INNOVATION. FIRM SIZE AND MARKET STRUCTURE: SCHUMPETERIAN HYPOTHESES AND SOME NEW THEMES

George Syrneonidis

TABLE OF CONTENTS

Introduction ................................. 36
Innovation, firm size and market structure ................................. 37
   Innovation and firm size .................................................. 37
   Innovation and market structure ........................................ 39
   Summary and an assessment ............................................... 41
Specific mechanisms in the Schumpeterian tradition ....................... 43
   The cost of R&D ....................................................... 43
   Economies of scale and scope in the production of innovations ...... 45
   Diversification .......................................................... 47
   Financial constraints .................................................... 48
   Appropriability conditions .............................................. 52
Other themes ......................................................... 54
   Market share dynamics and technological leadership ................ 55
   Demand characteristics .................................................. 56
   Stochastic patterns of firm innovation and growth .................. 57
   Summary ............................................................... 59
Concluding remarks and policy implications ............................... 59
Bibliography ......................................................... 65

The author is lecturer at the University of Essex. He would like to thank Sveinbjorn Blöndal and Hannes Suppanz for helpful comments on previous drafts.
INTRODUCTION

A central issue in industrial economics is how different market structures affect economic performance and social welfare. The exercise of monopoly power is known to result in static allocative inefficiency, although empirical estimates of the associated welfare loss vary widely. However, the static analysis of the social costs of monopoly (or oligopoly) fails to take into account the implications of alternative market structures for dynamic efficiency. A common argument in this context is that concentrated market structures may be favourable to technological progress, and hence economic growth and higher welfare. This implies that there may be a trade-off between short-run allocative gains from increased price competition and long-run welfare improvements from a higher rate of innovation under a more concentrated structure.

There are obviously important policy implications of the relationships between innovation, on the one hand, and market structure and firm size, on the other. As the role of innovation in promoting competitiveness and growth has become increasingly apparent, a number of policymakers have argued that antitrust laws should be relaxed, or should not be reinforced, as the short-run gains from more price competition are offset in the longer term by a slower rate of technological progress. The relevance of such views depends on the validity of the hypothesis of a positive effect of market power, concentration or firm size on innovation.

Moreover, the endogeneity of market structure has often been neglected in discussions of the welfare consequences of market power. An implication of this endogeneity, however, is that there may be a limit to what can be achieved by competition and industrial policies. In particular, it has been argued that there is, in any given industry, a range of market structures which are not sustainable, and that this may be part of the reason why industrial policies of support for specific firms or groups of firms through public procurement and subsidies have sometimes failed in the past. An adequate understanding of the economic mechanisms that determine market structure is therefore essential for designing the correct policies for R&D-intensive industries.

This paper is a critical survey of the large empirical literature on the links between innovation, market structure and firm size. The issue was brought into mainstream economics by Schumpeter, who, writing in 1942, argued that the large
firm operating in a concentrated market is the main engine of technological progress. A number of specific hypotheses as to why this may be the case have been advanced, most of which were already present in Schumpeter’s own work. Counter-arguments have also been suggested. The first part of the present paper reviews studies which have examined directly the impact of market structure or firm size on innovation. The second part then reviews studies which have focused on specific hypotheses as to why a positive effect may exist. The main questions examined in this relate to the optimal scale for R&D or innovation, the role of diversification, appropriability conditions and the role of financial constraints.

What is common to all studies in the Schumpeterian tradition reviewed in the first and second parts of the paper is the emphasis on a one-way direction of causality: from market structure and firm size to innovative performance. Recent work, however, suggests that both innovation and market structure must be seen as jointly determined endogenous variables. The third part of the present paper surveys empirical studies that have thrown some light on mechanisms relating innovation and market structure with an emphasis on the endogeneity of both these variables. The paper concludes with a discussion of policy implications in the light of the evidence presented both on the “Schumpeterian hypotheses” of the traditional literature and on the “new themes” of the recent literature on market structure and innovation.

INNOVATION, FIRM SIZE AND MARKET STRUCTURE

This part critically reviews the empirical literature on the Schumpeterian relationships between innovation, on the one hand, and firm size or market structure, on the other. It will be relatively brief, as the issues have been mostly examined in other surveys.¹

Innovation and firm size

Empirical studies of the relationship between innovation and firm size typically relate some measure of innovative activity to a measure of size, using a cross-section of firms from one or several industries.² Most early studies found little support for the Schumpeterian hypothesis of a more than proportionate effect of firm size on innovative activity. Some authors found an inverted-U relationship between firm size and R&D intensity, i.e. the ratio of R&D expenditure or personnel to size, or between firm size and the ratio of patents to size (for example, Scherer 1965a, 1965b). Others found a positive relationship up to a certain threshold, and no significant effect for larger firms. Moreover, industry studies revealed considerable differences in the tested relationships across industries. Many of these early studies were subject to serious limitations, such as the lack of adequate controls for
industry effects when using cross-industry samples of firms, or the low levels of statistical significance when using small samples of firms from a single industry. Despite these limitations, Kamien and Schwartz were able to conclude in their 1982 survey that, with the exception of the chemical sector, there was little support for the hypothesis of a more than proportionate effect of firm size either on R&D or on innovative output.

Others were less keen to totally dismiss Schumpeter, however. Soete (1979) found that R&D intensity increased with size in a number of sectors in the United States, although this was not always true for the largest firms, and decreased with size in others. Freeman (1982) and Rothwell and Zegveld (1982) pointed out that much of the empirical work on innovation and firm size had focused on firms which engaged in formal R&D and were relatively large. Small firms with less than 100 employees, on the other hand, typically do not perform any formal R&D and produce a less than proportionate number of innovations, which seems to be in line with the Schumpeterian view. It was acknowledged, however, that once the vast majority of firms that perform no formal R&D are excluded from the analysis, the relationship between size and R&D intensity is weak, non-existent or even negative.

The Schumpeterian hypothesis has received little support in recent studies using measures of innovative output such as innovation counts. Acs and Audretsch (1987, 1990) found that smaller firms (less than 500 employees) had a higher number of innovations per employee than larger firms (more than 500 employees) in 156 U.S. industries, while the reverse was true in 122 industries. They interpreted these results as evidence that the whole debate regarding which firm size is more conducive to innovation is pointless and the real issue is which industry characteristics favour either large or small innovators. In a later study, Acs and Audretsch (1991) found that innovations increased less than proportionately with firm size in each of 14 U.S. sectors, except, in some sectors, for the largest one or two firms, which had produced a more than proportionate number of innovations.

Similarly, Pavitt et al. (1987) found that innovation intensity in the United Kingdom was greater for large firms (more than 10,000 employees) and small firms (between 100 and 2,000 employees) and smaller for medium-sized and very small firms. Moreover, innovation intensity had increased steadily for firms with less than 500 employees over the 1956-83 period. The Pavitt et al. study also revealed the existence of important sectoral differences, which reflected sector-specific characteristics other than technological opportunity. The authors suggested that the differences related to appropriability conditions and the scope for diversification. These ideas should be seen in conjunction with earlier hypotheses advanced by Pavitt (1984) linking the scope for diversification and the typical size of innovating firms to technology and demand characteristics such as the extent to which innovative activity is science-based or driven by customer requirements, and
the extent to which appropriability is facilitated by the existence of large fixed costs and significant learning economies. Clearly, in trying to explain the observed sectoral differences in the size of innovating firms by reference to primary factors such as technology and demand, the studies by Pavitt (1984) and Pavitt et al. (1987) have identified a number of important elements for a theory of market structure and innovation in R&D-intensive industries.

Weak or non-existent links between size and R&D or innovation intensity were also confirmed in a number of recent US studies which have exploited richer data sets and/or have applied more sophisticated econometric techniques. Bound et al. (1984) found that both small and large US firms were more R&D-intensive than medium-sized firms. Scherer (1984) examined the effect of business unit size on business unit R&D intensity, using data collected by the US Federal Trade Commission. Despite the fact that for many industries there were only a few observations, Scherer ran simple regressions of R&D expenditure on sales for 196 industries. R&D intensity increased with sales in 20.4 per cent of the industries and decreased with sales in 8.2 per cent of the industries. In all the remaining cases the size coefficient was not statistically significant (it was positive in roughly half of the cases). Scherer also ran regressions with patent counts as the dependent variable. The number of patents was found to increase more than proportionately with sales in 11.3 per cent of the industries, less than proportionately in 15.3 per cent of the industries, while in the remaining cases no statistically significant departure from proportionality was detected (in slightly more than half of these cases the number of patents increased less than proportionately).

Finally, Cohen et al. (1987) used the FTC data to compare firm size to business unit size effects. Neither firm nor business unit size had a statistically significant effect on R&D intensity. Appropriability and technological opportunity at the industry level explained much of the variance in R&D intensity between firms. However, a threshold effect was identified: the size of the business unit, but not the overall size of the firm, had a positive and significant effect on the probability of conducting R&D. A subsequent study by Cohen and Klepper (1996) confirmed that any positive relationship between R&D and firm size is mainly driven by factors which are relevant at the business unit rather than at the firm level.

Innovation and market structure

Many studies in this area have used single-equation models to relate some measure of innovative activity to some concentration index. Early studies produced mixed results (see, for example, Scherer 1965b, 1967). Some found evidence of a positive effect of concentration on research intensity, while others failed to provide any support for the Schumpeterian hypothesis. On the whole, support for Schumpeter’s thesis seemed to be weaker when the number of patents rather than
RCD intensity was used as the dependent variable. Some studies have also attempted to analyze whether the RGD-concentration relationship depends on industry-specific characteristics, for example advertising intensity (Comanor, 1967) or the type of product (for example, Shrieves, 1978, distinguished between consumer and producer goods and between durables and non-durables).

A number of recent studies have found that the concentration-innovation link is weak or non-existent when controlling for industry effects. Scott (1984), who used the US FTC data at the business unit level, found that the effect of concentration on RGD disappeared when fixed sector and firm effects were included in the regressions. Similarly, Levin et al. (1985), who used the FTC data at the industry level, found an inverted-U relationship when including dummy variables at the sector level among the regressors, but no relationship when including a set of variables intended to capture inter-industry differences in appropriability and technological opportunity. For the United Kingdom, Geroski (1990) found evidence of a negative effect of concentration on innovation counts when controlling for industry effects. This study also explicitly recognized that market power cannot be fully captured by just one variable, namely concentration, so six different measures of market power were used, namely the extent of market penetration by entrants, the market share of imports, the relative number of small firms, the change in concentration, the market share of exiting firms and, finally, the concentration ratio (the effect of all six market power variables on innovation was in the same direction, but only in the case of the concentration ratio was the effect statistically significant).

While most of the literature has focused on the impact of market structure on innovative activity, some studies have explored the possibility of a two-way direction of causality. Econometric analyses of the relationship between innovative activity and market structure in the context of simultaneous-equation models have generally produced mixed results (see, for instance, Farber, 1981; Lunn, 1986; Levin and Reiss, 1988). An interesting aspect of the study by Lunn, who used the number of patents as his measure of innovative activity, was the distinction between process and product innovation. In particular, Lunn found a positive and statistically significant two-way relationship between concentration and process innovation and no effect either way between concentration and product innovation.

While the main body of the literature on the innovation-market power relationship has simply used concentration as a proxy for market power, a small number of recent studies have instead focused on the impact on innovation of changes in the international environment faced by innovating firms. In particular, these studies have examined how domestic innovators respond to increased import competition, which presumably erodes their market power. The results were, once again, mixed. Scherer and Huh (1992) analyzed the RGD spending reactions of US firms to increased high-technology import competition over 1971-87. They found that most of the variation in company RCD intensity change was unsystematic and that rising
import competition had a negative but statistically insignificant effect on R&D intensity in the short-run. Bertschek (1995), on the other hand, found a positive and statistically significant effect of the import penetration ratio on the production of both product and process innovations by German firms over 1984-88. However, both these studies have focused on the short-run. The analysis of long-run effects would require not only data over a long period, but also a different econometric specification, as some of the variables often included in short-run empirical models are endogenous in the longer term.

Summary and an assessment

A number of results seem to have emerged on the impact of firm size or market structure on innovative activity. First, very small firms do little R&D and produce a less than proportionate number of innovations. Second, above a certain threshold firm size, R&D seems to rise more or less proportionally, on the whole, with size, although there are variations of this pattern across industries, time periods and countries. Third, the evidence on the relationship between innovative output and size is inconclusive; most authors would probably agree that innovative output tends to rise less than proportionately with firm size, although other patterns have also been suggested for particular industries, periods or countries. Fourth, there is little evidence of a positive relationship between R&D intensity and concentration in general, although there may be circumstances where such a relationship exists. Fifth, there is even less evidence of a positive impact of concentration on innovative output. And finally, industry characteristics such as technological opportunity may explain much more of the variance in R&D intensity or innovation than market structure or firm size.

The studies reviewed in this part are subject to a number of limitations. First, there are problems with measuring innovative activity. Second, an assumption implicit or explicit in much of the literature is that firm size and market structure are exogenous. Now it is clear that past innovation affects firm growth and hence firm size (Scherer, 1992). If, as seems likely, innovative activity is influenced by unspecified factors which persist across time, firm size is correlated with these factors in any given year, so a regression of innovative activity on size will produce biased results. One response to this criticism has been to argue that the effect of innovative activity on firm size occurs with a multi-year lag, so the endogeneity problem is probably not very serious. The endogeneity of concentration is probably a more serious limitation. It has long been argued that the relationship between market structure and innovation is not a simple one-way causal relationship, because market structure is affected by innovative activity. Recognising the problem, some authors have used techniques that control for the endogeneity of concentration in the context of single-equation models, while others have estimated simultaneous-
equation systems in which both innovation and market structure are treated as endogenous.

Another difficulty arises from the need to control for industry-specific characteristics which affect innovation and may be correlated with concentration or firm size. It is important to control for industry effects to avoid obtaining biased results when using a sample covering many industries. It is, however, difficult to control properly for industry effects in studies of the size-innovation relationship, because, while most large firms are diversified and operate in more than one industries, data at the business unit level are rarely available. In studies of the impact of market structure on innovation, on the other hand, there is sometimes a somewhat different problem. A number of studies have controlled for technological opportunity, appropriability conditions, or both. However, while technological opportunity can be treated as exogenous, appropriability may itself reflect the degree of market power, and therefore the results may be difficult to interpret.8

A final problem arises from the fact that while one of the Schumpeterian hypotheses is that innovation is higher in the presence of market power, most of the literature has actually tested a different proposition, namely that innovation is higher in concentrated markets. The implicit assumption is that market power, i.e. the profit margin or mark-up, is greater in concentrated markets. This is not obvious in a framework where market structure is seen as endogenous. For example, there is evidence which supports the view that more intensive competition can result in higher concentration levels, as margins are squeezed and firms cannot cover their fixed costs unless their number falls through merger or exit (Sutton, 1991; Symeonidis, 1996, 1996c). These recent results and the inconclusiveness of the empirical literature on the profitability-market structure relationship (Schmalensee, 1989) suggest that concentration is at best an imperfect proxy for market power in cross-industry studies.

Apart from these problems, there are more fundamental limitations. First, with the exception of very few studies, this literature has not attempted to examine the specific mechanisms that presumably relate innovation to firm size or market structure. Second, no account has been taken of the fact that market power, market structure and innovation are all endogenously determined within a complex equilibrium system. Third, since it is clear that the size-innovation relationship tends to vary considerably across industries, it may be more useful to focus on particular factors that bear on the links between innovation, firm size and market structure, rather than trying to establish global patterns which may not exist. This is not to deny that this literature has been useful in establishing that there is probably no general advantage related to size, at least for firms above a moderate size threshold, and that innovation is not generally associated with a higher level of concentration.
SPECIFIC MECHANISMS IN THE SCHUMPETERIAN TRADITION

A variety of arguments have been advanced in support of the view that innovation is favoured by high concentration and large firm size (see Kamien and Schwartz, 1982; Rothwell and Zegveld, 1982, for comprehensive discussions). This part will review the empirical literature on the most important hypotheses as to why firm size or market power may have a positive effect on innovative activity. In particular, it has been argued that innovation increases more than proportionately with firm size because:

1. R&D projects typically involve large fixed costs, and these can only be covered if sales are sufficiently large.
2. There are economies of scale and scope in the production of innovations.
3. Large diversified firms are in a better position to exploit unforeseen innovations.
4. Large firms can undertake many projects at any one time and hence spread the risks of R&D.
5. Large firms have better access to external finance.

In addition, innovative activity may be higher in concentrated industries because:

6. Firms with greater market power are better able to finance R&D from own profits.
7. Firms with greater market power can more easily appropriate the returns from innovation and hence have better incentives to innovate.

What all these mechanisms have in common is the fact that they involve a one-way causality, from market structure to innovative activity.

The cost of R&D

According to one view, R&D projects typically involve large fixed costs, i.e. costs which are independent of the size of the market for the innovation. The disadvantage of small firms stems from the fact that, given the gross rate of return, their expected sales are not sufficiently large to allow them to cover these costs. Note the two implicit assumptions in this line of reasoning: the first is that firms mainly exploit their innovations through their own output, the second is that current firm size limits firm growth (Cohen and Klepper, 1996). There is, in fact, empirical evidence which supports these assumptions (see Cohen, 1995). The question then is whether it is actually the case that the costs of R&D projects are typically large. Another question is to what extent relatively large R&D projects can be jointly carried out so that the costs are spread among a number of smaller firms.
There is hardly any systematic evidence on the costs of R&D projects across industries, but case study evidence suggests that the notion of a typical cost of R&D is probably meaningless. In some industries such as aircraft the cost of a project may be so large that even large firms find it difficult to conduct independent R&D, while in other industries, such as instruments, R&D costs are relatively small.\(^1\) There may also be important variations from one project to another within the same industry (see, for example, Mansfield et al., 1971). Moreover, it is often the case that after the introduction of a major (and costly) innovation many minor improvements are made, and these may have a greater market value than the initial innovation. Regarding the evolution of R&D costs, Scherer (1991) has estimated that the average cost of significant innovations in the United States has risen at about 1 per cent per year between 1969 and 1986. However, in some industries, such as aircraft and pharmaceuticals, the increase in R&D costs has been much higher. According to Mowery and Rosenberg (1989), the same is true for information technology industries.

It may also be asked whether, in industries where the cost of R&D projects is high, the disadvantage of small firm size can be overcome through co-operative R&D agreements allowing for the spreading of R&D costs among a number of smaller firms. This issue has not been systematically examined in the literature on co-operative R&D. It is possible to think of reasons why joint small firm R&D may not be a substitute for large firm R&D; for instance, management of co-operative agreements may be difficult. There is also some evidence that the propensity to participate in R&D agreements or the fraction of joint R&D to total R&D at the firm level may be positively related to firm size (see Colombo, 1995; and Link and Bauer, 1989, but note that a limitation of both these studies is the fact that they have used small samples of relatively large firms). While some authors have suggested that many small R&D-intensive firms are engaged in joint R&D (see Dogson, 1993), this often involves co-operation with large firms rather than other small firms.

In conclusion, there are large differences in the cost of R&D projects across industries and possibly also within the same industry. It might be misleading to conclude that when these costs are high, large firms have an advantage in innovation, since firm size and market structure should be seen as endogenous: when the cost per R&D project is high, we simply expect innovative firms to be large, other things being equal.\(^2\) There is hardly any direct evidence on the extent to which the disadvantages of small firm size in industries with costly R&D projects can be overcome through joint R&D.
Innovation, firm size and market structure

Economies of scale and scope in the production of innovations

A number of studies have focused on the relationship between innovative inputs, output and firm size in an attempt to investigate the existence of scale and scope economies in the production of innovations. There are, in fact, two types of studies in this area: on the one hand, those that examine the relationship between the “productivity” of R&D (i.e. the ratio of innovative output to R&D expenditure or employment) and firm size; on the other, those that analyse the relationship between the productivity of R&D and expenditure or employment, or simply between innovative output and inputs. The expression “scale economies” may be somewhat misleading here, because the main issue is not whether innovation increases more or less than proportionately with R&D for a given firm size; it is rather whether the ratio of innovations to R&D increases or decreases with firm size.13

Various arguments have been proposed to support the hypothesis of economies of scale and scope in the production of innovations. These include the existence of positive spillovers between the various research projects within a firm, and the positive effect on researchers’ productivity of the interaction and complementarities within a large team. Counter-arguments have also been suggested. These usually refer to various organisational problems related with size, such as the loss of managerial control and the bureaucratisation of innovative activity.

Most of the early cross-industry studies have found no evidence of economies of scale in the production of innovations (for example, Scherer, 1965b), and some have even found evidence of diseconomies of scale (Schmookler, 1972) These results have been replicated by more recent studies. Scherer (1983), using the US FTC data at the business unit level, complemented by equally rich data on patents, examined the impact of R&D expenditure on patent counts in 124 industries. He found increasing returns in 15.3 per cent of the industries, decreasing returns in 25 per cent of the industries, while in the remaining cases there was no statistically significant departure from constant returns. It should also be noted that the FTC data cover only relatively large firms. Hausman et al. (1984) found evidence of decreasing returns in the patents-R&D relationship for a sample of 128 firms over the period 1968-74. Acs and Audretsch (1991) obtained similar results for the innovations-R&D relationship using a sample of 732 US firms in 14 sectors. These authors also estimated the model separately for each sector and found evidence of decreasing or constant returns in most sectors.14

It may be asked whether innovations introduced by large firms are, on the whole, more important or have higher market value than innovations by small firms. The evidence on this issue is mixed (see Gellman Research Associates, 1976; Acs and AUDRETSCH, 1988; Scherer, 1983) Since it has also been argued that major innovations are often inefficient initially, and the market value of small and gradual
improvements is greater than the value of the basic innovation, it is somewhat difficult to interpret this evidence anyway.

It should also be noted that the literature on R&D productivity is subject to practical problems of measurement of innovative inputs and output. A number of authors seem to agree that much of the R&D of small firms is informal and goes unreported in the statistics. It has also been argued (Schmookler, 1966) that large firms may have a lower propensity to patent their innovations than small firms. Finally, the nature of innovative activity may be different in large and small firms (for example, some of the research of large firms is basic research), so the respective rates of return on R&D may not be directly comparable.

On the whole, and to the extent that the measurement problems mentioned above are not so severe as to result in serious biases, the cross-industry studies suggest that innovative output per unit of R&D expenditure does not generally increase with firm size, and it may even decrease. There seems to be considerable variation across industries, however, although none of the cross-industry analyses has attempted to examine the factors that may account for this.

A recent study of the pharmaceutical industry by Henderson and Cockburn (1993) illustrates the advantages of detailed industry-level studies. The authors used highly disaggregated data at the research program level on patents and research expenditure for the discovery, as distinct from the development, of new drugs for ten major pharmaceutical companies over a period of 30 years. They regressed the number of patents on research expenditure, a proxy for firm size, diversification (i.e. the number of research areas) at the firm level, firm dummies and a time trend. The use of data on innovative inputs and output at the research program rather than the firm level allowed for interesting interpretations of the coefficients on firm size and diversification. The coefficient on firm size was positive and significant, implying that a given research program is more productive the larger the firm, while the relationship between patents and diversification, given firm and research program size, had an inverted-U shape. These results suggest, according to the authors, that there are economies of scale and scope up to a certain point in the research for discovery of new drugs.

Finally, it should be stressed that the question of the impact of firm size on R&D productivity must not be confused with the more general question of the “optimal” scale for innovation. In a recent study Cohen and Klepper (1996) have argued that the apparently negative relationship between R&D productivity and firm size need not imply a disadvantage of large firms in R&D. Suppose that the marginal productivity of R&D is diminishing, i.e. innovative output increases less than proportionately with R&D. Then R&D productivity falls as firm size increases simply because large firms do more R&D (in absolute terms) than small firms. Cohen and Klepper have argued that since R&D is a fixed cost, net returns to R&D are higher the larger the level of output over which this cost can be spread. In other words, net
returns to RGD are higher for larger firms, which is the main reason why these firms spend more on RCD than smaller firms, and are thus subject to lower RCD productivity. It follows that this lower productivity does not reflect any disadvantage of large firms. Clearly, the idea that cost spreading is an important factor in the R&D-size relationship is crucial to this whole argument. Cohen and Klepper have found support for their cost spreading model using the FTC data. The implication of their study is that it may be misleading to focus on RGD productivity alone, and that the existence of indivisibilities in RGD should also be taken into account when drawing inferences concerning the "optimal" scale for innovation.

**Diversification**

The focus on diversification has been motivated by at least three different hypotheses associating diversification with the innovation-firm size relationship. The first, already suggested by our discussion of the Henderson and Cockburn study above, is that large diversified firms can benefit from positive spillovers between the various research programs they undertake; this is then an argument about the existence of economies of scope. The second hypothesis is that large diversified firms are in a better position to exploit unforeseen innovations. The idea here is that the results of research, especially basic research, are often unpredictable, and a diversified firm has more opportunities to exploit the new knowledge outside the product line for which the original research was done. The third hypothesis is that large firms can undertake many projects at any one time and hence spread the risks of RGD. All three hypotheses imply that RGD intensity is higher in large than in small firms because large firms can diversify into many research areas or markets.

It is difficult to distinguish between these various mechanisms empirically, so most studies in this area have simply examined the relationship between diversification and innovative activity. The results have been inconclusive (see Kamien and Schwartz, 1982; Cohen and Levin, 1989). Moreover, a serious problem in this literature is that the measure of diversification may partly reflect pure industry or product line effects: if a firm whose primary product line requires a high (low) level of R&D intensity diversifies into unrelated product lines, its RCD intensity will most probably fall (rise) for simple statistical reasons. As, in addition, many studies have used rather crude measures of diversification, such as the number of industries in which a firm operates, it is questionable whether much can be learned from most of the literature in this area.

A number of recent studies have tried to tackle these problems. Henderson and Cockburn, in the study discussed above, have used disaggregated data at the research programme level. Another approach has been to distinguish between diversification in technologically related and in technologically unrelated industries. Scott (1993) divided his sample of 355 firms taken from the FTC Line of Business...
data into "purposively" and "non purposively" diversified. The former group contained those firms which had diversified into related industry categories, the latter included firms diversified into unrelated industries and undiversified firms. Scott found that, although purposively diversified firms may have higher or lower R&D intensities in particular industries than firms that have not purposively diversified, there is a significant overall tendency for the R&D intensities of the purposively diversified firms to be higher. It may be tempting to interpret this result as evidence consistent with those hypotheses emphasising the role of diversification in exploiting complementarities, although, on the other hand, it cannot be said to refute the risk spreading hypothesis. There is, however, one important caveat: diversification, especially in related industries, may itself be the result of accumulated R&D capital stock in the primary industry; or, as Pavitt (1984) has argued, diversification in related industries may be favoured in science-based high-technology sectors.

The hypothesis that large firm size may be conducive to innovation because of risk spreading has been tested in an entirely different framework in a study by Wedig (1990). He found strong evidence that R&D resulted in a considerable systematic risk premium at the firm level, but only weak evidence that concentration and firm size reduced this risk premium (the relevant regression coefficients had the correct sign but were more often statistically insignificant than significant).

In conclusion, the evidence on the hypothesis that diversification favours innovative activity is inconclusive. Because of measurement and endogeneity problems, it will be hard to obtain definitive results in this area. There are signs that there may be a positive correlation between diversification into technologically related industries and R&D, but it is not clear whether this reflects a causal relationship and, if so, what is the direction of causality. There may also be considerable variation in the relationship across industries.

Financial constraints

Some of the most frequently invoked arguments to support the Schumpeterian hypotheses of a positive effect on innovative activity of market power, on the one hand, and firm size, on the other, start from the observation that capital markets are imperfect. R&D activities, on the other hand, involve a larger or smaller amount of sunk costs, i.e., costs incurred before production takes place and independent of the gross returns from the innovation. In addition, R&D activities involve a high level of risk. It seems then likely that the availability of internal and external finance will impose a major constraint upon plans to undertake or expand R&D activities. This is the reason why firm size or market power may matter. Two different hypotheses are typically advanced in this context. It is argued, first, that R&D intensity is higher in concentrated industries because firms with greater market power are better able
to finance R&D from own profits; and, second, that R&D intensity is higher in large than in small firms because large firms have better access to external finance.

The preponderant role of internal funds for the financing of innovation is well documented, although it is hard to say to what extent this is due to the difficulty of obtaining external funds for risky projects rather than to the firms' own reluctance to raise debt for such projects. In a study of investment, R&D investment and sources of finance in a large panel of 1,678 US firms with significant R&D programs over the period 1976-87, Hall (1992) found strong evidence that more R&D-intensive firms had relatively less debt in their capital structure than less R&D-intensive ones. In a comparative study of 130 high-technology small US and UK firms, Oakey et al. (1988) found that more than two thirds of the firms relied on own profits as the main source of investment finance. Finally, using data on some 350 R&D performing UK and US firms, Board et al. (1993) found a negative relationship between R&D intensity and gearing (i.e., the proportion of debt in a firm's capital structure) in both countries.

The hypothesis that firms with greater market power are more innovative because they are better able to finance R&D from own profits has provided the rationale for examining the relationship between cash flow and innovative activity. A positive relationship would be taken to imply that liquidity constraints do affect R&D and that firms which are less constrained spend more on R&D, everything else being equal. From this, and the implicit assumption that cash flow is increasing in the degree of market power of the firm, a confirmation of the initial hypothesis that the existence of financial constraints favours innovation by firms with market power would follow.

The literature on innovation and cash flow has, on the whole, provided some support for the hypothesis of a positive relationship between these two variables. Results from early studies have been mixed (see Scherer, 1965b, Grabowski, 1968; Elliott, 1971). Kamien and Schwartz concluded in their 1982 survey that the evidence of a positive effect of either liquidity or profitability on innovation was weak, but they also pointed out that liquidity or profitability may be "threshold factors," i.e., necessary in some degree for R&D activity, but having little effect on R&D intensity above the threshold.

Recent studies have been, on the whole, rather more favourable to the hypothesis of a positive innovation-cash flow relationship. Hall (1992), in the study mentioned above, found a positive and statistically significant effect of the ratio of cash flow to capital stock on the ratio of R&D investment to capital stock. Switzer (1984), found weak evidence of a positive relationship between R&D expenditure and internal financing for a sample of 125 firms (the relevant coefficient was significant at the 20 percent level). Himmelberg and Petersen (1994) found a highly significant positive effect of cash flow on R&D investment in a sample of 179 firms from four R&D-intensive 2-digit sectors over the period 1983-87. The authors also mentioned...
that a comparison of these results with their own preliminary results for larger firms and with Hall (1992) suggested that larger firms were financially constrained as well, although to a lesser degree than smaller firms: the cash flow coefficient was lower but still significant for larger firms. Not all recent studies have found a positive innovation-cash flow relationship, however, and a few authors have even reported negative coefficients on cash flow. Antonelli (1989) obtained an inverted-U relationship in his study of 86 R&D performing Italian firms over three years, which he interpreted as evidence that low-profit “traditional” firms were involved in heavy R&D investments in an attempt to adapt to the more competitive environment of the 1980s, while for “modern” profitable firms the Schumpeterian view of cash flow constraining innovation was confirmed.

The main limitation of this literature is the interpretation of the cash flow variable. For example, although many authors would interpret a positive coefficient as evidence of financial constraints on R&D activity, it might well reflect either expected future profitability or the profitability of past R&D investment. Cohen (1995) has suggested that distinguishing between the cash flow of the firm as a whole and that of the business unit may help resolve these ambiguities, because firm cash flow is less likely than business unit cash flow to reflect past R&D productivity or expected future profitability. This has not yet been done in the literature, however.

The second hypothesis regarding financial constraints is that R&D intensity is higher in larger firms because these firms have better access to external finance. Unfortunately, there is only limited direct evidence on this hypothesis. While a number of studies have examined various aspects of the financing of R&D either in general or for small firms in particular, systematic comparisons between small and large firms have rarely been made.

Several authors have suggested that the existence of financial constraints may restrict innovation by small firms. Based on results from a large survey of “barriers to innovation” in 9 EC countries in the late 1970s, Piatier (1984) concluded that the greatest barriers were related to external finance and that smaller firms had greater difficulty in obtaining funds for R&D than larger firms and paid higher interest rates. Rothwell and Zegveld (1982) pointed out that during the 1970s the bulk of government subsidies for R&D in most industrial countries had gone to financing projects in large firms. According to Prakke (1988), conditions improved during the 1980s, as a considerable growth in the supply of venture capital took place, first in the United States, then in Europe, which improved the chances of new small high-technology firms to obtain external financing. In Europe, governments have been heavily involved in the supply of venture capital: 21 per cent of total venture capital funds in 1985 were public funds. This has been part of a more general shift in public R&D policy in favour of small firms and co-operative research (Oakey et al., 1988; see also...
Innovation. firm size and market structure

Metcalfé, 1994, for the United Kingdom; Malerba, 1993, for Italy; Bernard and Quéré, 1994, for France).

There is some evidence, however, that problems remain. Himmelberg and Petersen, in the study discussed above, have found that the impact of cash flow on R&D appears to be smaller for larger firms, which they have interpreted as evidence that larger firms are financially constrained to a lesser degree than smaller firms. Moore (1994), using data from a survey of some 2,000 UK small firms, found that, within the class of small high-technology firms, smaller firms were more likely to be constrained during early growth and subsequent maturity than larger ones, other things being equal. Conditions are likely to vary across countries. For example, in a comparison of US and UK sources of finance for “new technology-based firms”, Mason and Harrison (1994) have argued that UK firms are likely to experience greater financial constraints because informal venture capital (i.e., funds invested by wealthy individuals in high-technology firms) is much smaller in the United Kingdom than in the United States.

There is also evidence from the literature on small firm finance that interest rates on bank loans tend to be negatively related to firm size in many countries. This is to a large extent due to the fact that small firms are correctly evaluated by the banks as more risky, on average, than large firms, since they have a higher probability of failure. However, it may also be partly due to fixed costs in assessing applications for finance, the bargaining power exerted by large buyers, or the use by bank managers of performance criteria unfavourable to smaller firms. One study which has attempted to assess the extent to which differences in loan rates reflect differences in risk is the study by Bardos (1990). Using data from French surveys, this author classified some 8,000 firms into seven groups according to an index of risk constructed in a somewhat ad hoc manner on the basis of various economic and financial performance criteria. The sample was not confined to innovative firms and contained both large and small firms. Bardos concluded that the negative relationship between firm size and the interest rate persisted even when controlling for the level of risk.

The extent to which small firms face difficulties in raising finance for start-up or growth has also been examined in a number of UK studies (see Hall, 1989; Mayer, 1992; and DTI, 1991, for reviews of this literature). The main findings can be summarised as follows. First, small firms are required to pay a higher interest rate and offer a higher level of security than larger firms, although it is difficult to know to what extent this is a result of market failure rather than risk differentials. Second, there are many more financial avenues open to large firms than to small ones and there is also evidence that a considerable fraction of the overall supply of venture capital has not actually been directed towards small innovative firms during the 1980s. Third, while there is no evidence of a general shortage of funds for small firms, it seems that for certain classes of small firms, namely fast-growing, innova-
tive or starting-up firms, there are sometimes difficulties in raising finance, although again this does not necessarily imply that the financial markets are not operating efficiently in these cases.

To sum up, there appears to be some evidence that the existence of financial constraints may restrict innovation by small firms and firms with little market power. This evidence is consistent with the Schumpeterian hypotheses on the role of financial constraints. A note of caution is necessary here, however, as some of the results, namely those regarding the innovation-cash flow relationship, are open to alternative interpretations, and others, in particular those on the availability of external finance for small innovative firms, are not conclusive in the absence of systematic comparisons with large innovative firms. It should also be emphasized that the existence of financial constraints does not necessarily imply that financial markets are operating inefficiently.

**Appropriability conditions**

According to another hypothesis, R&D intensity is higher in concentrated industries because firms with greater market power are better able to appropriate the returns from innovation and hence have better incentives to innovate. A number of issues need to be examined in this context. First, the role of the patent system in ensuring appropriability and the extent to which this system may be favourable to firms with greater market power. Second, the role of other mechanisms ensuring appropriability, such as secrecy, lead time, investment in marketing and customer service, learning by doing, and control of distribution channels, and the extent to which firms with market power are in a better position to benefit from such mechanisms. Third, the effect of appropriability conditions on the incentives to perform R&D.

There is strong evidence that industries differ considerably with respect to the role of patents in ensuring appropriability and that in many industries patents are seldom regarded as an essential incentive for innovation (Taylor and Silberston, 1973; Mansfield et al., 1981; Mansfield, 1986; Levin et al., 1987). The most comprehensive study of appropriability conditions is the study by Levin et al. The authors conducted a survey of R&D executives in 130 industries. Patents were thought to be highly effective as a means of preventing duplication in only five industries, and quite effective in another 20 industries. For process innovations, only three industries regarded patents as highly or quite effective. The overall sample means on a seven-point scale of effectiveness were 3.52 for process patents and 4.33 for product patents.

Despite the fact that patents do not seem to be regarded as a necessary incentive for innovation in most industries, it may be asked whether the patent system tends to be more favourable to firms with greater market power. Very little
direct evidence on this question is available. Scherer (1983) found that industry concentration had no impact on the number of patents per unit of R&D expenditure at the business unit level. On the other hand, Mansfield (1986), using a small sample of US firms, found a positive and statistically significant correlation between firm size and the percentage of patentable innovations which were patented in three sectors (pharmaceuticals, chemicals, petroleum) where patents seemed to be most important. However, these results are not very informative as to whether the patent system does or does not favour firms with higher market power, since many effects are involved in these relationships. An important issue in this context is whether firm size or market power is related to the ability of patent holders to enforce their rights against alleged infringers. This may depend on specific characteristics of the patent system in each country. For the United States, Scherer (1991) analysed 148 patent case decisions between 1983 and 1988 and found no evidence that smaller firms were at a disadvantage relative to larger firms.

The main reason for the limited effectiveness of patents seems to be that competitors can legally “invent around” patents (Levin et al., 1987). Other reasons include possibly stringent legal requirements for proof that a patent is valid or that it is being infringed, and the disclosure of information through patent documents. In many industries, however, firms find other means of appropriation to be quite effective. Investment in complementary sales and service efforts was regarded as highly effective in 80 per cent of the 130 industries in the Levin et al. sample. Other mechanisms mentioned in this study included secrecy, imitation costs and lags, and the ability to move down the learning curve. The effectiveness of each of these mechanisms was shown to vary across industries. However, the relationship between appropriability conditions and other industry-specific characteristics such as market structure was not examined.

Levin et al. have also argued that a shortcoming of their study is that it has focused on relatively large firms and may have thus understated the role of patents in ensuring appropriation of returns to R&D. The idea here is that it may be unfeasible for a small firm to invest in complementary sales or customer service or to move down a learning curve. This criticism also applies to Mansfield (1986) and Mansfield et al. (1981). Both these studies examined whether patent protection is more important for smaller firms than for larger ones, and found little evidence to support this view. On the question of the links between firm size or market power and the effective use of appropriation mechanisms such as secrecy, lead time and investment in marketing and customer service, there is hardly any systematic evidence.

Finally, it should be noted that while the literature on the Schumpeterian hypotheses has typically taken for granted that appropriability has a positive effect on R&D incentives, the scant empirical evidence on this issue is mixed (see...
Moreover, little is known on the extent to which the effect depends on industry-specific characteristics.

In conclusion, we still know very little on the factors that account for the observed inter-industry variation in appropriability conditions. It remains an open question whether firms with greater market power are in a better position to benefit from patents or other mechanisms of appropriating returns from innovation.

OTHER THEMES

This part will discuss recent empirical work which has thrown some light on a number of mechanisms relating innovation to market structure with an emphasis on the endogeneity of both these variables. Most of these studies are based on theoretical models of market structure and the innovative process and all of them go beyond the Schumpeterian hypotheses by emphasising primary factors like technology, the characteristics of demand, the institutional framework, strategic interaction and chance. Three themes will be explored:

1. First mover advantages, such as learning by doing, and disadvantages, such as organisational inertia, and their implications for the evolution of market shares and technological leadership.

2. Demand characteristics, such as the degree of product differentiation, and their implications for the trade-off between product enhancement and product proliferation and for market structure.

3. Stochastic patterns of firm innovation and growth and their implications for the evolution of market structure over the industry life cycle.

An advantage of these recent studies over the cross-industry econometric literature on the impact of innovative activity on market structure is the fact that they focus on specific mechanisms relating the two variables.20,21

A general underlying theme in much of the recent literature on market conduct and structure is the notion that equilibrium outcomes in oligopolistic industries depend on a number of factors, some of which are either not observable or not systematic. In fact, the empirical study of industrial structure and business strategy provides numerous examples of how the evolution of an industry can be affected by accident and personality.22 This dependence on unobservable or non systematic influences has important implications for the analysis of markets, because it means that, for any given set of values of the measurable and systematic exogenous factors, there is a multiplicity of possible outcomes. While for some problems the best approach may still be to regard these influences as random noise, for others a different approach may be more useful. Sutton (1991, 1995) has advocated the use of a "bounds" approach for the study of market structure. This involves placing bounds on the space of outcomes and thus excluding certain outcomes as either
impossible or unlikely. For example, in Sutton (1991) a lower bound to concentration is proposed, which corresponds to a situation where each of a number of symmetric single-product single-plant firms operates at a level of output equal to the minimum efficient scale level. While the equilibrium level of concentration cannot be lower than this bound, it can be higher to the extent that an industry comprises asymmetric multi-plant or multi-product firms or firms operating above the minimum efficient scale level.

**Market share dynamics and technological leadership**

A number of studies have focused on factors determining the evolution of technological leadership and market shares in RCD-intensive industries. Gruber (1992a, 1992b, 1995) has focused on firm-specific learning by doing as a mechanism which generates stability of market share patterns over a sequence of product innovations. In particular, he has examined the relevance of learning by doing in the production of two different types of semiconductor memory chips. Firm-specific learning by doing is very important in the production of one of these types, and this helps explain the regularity of market share patterns over successive generations of the product. Learning by doing is less relevant in the case of the other type of chip, however, and this may be part of the reason why “leapfrogging” has occurred in this market.

Learning by doing may generate a first mover advantage in products where quality improvement is a relatively continuous and largely predictable process. Other factors, however, may favour the leapfrogging of a technological leader, especially when improvements are less predictable and technological discontinuities, i.e. changes from one group of products or processes to another, occur. Foster (1986) has looked at a large number of cases of leapfrogging across a wide range of industries and has identified a number of potential incumbent firm disadvantages in innovation. Due to the uncertainty involved in all innovative activity, a technological leader may have difficulty realising that he has reached the region of diminishing returns to the effort put into improving an existing product or process. In addition, a technological leader may be tempted to stick to an existing technology too long because of organisational inertia or in order to maximise the returns on his sunk investment. As a result, the leader is likely to be leapfrogged when a significant change in technology takes place.

Similar ideas have been advanced in a recent study by Swann and Gill (1993a, 1993b). Their main hypothesis is that technological change which is consistent with a widely accepted vision of the future of a technology will tend to favour incumbent firms, because the existence of static scale economies and dynamic learning economies, and hence be concentrating, at least in the short to medium term. On the other hand, technological change which conflicts with widely held
expectations about the future of a technology and hence is disruptive to existing organisational structures (for example, a radical redefinition of the product or the introduction of a new process) will have a deconcentrating effect, as entrants will take opportunities that incumbent firms will fail to exploit because of organisational inertia. Swann and Gill examined evidence from case studies of five high-technology markets, namely microprocessors, memory chips, standard logic chips, PC software and biotechnology, which seemed, on the whole, to be consistent with their hypothesis. In the PC software case, for instance, rapid incremental change of a predictable sort, namely the sequence of upgrades to a particular software package, has tended to be concentrating, while radical innovations which have created new software categories have tended to be deconcentrating.

These recent studies have taken a first step towards a better understanding of the factors determining the evolution of technological leadership and the degree of turbulence in R&D-intensive industries. The main focus has so far been on the characteristics of technology itself, such as the degree of continuity of technology and the extent of learning economies in innovation, and on the organisational structure of the firm. Future work may analyse the role of additional factors, such as appropriability conditions, and also examine more closely the link between turbulence and concentration in R&D-intensive industries, since one can think of situations where high turbulence is not deconcentrating in the medium to long term or persistence of technological leadership is not concentrating.

**Demand characteristics**

The second theme explored in the recent literature relates to the implications of demand characteristics, such as the degree of product differentiation, for the trade-off between product enhancement and product proliferation and for market structure. Sutton (1996) identified two key industry-specific exogenous parameters that are important determinants of both R&D intensity and concentration. The first is the elasticity of the R&D cost function, a technological characteristic. Broadly interpreted, this is a measure of the returns from improving the technical performance or quality of an existing product. The second is the degree of horizontal product differentiation, a demand characteristic. This is a measure of the degree to which demand is fragmented due to the existence of a variety of customer tastes or requirements; it therefore indirectly reflects the returns from introducing a new product. Sutton argues that an important trade-off in high technological opportunity industries is between spending in R&D to enhance the quality of an existing product and spending to develop a new product. The former strategy will result in a relatively high level of concentration, as incumbent firms get caught in a process of R&D escalation leading to a rise in R&D expenditures. This will cause a rise in concentration, since firms cannot cover these expenditures unless their number falls and each firm produces more output. The latter strategy will result in a
relatively low level of concentration, as the possibility of developing new varieties that match the needs of particular buyers leads to a proliferation of varieties and attracts new entry. Both strategies are associated with a high level of R&D intensity, which is itself endogenously determined. This whole argument is cast by Sutton within the “bounds” approach to market structure discussed above.

The empirical analysis of these ideas in Sutton (1996) involved both the formulation of theoretical predictions that could be tested across industries and the analysis of case studies. The main prediction related to the relationship between the degree of product differentiation and market structure in R&D-intensive industries: the lower bound to concentration is expected to decrease as product differentiability rises and product proliferation increasingly dominates product enhancement. On the other hand, there is no reason to expect any relationship between product differentiation and market structure in low-R&D industries. Sutton tested this hypothesis using the FTC Line of Business data on R&D intensity at the industry level and matched data on horizontal product differentiation (proxied as the number of 7-digit product lines within each industry. He found that the lower bound to concentration decreased with product differentiability in the group of industries with R&D to sales ratio higher than 4 per cent, while there was no relationship between the two variables in the control group of industries with R&D to sales ratio lower than 4 per cent.

There may also be a number of additional implications of these ideas. For instance, it seems likely that demand characteristics play a less important role in industries with process R&D than in industries with product R&D. One might then expect a stronger correlation between R&D intensity and concentration in the former group of industries than in the latter. Finally, one qualification should be noted. Even though R&D outlays are endogenously determined as a result of firms’ strategic choices, there is an exogenous component as well, mainly because of indivisibilities in R&D projects (see the part on R&D costs above). The average level of costs per R&D project in each industry is likely to be an important factor in explaining inter-industry differences in market structure, in addition to technological opportunity and demand characteristics.

Stochastic patterns of firm innovation and growth

The third theme examined in recent work relates to the implications of technological change and stochastic patterns of firm innovation and growth for the evolution of market structure over the industry life cycle. The starting point was a number of empirical regularities regarding the industry life cycle that seem to hold across a considerable number of high technological opportunity industries (see, for example, Gort and Klepper, 1982; Klepper and Graddy, 1990; Utterbach and Suarez, 1993). A first set of regularities relates to the evolution of market structure in technologically
progressive industries. When such an industry is new, there is a lot of entry and turbulence, but subsequently entry slows and exit starts to prevail despite continued market growth, thus leading to a “shake-out” in the number of firms; eventually, market structure stabilizes. A second set of regularities pertains to technological change in industries with rich opportunities for both product and process R&D. Initially, many varieties of the industry’s main product are offered and the rate of product innovation is high, but subsequently increasing effort is devoted to process relative to product innovation and the diversity of competing versions of the product declines.

There are a number of studies that try to explain these and related patterns. According to one story, during the early stages of industry development there is experimentation with various product designs, as uncertainty is high and demand is fragmented across a number of variants. Firms are relatively small and entry is easy, as there are no significant advantages to incumbency. At some point in time, however, a dominant design or a narrow class of designs emerges and finally becomes established. Only firms that can adopt the dominant design stay in the industry. In addition, competition shifts from design to price, process innovation increasingly dominates product innovation, and economies of scale and learning become important. This induces further exit and also puts potential entrants at a disadvantage. After the shake-out, structure stabilizes. It has been argued that this story describes the evolution of the US automobile industry and other industries where consumer preferences are relatively homogeneous.

Klepper (1996) has proposed another model to explain the observed regularities in industry evolution. In contrast with the dominant design theory, there are no exogenous changes in technology, i.e. firms face opportunities for product and process innovation which do not change over time. The basic mechanism shaping industry evolution is as follows. After an initial stage with a lot of entry, exit and product innovation, all principally driven by random firm differences in innovative capabilities and the timing of entry, consistently successful firms gradually get bigger. As returns to process RGD are assumed to be positively related to firm size, larger incumbents increase their process RGD and hence gradually magnify their cost advantage over potential entrants. So eventually entry stops, the number of firms falls, and the rate of product RGD declines. Klepper and Simons (1993) argued that this model explained the evolution of four US industries (automobiles, tyres, televisions and penicillin) better than the dominant design theory or an alternative story by Jovanovic and MacDonald.24

Although industry life cycle theories seem to explain certain patterns which have been observed across a considerable range of industries, they have also been criticised on various grounds. It has been argued that the dominant design story is not valid in the case of industries where demand is fragmented due to the existence of diverse consumer needs. The Klepper model, on the other hand, may not be
relevant for industries where the scope for process innovation or economies of scale is limited. More generally, these models may oversimplify the evolutionary process by implicitly suggesting that there are no discontinuities in the technology of the industry during the later stages of industry development, and hence incumbents can safely maintain their dominant position. But while the generality of the particular mechanisms and evolutionary paths proposed by these studies is questionable, the focus on a number of factors that have not been much analysed in the literature, namely technological events, predictable or unpredictable, and random differences between firms in innovation and growth, is certainly welcome. It is now clear that a full understanding of the interaction between innovation and market structure involves examining the ways in which chance and technological change affect the evolution of market structure and innovative activity.

Summary

A number of recent studies have emphasized the endogeneity of both market structure and innovation and have examined certain mechanisms relating the two variables. One key result is that certain characteristics of technology, such as the degree of continuity and predictability of technology and the extent of learning economies in innovation, are important determinants of the evolution of technological leadership and the degree of turbulence in R&D-intensive industries. Another key result is that demand characteristics such as the degree of product differentiation affect the extent to which high R&D intensity will be associated with a high level of market concentration. A third result is that technological events and random differences between firms in innovation and growth play an important role in shaping the evolution of market structure and innovative activity in technologically progressive industries.

CONCLUDING REMARKS AND POLICY IMPLICATIONS

The present literature survey suggests that, on the whole, there is little empirical support for the view that large firm size or high concentration are factors generally associated with a higher level of innovative activity. Moreover, even though there are circumstances where a positive association exists, this by no means implies the existence of a causal relationship. There is then no general trade-off between competition policy and technical progress. However, in some R&D-intensive industries a high level of concentration may be inevitable, mainly because of the existence of indivisibilities in R&D and the fact that R&D projects may involve large fixed costs. It follows that merger and antitrust policies should be used with a view to the specific circumstances of each industry.

Competition and industrial policies can influence both market structure and innovation, but they must be used with care. For instance, industrial policy aimed at
promoting co-operative research may in some cases allow smaller firms to share the
cost of an RGD project. However, little is known on the effectiveness of such a
policy, or on its impact on the total level of RGD or the intensity of price competi-
tion. Policy with respect to the financing of R&D can increase the availability of
external finance to financially constrained small innovative firms. However, such a
policy must be carefully designed to avoid creating distortions, since the existence
of financial constraints need not imply that financial markets operate inefficiently.
These are important issues, given the recent shift in public R&D policy in favour of
small firms and co-operative RGD.

Finally, the policy implications of the endogeneity of market structure and firm
size should also be emphasized. Recent work in industrial economics suggests that
market structure and R&D intensity are jointly determined by technology, demand
characteristics, the institutional framework, strategic interaction and chance. A
major policy implication of the recent literature is that there are limits to what can
be achieved through competition and industrial policies, since there is only a
certain range of sustainable market structures in any given industry.
NOTES


2. Measures of innovative or technological activity can be classified as measures of either innovative inputs or output. The former include R&D expenditure and personnel involved in R&D, the latter include the number of patents and the number of significant innovations.

3. Pavit et al. found that smaller firms were important innovators in machinery, instruments and construction, larger firms in food products, chemicals, metals, electrical engineering and aerospace. Now both machinery and instruments, on the one hand, and chemicals and electrical engineering, on the other, are sectors of high technological opportunity. However, in chemicals and electrical engineering, firms typically produce innovations across a range of technologically related 3-digit industries within their principal 2-digit sector and, in addition, technological imitation by firms producing outside these sectors is difficult. By contrast, in machinery and instruments, firms are much more specialised technologically within their principal 3-digit industry and imitation by firms producing outside these sectors is relatively easy.

4. Since large firms are typically diversified, with business units in several industries, this data set had two advantages. First, it was possible to control properly for industry effects. Second, the effects on R&D intensity of both business unit size and firm size could be examined. In fact, while some of the arguments advanced to justify the hypothesis of a positive effect of firm size on innovative activity refer to the overall size of the firm (e.g. better access to external finance), others refer to the business unit (e.g. fixed R&D costs), so this was an opportunity to examine the validity of some of these arguments. A potential shortcoming of the FTC data set was that only relatively large firms were covered.

5. Both Geroski and Levin et al. used two-stage least squares to control for the potential endogeneity of concentration.

6. None of the measures of innovative activity is entirely satisfactory. One problem with R&D data is that some R&D activity takes place outside a firm’s formal R&D operation. Since many small firms do not have R&D departments, even though informal R&D is carried out within the firm, the amount of small firm R&D may be underestimated in some data sets (Kleinknecht, 1987, Kleinknecht; and Verspagen, 1989). The main
problems with patent counts are that patents differ greatly in their economic value and that the propensity to patent varies significantly across industries. Finally, attempts to count the number of significant innovations are subject to some arbitrariness and possible biases in the evaluation procedure.

7. The evidence on the average time between the beginning of a research project and the beginning of the associated revenue stream is somewhat mixed, with estimates ranging from 2 to 4 years. On the other hand, the average time between the completion of a project and an effect on profits is less than a year (see Mansfield et al., 1971, Pakes and Schankerman, 1984). This suggests that the endogeneity problem may be more serious in studies using measures of innovative output than in studies using measures of innovative inputs.

8. One of the arguments most frequently put forward to rationalise the hypothesis of a positive effect of market power on innovative activity is that firms with market power can more easily appropriate the returns from their innovations. Suppose now that the inclusion of a measure of appropriability among the explanatory variables in a regression of innovative activity on concentration reduces substantially the coefficient and the t-statistic on concentration. It is not clear how such a result should be interpreted. For instance, this might indicate that market power matters for innovation and the effect works through appropriability rather than the ability to finance R&D. Alternatively, it might indicate that concentration does not matter, as a correlation between concentration and appropriability does not necessarily imply a causal relationship between these variables.

9. Many counterarguments have also been suggested. Some of the most important, such as the notion that market power caused by the absence of competitive pressures may lead to inertia and hence to lower innovative activity, are difficult to examine empirically.

10. In fact, the instrument engineering sector is the typical example of a sector with high R&D intensity and low levels of concentration.

11. In fact, most of the literature on small firm cooperative R&D has focused on large firm-small firm linkages. More generally, some authors have argued that the pattern towards increased R&D cooperation among firms, including cooperation between large and small firms, has minimised the relevance of the whole debate on the relationship between firm size and innovation. These authors emphasise the existence of complementarities in R&D between firms of different sizes (see, for example, Teece, 1992; Rothwell, 1989).

12. The cost of an R&D project in any given industry is itself partly endogenous, since it is affected by the past, and possibly also the current, strategic choices of the firms. It is probably correct to assume, however, that at any given point in time this cost cannot be lower than a certain minimum level, which is taken as exogenous by the firms.

13. On the other hand, a number of studies (not reviewed here) have not in fact examined the impact of firm size on R&D productivity. For example, some authors control for firm size in regressions of innovative output on R&D outlays, or regress innovative output on R&D intensity. They therefore obtain results on the impact of R&D on innovative output when firm size is kept constant, which is not the key issue in this line of research.
In particular, Acs and Audretsch found decreasing returns in chemicals, computers and office equipment, electrical equipment, motor vehicles, and instruments; increasing returns in petroleum; and constant returns in pharmaceuticals, foods and tobacco, and an aggregation of all industries with \( R&d \) to sales ratio lower than 1 per cent. However, in some of the sectors exhibiting decreasing returns the single largest firm was an exception to this pattern. This may indicate that scale economies exist at very high levels of \( R&D \) expenditure in some sectors, but it may also be due to the exceptional \( R&D \) productivity of a small number of leading industrial firms.

In particular, they found that firm size explains an average of 15 per cent of the variance in business unit \( R&D \), while business unit size explains an average of 65 per cent of this variance. Hence \( R&D \) seems to be tied to business unit size rather than firm size, which is consistent with the cost spreading hypothesis. They also found that the relationship between \( R&D \) and business unit size is weaker in those industries and types of innovations where cost spreading matters less: industries where the expected growth rate due to innovation is high or where innovations can be sold in disembodied form.

One study which has focused on basic research, as distinct from applied research and development activities, of 250 US firms has found a positive relationship between diversification, measured as the number of 4-digit industries in which a firm operates, and basic research intensity, i.e. the ratio of expenditure on basic research to firm size (Link and Long, 1981).

They also found, however, that when \( R&D \) performing firms were compared to firms not engaged in \( R&D \), average gearing was lower for the former group in the United States but not in the United Kingdom.

On the other hand, Moore found little evidence of a significant difference in availability of finance between high-technology and low-technology small firms; only within the class of fast-growing small firms were high-technology firms more likely to be constrained than conventional firms.

Mansfeld et al. (1981) found no significant effect of firm size on the probability that an innovation would have been introduced without patent protection, but their sample was quite small. Mansfield (1986) correlated firm size with the proportion of a firm's inventions that would have not been developed in the absence of patents in each of 10 sectors. The correlation coefficient was negative and statistically significant in one sector, positive and significant in two sectors, and insignificant in the remaining seven sectors. However, the sample did not include any firms with sales below $25 million.

Early studies which have examined specific mechanisms relating market structure and innovation include the Phillips (1971) monograph on the evolution of the US commercial aircraft industry and the Mansfield (1962) analysis of the impact of innovation on firm growth in the US steel and petroleum refining industries.

Innovation may affect market structure directly or indirectly and in various ways: because of fixed costs of \( R&D \), or by affecting the pattern of firm growth in an industry, or by increasing or decreasing the minimum efficient scale of production, or because innovative activity can either facilitate or hinder entry of small firms. The existence of a variety of factors may be one reason why the econometric literature has been inconclu-
sive. For example, while Levin and Reiss (1984, 1988) and Farber (1981) found a positive effect of R&D intensity on concentration, Mukhopadhyay (1985) found a negative effect of R&D intensity on concentration change and Geroski and Pomroy (1990) found a negative effect of innovation counts on concentration change. Also, the econometric results are not very informative as to the specific ways in which innovation and market structure interact.

22. These empirical observations correspond to a well known feature of theoretical models of oligopoly, namely that equilibrium outcomes very often depend on the order of moves or on assumptions regarding the agents' beliefs about their rivals.

23. It is interesting, in this respect, that Lunn (1986) has found a statistically significant positive two-way relationship between process innovation and market structure and no effect either way between product innovation and market structure, in his study of 191 US industries.

24. According to the Jovanovic and MacDonald (1994) story, shakeouts are caused by major refinements to the technology of the industry. These open up new opportunities for innovation, which many firms are not successful in exploiting, and may also cause an increase in the minimum efficient scale of production.
BIBLIOGRAPHY


agreements in information technology industries”, *International Journal of Economics and
Business*, 2, pp. 3-29.


DTI (Department of Trade and Industry) (1991), *Constraints on the Growth of Small Firms*,
HMSO, London.


expenditures in the firm”, *Southern Economic Journal*, 37, pp. 409-422.


London.

GELLMAN RESEARCH ASSOCIATES (1976), *Indicators of International Trends in Technological
Innovation*, Final Report to the National Science Foundation, Jenkintown, Gellman
Research Associates, Penn.


GEROSKI, P.A. and R. POMROY (1990), “Innovation and the evolution of market structure”,


study of the chemical, drug and petroleum industries”, *Journal of Political Economy*, 76,
pp. 292-306.

Economics*, 40, pp. 359-375.


HALL, B.H. (1992), “Investment and research and development at the firm level: Does the

in *J. Barber, J.S. Metcalfe and M. Porteous (eds.), Barriers to Growth in Small firms*,
Routledge, London.


PIATIER, A. (1984), Barriers to Innovation, Frances Pinter, London.


