not only generate externalities, but studies by Audretsch and Feldman (1996), Jaffe (1989), Audretsch and Stephan (1996), Anselin, Varga and Acs (1997 and 2000), and Jaffe, Trajtenberg and Henderson (1993) suggest that such knowledge spillovers tend to be geographically bounded within the region where the new economic knowledge was created. That is, new economic knowledge may spill-over but the geographic extent of such knowledge spillovers is limited.

Krugman [1991a, p. 53] has argued that economists should abandon any attempts at measuring knowledge spillovers because "...knowledge flows are invisible, they leave no paper trail by which they may be measured and tracked." But as Jaffe, Trajtenberg and

Studies identifying the extent of knowledge spillovers are based on the knowledge production function. Jaffe (1989) modified the knowledge production function approach to a model specified for spatial and product dimensions:

\[ I_{si} = IRD_{si}^{\beta_1} * UR_{si}^{\beta_2} * (UR_{si} * GC_{si}^{\beta_3}) * \varepsilon_{si} \]  \hspace{1cm} (1)

where \( I \) is innovative output, \( IRD \) is private corporate expenditures on R&D, \( UR \) is the research expenditures undertaken at universities, and \( GC \) measures the geographic coincidence of university and corporate research. The unit of observation for estimation was at the spatial level, \( s \), a state, and industry level, \( i \). Estimation of equation (1) essentially shifted the knowledge production function from the unit of observation of a firm to that of a geographic unit.

Implicitly contained within the knowledge production function model is the assumption that innovative activity should take place in those regions, \( s \), where the direct knowledge-generating inputs are the greatest, and where knowledge spillovers are the most prevalent. Audretsch and Feldman (1996), Anselin, Acs and Varga (1997 and 2000) and Audretsch and Stephan (1996) link the propensity for innovative activity to cluster together to industry specific characteristics, most notably the relative importance of knowledge spillovers.

5. The Knowledge Production Function Reconsidered
The model of the knowledge production function becomes even less compelling in view of the evidence documented in Section 3 that entrepreneurial small firms are the engine of innovative activity in some industries, which raises the question, "Where do new and small firms get the innovation producing inputs, that is the knowledge?"

The appropriability problem, or the ability to capture the revenues accruing from investments in new knowledge, confronting the individual may converge with that confronting the firm. Economic agents can and do work for firms, and even if they do not, they can potentially be employed by an incumbent firm. In fact, in a model of perfect information with no agency costs, any positive economies of scale or scope will ensure that the appropriability problems of the firm and individual converge. If an agent has an idea for doing something different than is currently being practiced by the incumbent enterprises -- both in terms of a new product or process and in terms of organization -- the idea, which can be termed as an innovation, will be presented to the incumbent enterprise. Because of the assumption of perfect knowledge, both the firm and the agent would agree upon the expected value of the innovation. But to the degree that any economies of scale or scope exist, the expected value of implementing the innovation within the incumbent enterprise will exceed that of taking the innovation outside of the incumbent firm to start a new enterprise. Thus, the incumbent firm and the inventor of the idea would be expected to reach a bargain splitting the value added to the firm contributed by the innovation. The payment to the inventor -- either in terms of a higher wage or some other means of remuneration -- would be bounded between the expected value of the innovation if it implemented by the incumbent enterprise on the upper end,
and by the return that the agent could expect to earn if he used it to launch a new enterprise on the lower end

A different model refocuses the unit of observation away from firms deciding whether to increase their output from a level of zero to some positive amount in a new industry, to individual agents in possession of new knowledge that, due to uncertainty, may or may not have some positive economic value. It is the uncertainty inherent in new economic knowledge, combined with asymmetries between the agent possessing that knowledge and the decision making vertical hierarchy of the incumbent organization with respect to its expected value that potentially leads to a gap between the valuation of that knowledge.

Divergences in the expected value regarding new knowledge will, under certain conditions, lead an agent to exercise what Albert O. Hirschman (1970) has termed as exit rather than voice, and depart from an incumbent enterprise to launch a new firm. But who is right, the departing agents or those agents remaining in the organizational decision making hierarchy who, by assigning the new idea a relatively low value, have effectively driven the agent with the potential innovation away? Ex post the answer may not be too difficult. But given the uncertainty inherent in new knowledge, the answer is anything but trivial a priori.

This initial condition of not just uncertainty, but greater degree of uncertainty vis-à-vis incumbent enterprises in the industry is captured in the theory of firm selection and industry evolution proposed by Boyan Jovanovic (1982). The theory of firm selection is particularly appealing in view of the rather startling size of most new firms. For example,
the mean size of more than 11,000 new-firm startups in the manufacturing sector in the United States was found to be fewer than eight workers per firm. While the minimum efficient scale (MES) varies substantially across industries, and even to some degree across various product classes within any given industry, the observed size of most new firms is sufficiently small to ensure that the bulk of new firms will be operating at a suboptimal scale of output. Why would an entrepreneur start a new firm that would immediately be confronted by scale disadvantages?

An implication of the theory of firm selection is that new firms may begin at a small, even suboptimal, scale of output, and then if merited by subsequent performance expand. Those new firms that are successful will grow, whereas those that are not successful will remain small and may ultimately be forced to exit from the industry if they are operating at a suboptimal scale of output.

An important finding of Audretsch 1995) verified in a systematic and comprehensive series of studies contained in the reviews by Caves (1998), Sutton (1997) and Geroski (1995) is that although entry may still occur in industries characterized by a high degree of scale economies, the likelihood of survival is considerably less. People will start new firms in an attempt to appropriate the expected value of their new ideas, or potential innovations, particularly under the entrepreneurial regime. As entrepreneurs gain experience in the market they learn in at least two ways. First, they discover whether they possess the right stuff, in terms of producing goods and offering services for which sufficient demand exists, as well as whether they can product that good more efficiently than their rivals. Second, they learn whether they can adapt to market conditions as well as to strategies engaged in by rival firms. In terms of the first type of learning,
entrepreneurs who discover that they have a viable firm will tend to expand and ultimately survive. But what about those entrepreneurs who discover that they are either not efficient or not offering a product for which there is a viable demand? The answer is, *It depends -- on the extent of scale economies as well as on conditions of demand.* The consequences of not being able to grow will depend, to a large degree, on the extent of scale economies. Thus, in markets with only negligible scale economies, firms have a considerably greater likelihood of survival. However, where scale economies play an important role the consequences of not growing are substantially more severe, as evidenced by a lower likelihood of survival.

What emerges from the new evolutionary theories and empirical evidence on the role of small firms is that markets are in motion, with a lot of new firms entering the industry and a lot of firms exiting out of the industry. The evolutionary view of the process of industry evolution is that new firms typically start at a very small scale of output. They are motivated by the desire to appropriate the expected value of new economic knowledge. But, depending upon the extent of scale economies in the industry, the firm may not be able to remain viable indefinitely at its startup size. Rather, if scale economies are anything other than negligible, the new firm is likely to have to grow to survival. The temporary survival of new firms is presumably supported through the deployment of a strategy of compensating factor differentials that enables the firm to discover whether or not it has a viable product.

The empirical evidence (Caves, 1998; Sutton, 1997 and Geroski, 1995) supports such an evolutionary view of the role of new firms in manufacturing, because the post-entry growth of firms that survive tends to be spurred by the extent to which there is a
gap between the MES level of output and the size of the firm. However, the likelihood of any particular new firm surviving tends to decrease as this gap increases. Such new suboptimal scale firms are apparently engaged in the selection process. Only those firms offering a viable product that can be produced efficiently will grow and ultimately approach or attain the MES level of output. The remainder will stagnate, and depending upon the severity of the other selection mechanism -- the extent of scale economies -- may ultimately be forced to exit out of the industry. Thus, the persistence of an asymmetric firm-size distribution biased towards small-scale enterprise reflects the continuing process of the entry of new firms into industries and not necessarily the permanence of such small and sub-optimal enterprises over the long run. Although the skewed size distribution of firms persists with remarkable stability over long periods of time, a constant set of small and suboptimal scale firms does not appear to be responsible for this skewed distribution. Rather, by serving as agents of change, entrepreneurial firms provide an essential source of new ideas and experimentation that otherwise would remain untapped in the economy.

6. **Conclusions**

Within a generation, scholarship has produced theories, evidence and new insights that have dramatically changed the prevalent view about the role of entrepreneurship in innovation and technological change. The conventional wisdom held that small firms inherently have a deficit of knowledge assets, burdening them with a clear and distinct disadvantage in generating innovative output. This view was certainly consistent with the
early interpretation of the knowledge production function. As Chandler (1990) concluded, “to compete globally you have to be big.”

More recent scholarship has produced a revised view that identifies entrepreneurial small firms as making a crucial contribution to innovative activity and technological change. There are two hypotheses why scholarship about the role of small firms has evolved so drastically within such a short period. This first is that, as explained in this Chapter, the measurement of innovative output and technological change has greatly improved. As long as the main instruments to measuring innovative activity were restricted to inputs into the innovative process, such as expenditures on formal R&D, many or even most of the innovative activities by smaller enterprises simply remained hidden from the radar screen of researchers. With the development of measures focusing on measures of innovative output, the vital contribution of small firms became prominent, resulting in the emergence of not just the recognition that small firms provide an engine of innovative activity, at least in some industry contexts, but also of new theories to explain and understand how and why small firms access knowledge and new ideas. This first hypothesis would suggest that, in fact, small firms have always made these types of innovative contributions, but they remained hidden and mostly unobserved to scholars and policy makers.

The alternative hypothesis is that, in fact, the new view towards the innovative capacity of small firms emerged not because of measurement improvements, but because the economic and social environment actually changed in such a way as to shift the innovative advantage more towards smaller enterprises. This hypothesis would say that the conventional wisdom about the relative inability of small firms to innovate was
essentially correct – at least for a historical period of time. Rather, the new view of small firms as engines of innovative activity reflect changes in technology, globalization and other factors that have fundamentally altered the importance and process of innovation and technological change. As Jovanovic (2001, pp. 54-55) concludes, “The new economy is one in which technologies and products become obsolete at a much faster rate than a few decades ago...It is clear that we are entering the era of the young firm. The small firm will thus resume a role that, in its importance, is greater than it has been at any time in the last seventy years or so.”

Future research may sort out which of these two hypotheses carries more weight. However, one important conclusion will remain. Scholarship has clearly changed in its assessment of the role of small firms in the process of innovation and technological change from being mostly unimportant to carrying a central role.
References


Discussion Papers on Entrepreneurship, Growth and Public Policy


