Innovate or Die?
A critical review of the literature on innovation and performance
Innovate or Die?
A critical review of the literature on innovation and performance∗

Stefano Brusoni\textsuperscript{a}, Elena Cefis\textsuperscript{b} and Luigi Orsenigo\textsuperscript{c}

\textsuperscript{a} CESPRI - CRORA, Bocconi University and Silvio Tronchetti-Provera Foundation
\textsuperscript{b} Utrecht School of Economics, Utrecht University and Bergamo University
\textsuperscript{c} University of Brescia and CESPRI, Bocconi University

∗ We would like to thank Andrea Pozzi for his research assistance. Also, we are grateful to Andrea Sironi for useful comments and to all the participants of the Conference “Business Performance in the 20\textsuperscript{th} century: A comparative Perspective”. Two anonymous referees gave us interesting comments and suggestions. E.Cefis also wishes to thank Karen Klomp, research assistant at USE, Utrecht University. The financial supports of the University of Bergamo (grant ex 60\%, n. 60CEF104, Dept. of Economics) and the Silvio Tronchetti Provera Foundation are gratefully acknowledged.
Abstract

The idea that innovation leads to positive economic performance has become a sort of truism in recent years. However, empirical evidence showing that innovating organizations and countries outperform non-innovating ones remains scant and scattered. In many ways, the jury is still out. First of all, there is still little agreement about what ‘performance’ means. The range of indicators adopted in the literature varies widely: financial performance, market shares, new products introduced into the market, patents, GDP growth, and so on. Second, the time lag between innovative efforts and performance is often so large, and so industry specific, that it remains just very hard to produce reliable estimates. Third, it is still unclear at what level of analysis one should go looking for positive economic performance. Studies exist that look at the relationship between performance and innovation at the level of design teams, projects, firms, networks, industries, and countries. This paper aims at critically reviewing the wide, yet remarkably scattered literature that aims at measuring and explaining the relationship between innovation and performance. It builds upon an extensive review of contributions in economics, management and organisation sciences to identify trends and results that are consistent and robust. In a nutshell, this paper argues that country- and sectoral-level approaches which emphasize the role that knowledge, spillovers and human capital play in fostering economic growth through innovation need to consider the fundamental role played by competition among heterogeneous organisations in igniting the growth process. In this respect, micro- and firm-level studies can provide useful insights about how competition fosters learning, innovation and ultimately growth.

Keywords: innovation, growth, knowledge, performance, competition.

JEL codes: O14, O33, O40
1. Introduction

The idea that innovation leads to positive economic performance has become a sort of truism in recent years. Practitioners consider innovation as the only possible strategy to react to increasing competitive pressures from fast growing Asian economies characterised by very low labour costs. Policy makers in industrialised countries look at innovation as the panacea against slow growth, unemployment and environmental problems. Academics interested in the short and long term evolutionary dynamics of economic systems have produced an enormous amount of literature that highlight the pivotal role played by innovation in influencing processes as diverse as firms’ growth and survival, their profitability, their internal and external organisation, the evolution of their product portfolios but also sector- and country-level specialisation patterns, GDP growth and country level competitiveness.

However, empirical evidence showing that innovating organisations and countries outperform non-innovating ones remains scant and scattered. In many ways, the jury is still out. First of all, there is still little agreement about what ‘performance’ means. The range of indicators adopted in the literature varies widely: financial performance, market shares, new products introduced into the market, patents, GDP growth, and so on. Second, the time lag between innovative efforts and performance is often so large, and so industry specific, that it remains just very hard to produce reliable estimates. Third, it is still unclear at what level of analysis one should go looking for positive economic performance. Studies exist that look at the relationship between performance and innovation at the level of design teams, projects, firms, networks, industries, and countries.

This chapter critically reviews the wide, yet remarkably scattered literature that aims at measuring and explaining the relationship between innovation and performance. The starting point of this chapter is that the country- and sectoral-level evidence is dubious because of both data issues and theoretical problems. On the data side, most studies try to develop indicators that are aggregations of micro-level indicators (e.g. patent counts, R&D expenditure, human capital investments). The robustness of such method is questionable, if only for an issue of cross-sectoral compensation effects. On the theoretical side, some argue that the answer lies in shifting the analysis from the sectoral and country level, to the world level. This view emphasises the ‘free good’ nature of knowledge. Since innovation processes generate knowledge that is at least partially public in nature, externalities and spillovers become the key engine through which new knowledge is circulated. However,
the effects of such processes should be sought without being constrained by national borders. This approach would have the advantage to provide a solid quantitative evaluation of the effect of innovation on GDP growth.

Others argue that innovation should be analysed in terms not only of the adoption of new and better ‘techniques’, but rather as a co-evolutionary process of technologies and organisations.

However, such co-evolutionary process is likely to be extremely sector- and firm-specific. Hence, the need to shift the analysis from the macro- to the micro-level. The latter observation brings about the main theoretical issue this chapter builds upon. Firm-level heterogeneity plays a key role in explaining innovation. Thus, sectoral- or country-level aggregations might obfuscate the role played by one of the key engines of innovation. This approach is preferred by those who argue that innovation is the outcome of deliberate investments and strategic efforts.

This chapter will pursue – in an admittedly rather idiosyncratic way - the latter line of development, for two related reasons. First, on theoretical grounds, we believe that firm- and sector-level specificities are not a nuisance to be abstracted away in order to deliver cleaner and neater evaluations of the relationship under examination. Rather, they are a fundamental constituent of the engine that makes economies endogenously grow. To put it simply, economic systems that can exploit this sort of specificities (variety) are more likely to generate sustainable patterns of growth. Second, on empirical grounds, recent literature has finally begun to deliver robust findings in terms of the positive and persistent relationship between innovation and performance at the micro- (firm-) level (Cefis, 2003a; Cefis and Ciccarelli, 2005). Such literature, and some recent developments in innovation management and organisation sciences, could provide a viable starting point to shed light – and generate meaningful policy implications - on the relationship between innovation and performance.

Thus, this chapter will continue its exploration of the micro-level literature. First, evidence is emerging on the persistent relationship existing between innovation and profitability (Cefis, 2003a). However, most sectors seem to host relatively few innovators, while most companies tend not to be innovators. Or, at least, they aim to be fast followers rather than pioneers. This is no paradox. Innovation processes are and will remain inherently uncertain. Thus, pioneers face riskier challenges than followers. Industries seem to be characterised by the persistence of a relatively small core of ‘serial innovators’ that open the path to a long tail of (often very fast) imitators. In a rather speculative manner,
one might argue that the tension between would-be innovators and imitators is the fuel that starts the engine of endogenous growth. In a nutshell, the competition among heterogeneous organisations is the operating principle of the growth engine of countries. Aggregated approaches that identify the engine of growth in the innovation process very often fail to pay adequate attention to its fuel and combustion process, i.e. competition among heterogeneous organisations.

This chapter is organised as follows. Section 2 will focus on the macro- and meso-level analysis of the relationship between innovation and performance. Section 3 will move on to the micro-level issues. Section 4 will summarise and conclude.

2. Innovation and performance: macro- and meso-level debates

‘Technological innovation has been one of the most important contributors to the growth in employment and incomes in the U.S. economy.’ (Mowery and Rosenberg, 1989: vii).

‘[I]nnovation is of importance not only for increasing the wealth of nations in the narrow sense of increased prosperity, but also in the more fundamental sense of enabling people to do things which have never been done before.’ (Freeman and Soete, 1997: 2).

‘It is well known that some countries grow at a much faster rate than others, even for long periods of time. … How can one account for these differences? The suspicion that it may have something to do with technology has been around for a long time’ (Fagerberg, 1994: 1147)

These citations quite clearly make the point of the range of issues that revolve around the relationship between technological innovation and performance. They are quite representative of the breadth of meanings normally attached to the concepts of innovation and performance, as well as the breadth of economic and societal implications stemming from their relationship. In the space of few words, we are told that technological innovation plays a central role in explaining nothing less than the growth rates of income and employment in the world’s largest economy, the betterment of the quality of life of people, as well as the persistent differences among income growth rates at country level.

Similarly, the breadth of approaches that have looked theoretically and empirically to such relationships is such that it makes efforts to produce a comprehensive review of dubious utility. Suffice to say that technology has been conceptualised as a ‘free good’, as an externality of other economic activities, and as the result of R&D activities performed
by private firms and public research centres (Fagerberg, 1994: 1170). The theoretical underpinnings, behavioural rules and policy implications of each view are remarkably different. Yet, over the years, some degree of convergence in the literature has been observed.

Nowadays, nobody would question that a positive relationship exists between – broadly speaking- innovation on the one side and country-level performance (in terms of growth rates and trade patterns) on the other. This is an accepted stylised fact in various streams of literature. Nevertheless, we will show that the available empirical evidence is disappointingly timid on this issue. Besides, the evidence is much less clear about “how much” does innovation contribute to growth or export performance and even more so about the “mechanisms” through which innovation exerts its impact.

The key departing point of our analysis is that the link between empirical evidence and the predictions of macro models (and their micro-foundations) is –to cut short a long story- rather weak.

As it is well known, following Solow’s seminal contribution, growth accounting methodologies used to blackbox the effects of technological innovation into the residual that became known as Total Factor Productivity. The latter was considered to be the effect of technology. However, subsequent attempts to clarify and measure the nature of the residual have somewhat reduced the magnitude of the attributed contribution of technological change to growth, by emphasising instead the role played by variables like education, human capital, etc. Moreover, how and through which processes technology displayed its effects was not known.

Dissatisfaction with neoclassical models and with growth accounting exercises has always been voiced. For example, Kaldor’s work has inspired a wealth of studies that look at issues related to demand composition and cumulative causation. The idea that ‘success breeds success’ has come to play a fundamental role in more recent micro- and meso-analyses about the effects of technological change. It plays a major role in understanding the persistent difference in growth rates at country level as well as more recent results in terms of profitability, growth rates and innovativeness at firm-level on which the third section of this chapter will specifically focus. Moreover, growth is not explained by making references to the public nature of knowledge as some sort of deus ex machina that drives diffusion and thus country-level growth.

More recently, huge efforts have been devoted to developing the stream of research that normally goes under the heading of ‘new growth theory’. Such literature stems from
the dissatisfaction with the assumption of technological change as the *exogenous* engine of growth and development. To overcome this limitation, models have been developed in order to endogenise the very process of growth (Romer, 1986, 1990) made possible by various forms of technological change (e.g. capital deepening vs. capital widening, etc.). According to this line of research, the quasi public nature of (scientific and technological) knowledge generate positive externalities through spillovers and other non-market based mechanisms that would allow the emergence of increasing returns at industry level. In essence, the costs of generating knowledge need to be paid only once. Diffusion is virtually free and costless. Thus, also those countries and firms that do not generate new knowledge can benefit from it. Of course, the policy implications of such approaches focus on how to solve various forms of market failures associated with the under production of economically useful knowledge, given the weak incentives stemming from weak appropriability.

A further recent development – with strong kinship with endogenous growth theory - is represented by the literature on General Purpose Technologies. Such literature aims at striking a balance between neo-classical approaches focused on the idea that technological knowledge is to some extent a public good on the one side; and those approaches that emphasise instead cumulative causation and positive feedbacks. GPT approaches, building upon the concept of pervasive technology, argue that growth is led by the emergence of technologies that are general enough to allow for a variety of uses in the adopting industries (typical examples would be the steam engine, ICT, etc.). Setting aside the issue of how ‘general’ is a technology like the steam engine (not much, according to e.g. Nuvolari (2004), these models predict a cyclical evolution of growth patterns (with the upswing driven by the diffusion of a new GPT). However, they seem to also imply that GPTs replace each other, and cannot coexist. The validity of this point, we suspect, depend entirely on how broad a definition of GPT we are willing to accept (e.g. ICT vs. semiconductors).

Moreover, the notion that fundamental technological breakthroughs lead to cyclical growth is certainly not new in the literature, ever since Schumpeter and the debate on Kondratieff cycles and long waves. If anything, the contributions by Christopher Freeman, Carlota Perez and Luc Soete to the development of the notion of “techno-economic regimes” suggest that impact of what now is customarily called GPTs implies deep structural transformations in the economy and in institutions. And it usually takes a long time (including periods of falling productivity) before the potential offered by the new
technologies has begun to be reflected in productivity and growth (Freeman and Perez 1988; Freeman and Soete, 1997).

Against this background, recent empirical research has motivated what one might call a neoclassical revival. Notably, Mankiw, Romer and Weil (1992) use international cross-sectional evidence to argue that the steady state of the neoclassical growth model, when suitably augmented to allow for human-capital accumulation, can account for over three-quarters of the international variation in income per capita levels; that cross-sectional regressions yield plausible estimates of factor shares; and that measured rates of convergence, when appropriately conditioned on country-specific parameters, are consistent with the predictions of the model. In addition, Young (1995) and Jorgenson (1995) have used growth-accounting techniques to argue that technological progress measured by total factor productivity growth has been an insignificant source of income per capita growth relative to capital accumulation.

Klenow and Rodríguez–Clare (1997) react to these claims suggesting that the neoclassical revival has gone too far and conclude that the assumption of identical productivity levels across countries embedded in the neoclassical model was untenable. Similarly, Dinopoulos and Thompson (1998) reject the assumption of an exogenous and identical level of technology in all countries, and show that the results of Mankiw et al. (1992) depend critically on the measure of human capital which was used (the secondary school enrolment rates) and that are not robust to the use of other, more satisfactory indicators.

On somewhat different grounds, Jones (1995a, 1995b) has argued that some crucial conclusions of the conventional models of endogenous growth (Romer, 1990) rest on the built-in and empirically falsified assumption of the existence of scale-effects in R&D: doubling the resources devoted to R&D, the per capita growth rate of output should also double, at least in the steady state. A simple solution to the scale effects problem would be to assume that what matters for growth is the fraction of resources devoted to R&D, an approach that has been common in the empirical industrial organisation literature. Jones, however, rebuts by arguing that it is not easy to justify why one should need more workers to create a new process or a new product in a larger economy. The removal of scale effects makes growth only partially endogenous, in the sense that long run growth still depends on technological progress, which in turn is endogenously generated by the investment decisions of (rational) agents; but policy changes do not longer have permanent effects on long run growth.
A number of models have been proposed to deal with this problem (see for instance Dinopoulos and Thompson 1998 and 2000, Peretto 1996). What is important to emphasise here, though, is that the debate does not concern so much the existence of a link between - and the relevance- of technological change for economic growth, but rather the mechanisms through which technical progress impacts on growth. The specific processes (i.e. the engine) underpinning such link are probably hard to trace and observe at the level of analysis chosen by this literature and given the fundamental assumptions employed by these models (e.g. equilibrium, optimising behaviour, access to technical knowledge by countries, etc.).

On the evolutionary front, growth models have achieved some results in providing explanations to growth differentials at country level relying on more sensible microfoundations (Dosi et al., 1994). However, this literature also finds it difficult to supply neat quantitative results on the relationship between innovation and growth (Verspagen, 2001 and 2003). Evolutionary theories do provide some evidence about the relationship under examination, but with numerous caveats. At the country level, there is some evidence that technology does play a role in explaining growth differentials and convergence (or lack thereof). The related literature on catching up and (non)convergence of country growth rates also identified a relationship between various indicators of innovativeness and country-level performance (Fagerberg et al., 1994). However, also in this case the relationship appears to be time-dependent and hard to be precisely measured. For example, catching-up – rather than innovation per se - would seem to have a much bigger impact on countries' growth in the post- World War II period, up until the 1990s.

However, increasing R&D investments appear as a decreasingly good predictor of a country’s performance, if it is true that R&D investments are more and more necessary just to stand still (Verspagen, 2001).

One of the main problems is that very often country- and sectoral-level estimates adopt indicators of inputs to the innovation process (i.e. R&D investments, human capital, patents) that make sense at the micro-level, but are troublesome at the macro-level. We have already mentioned in the introduction the fact that the very act of aggregating micro-level indicators creates a sort of compensation effects that might cancel out sectoral- and firm-level heterogeneities. Moreover, technological learning involves many more elements than simply inventive discovery and patenting: equally important activities are imitation, reverse engineering, adoption of capital-embodied innovations, learning by doing, and learning by using (Freeman 1982; Dosi 1988; Pavitt 2001). And technological change goes
often together with organisational innovation. Finally, despite technological diffusion is taking place at a rather high rate, at least among OECD countries, important specificities in “national systems of innovation” persist related to the characteristics of the scientific and technical infrastructure, local user-producer relationships and other institutional and policy features of each country (Lundvall 1995; Nelson 1993; Freeman, 1995). Thus, measuring the relationship between technological progress and economic growth is inherently difficult.

One way out might consist in reducing the scale of aggregation, from countries to industries. The available evidence points to a significant relationship, but mainly in the case of high-tech sectors, e.g. telecommunications or biotechnologies (Laursen, 1999; Meliciani and Simonetti, 1997; Montobbio, 2003, Verspagen, 2003, Montobbio and Rampa, 2005). Similarly, at the sectoral level, the evidence points to a positive relationship between capability of innovating and quickly adopting new technologies and successful trade performance (Dosi et al. 1990). But this correlation does not hold for all industries and is much stronger (perhaps not surprisingly) for innovation intensive sectors.

The latter point is quite important for both theoretical and policy purposes. In fact, the idea that not all sectors are alike is not new. Many have raised the point that some sectors act as engines of growth generating breakthrough that are then adopted by laggards. This point has inspired both the evolutionary literature and the recent approaches that look at the impact of the so-called ICT revolution on the productivity of adopting sectors and entire countries. Some sectors appear to play a major role however as suppliers of key innovations to other sectors. However, the process of diffusion can hardly be understood in terms of spillovers and technology as ‘manna from heaven’ (Verspagen, 2001). In this respect, the canonical reference to ‘spillover effects’ hides more than it enlightens. As argued by Breschi and Lissoni (2001), under the spillover banner we can find a variety of channels through which knowledge is diffused. Most of these channels need investments and effort to be created, activated and maintained. Consistently, Verspagen (2001) also finds that R&D investments are becoming increasingly important not only to would-be innovation leaders, but also to followers as the process of diffusion of new technologies requires increasingly sophisticated skills and higher levels of absorptive capacity.

Besides issues of data availability, the general lesson we draw from the analysis of both mainstream and evolutionary accounts is that the relationship between innovation and performance at country-level remains quite elusive. In order to make sense of this complex pattern of relationships, we need to turn our attention to micro-level approaches.
3. The micro impact of innovation on performance

In the last two decades, thanks to the development of longitudinal micro data-sets, the relationship between innovation and performance at the firm level has been investigated. Indeed, efforts of economists and statisticians have been concentrated in collecting not only financial and economic firm data but also indicators of innovative activities carried out inside the firms. This is the main reason why in the last twenty years studies on the effects of innovation on some firm-level performance indicators have appeared, even if the links between economic/financial datasets and innovation datasets are still quite unexploited and a lot of work is still needed. Thus, a now large body of literature is available suggesting that indeed innovative firms enjoy better performances than non-innovative firms in terms of a variety of dimensions like profits, wages paid to employees, productivity, export performance, etc. However, below the surface, things appear to be much more complex than meets the eye. In particular, the beneficial effects of innovation on firms’ performance: are they transitory or permanent? Are they sector-, technology-, and firm-specific? Why certain performance variables seem to be affected by innovation but others – primarily and remarkably growth - do not? How does innovation impact on performance?

3.1 Innovation and profits

Quite a natural starting point in understanding the relation between innovation and performance is to study what the effects of innovation are on firm profits.

There are at least two alternative views about the way innovative activities affect the profitability of the firms. In the traditional view, innovations have only a transitory effect on the firm profitability by altering its competitive position in the short-run. The introduction of an innovation gives to the firm a temporary monopoly power, which, by increasing the firm’s market share, allows for higher profits until other firms can imitate the innovation (Aghion and Howitt, 1992; Klepper, 1996).¹

A second approach points out that innovations intrinsically ”characterise” a firm in that it creates a structural difference between innovating and non-innovating firms. According to this view, each firm owns different technological competencies that are firm

¹This approach is common in the literature on the first mover advantage (e.g. Gorecki, 1986, and Robinson, Kalyanaram and Urban, 1994) and on patent races (e.g. Tirole, 1988 (Ch.10), and Reinganum, 1989).
specific, cumulative, and emerge from the various learning processes the firm has passed through. These internal competencies, together with specific behavioural patterns, enable the firm to better face changes in the market in order to survive or even to obtain satisfactory profits over time (see Malerba and Orsenigo (1995), Cohen and Levin (1989), Dosi et al. (1995)).

The existence of correlation and the investigation of the channels of transmission between profit margins and innovations have been recurrent concerns of industrial policy makers. At a first glance, there is nothing that prevents non-innovating firms to perform as well as innovating ones, either because non-innovators could develop different capabilities, or because innovation may require a sunk-cost investment independent of innovation returns. On the other hand, answering these questions can be important for antitrust regulation to be able to discriminate among firms with relatively higher profits: between those who exploit a monopoly position, and those who own capabilities and competencies that make them systematically better than others. Previous studies (Geroski, Machin and Van Reenen (1993); Geroski, Van Reenen and Walters (1997)) have found positive, though not well-determined direct effects of innovations on profitability in the short run, and large indirect effects due to the relative insensitiveness of innovating firms to adverse macroeconomic shocks.

A potential limit of this stream of studies is given by the failure to fully control for heterogeneity at the sectoral- (e.g. most of these studies involve samples of firms active in very different sectors and technologies and heterogeneity is at best controlled for by dummy variables) and firm-level. Taking directly into account the heterogeneity among firms, a more recent study by Cefis and Ciccarelli (2005) shows, on the contrary, that there is a positive and well-determined effect of innovation on the firm profit margin. The effect is larger for innovation patented 2–3 years before and then decreases smoothly as time passes by. Furthermore, their analysis shows there exists a permanent difference between innovators and non-innovators. In the long run innovators and non-innovators converge (with different speeds) to different profitability steady states, the innovators steady state being higher than the non-innovators one. The effect of innovation on

---

2 Several empirical studies have shown that there is evidence of heterogeneity among firms due to industrial and size classification, the access to capital market, etc. Furthermore, technological and organisational firm-specific characteristics can cause heterogeneity at the firm level. Previous studies on profitability have not tackled the heterogeneity problem directly, but only indirectly considering an exogenous clusterisation of firms by size or industrial sectors. Instead, Cefis and Ciccarelli used a Bayesian approach that allows to control for general sources of firm heterogeneity in the population structure, acknowledging in this way that every firm can be generally different from the others.
profitability seems to be permanent over time. These results support the view that innovating firms have developed internal competencies and behavioural patterns that allow them to face the challenges of the market better than non-innovating firms. Moreover, the effect of innovation on profitability is greater the more persistently firms innovate.

3.2 Innovation and firm export performance

The capacity of a firm to generate export returns has been often interpreted as an important indicator of firm competitiveness and firm performance. Actually, the ability of a country to create earnings from exports has been considered one of the major proxies of international competitiveness at the country level. The relation between innovation and export performance has, indeed, been firstly and deeply studied at the macro-level. Macro studies, both theoretical (Posner, 1961 and Vernon, 1966) and empirical (among others Wakelin, 1998a, Greenhalgh, 1990) have identified innovation (proxied in different ways ranging from R&D expenditures to patenting activity) as a potential explanation for different world trade performances of the countries. On the theoretical side, models of international trade generally assumed that within a particular industry all firms were symmetric and if a sector was exporting, than all the firms within that sector were exporting. Beginning with Venables (1994), asymmetries among firm export behaviours were taken into consideration in theoretical models, while an increasing body of empirical literature has tested whether there were significant differences in export performances among firms, and whether these asymmetries were explained by firm specific innovation behaviour.

In the majority of the studies at the firm level (among others Wakelin, 1998b; Basile, 2001; Roper and Love, 2002; and Gourlay and Seaton, 2004), firm export behaviour has been defined in a dual way: both as a probability for a firm to export, and as the propensity to export for the exporting firms. The empirical relation between innovation and export has been tested using as control variables firm characteristics (as labour cost per unit of product, size, ownership structure, and firm location) and industry characteristics.

In the studies focused on European countries, innovation has been found a key determinant of firms export performance. In U.K., Gourlay and Seaton (2004, using a panel of 2134 UK firms between 1988 and 2001) and Greenhalgh et. al (1994) conclude that both the probability of exporting and firm export performance are positively influenced by respectively R&D expenditures and patents and successful innovations (according to the SPRU innovation database of innovations). In their study on Italian firms
Basile (2001) and Sterlacchini (1999) respectively find that the export intensity of innovating firms is systematically higher than that of non-innovating firms; and that small, non-R&D-performing firms are more likely to export when they have innovative activities (especially the amount spent on design, engineering and pre-production). In the French manufacturing sector, Barlet et. al. (1998) explain the share of new products in total sales and exports by the innovation types firms have implemented.

Studies on non-European countries include Ozcelik and Taymaz (2004) on Turkish firms, finding that variables such as size, advertisement intensities, ownership structures, and composition of employees have different impacts on export in innovator firms than in non-innovator firms. They conclude that in international competition, innovations and R&D activities are quite important. On the other hand, Hasan and Raturi (2003) studying the Indian manufacturing sector, show that the role played by technology (R&D and technology transfer agreements) does have a positive influence on the entry of firms into export markets, but only limited influence on the export volume. Factors such as labour intensity and firms size have more influence on both export participation and export volume.

Finally, considering the impact of innovation on firm export performance in an historical perspective, Fernandez-Perez (1999) studies the late 19th century Spanish sherry market. She shows that the single Spanish firm that was able to adopt technical innovation in production and distribution, was the only Spanish firm that was successful on foreign markets.

### 3.3 Persistence of innovation

The issue of persistence in innovative activities is particularly relevant in the context of the discussion about its relationship with firm’s profitability, but also, more in general, with regard to the properties of the patterns of innovative activities. Since Schumpeter the economic literature on technological change has developed two main views of the innovation process. Referring to what Schumpeter states in *The Theory of Economic Development* (1934) (known as the Schumpeter Mark I model), the process of technological change is considered to be a process of “creative destruction”. Conversely, referring to Schumpeter’s *Capitalism, Socialism, and Democracy* (1942), the process is seen as a process of “creative accumulation” (or Schumpeter Mark II model). The difference between the two depends on fundamental assumptions about the properties of
technology and of the innovative process\textsuperscript{3}.

In a rather simplified way, in Schumpeter Mark I technology is equally accessible to everybody and technological change is a random process, driven by a population of homogeneous firms that have a certain probability of realising technological opportunities. Innovation generates monopoly power that is at best only temporary, since it is quickly challenged and eroded by competitors. Since the relevant knowledge base is easily available, new innovators systematically substitute incumbents and typical innovators are small, newly established firms.

Conversely, in Schumpeter Mark II technical knowledge has a strong tacit component and is highly specific to individual firms. Innovation results from the accumulation of technological competencies by heterogeneous firms. Firm-specific technical change is cumulative, meaning that the generation of new knowledge builds upon what has been learned in the past, and accumulated competencies significantly constrain the future technological performance of the firm. Over time, the firm-specific, tacit, and cumulative nature of the knowledge base build high barriers to entry. A few (large) firms eventually come to dominate the market in a stable oligopoly.

Thus, the presence or absence of persistence in innovative activities is a major property of the innovative process and an important feature of the patterns of technological change, which has significant implications for both theory\textsuperscript{4} and policy-making.

Because of its cumulative nature, technological change is usually characterised by dynamic increasing returns (learning-by doing, learning-to-learn, research breeds new opportunities, etc.). Persistence in innovative activities might mean that technological change could be one source of increasing returns that can support persistent growth (Barro and Sala-i-Martin, 1995).

On the one hand, persistence, in general, would give some support to the "competence-based" theory of the firm at the microeconomic level (Nelson and Winter 1982, Teece and Pisano 1994), as well as to endogenous growth theories at the macroeconomic level. On the other hand, persistence in innovative activities would weaken those interpretations of the processes of growth of firms, industries, and countries (ranging from simple Gibrat-type processes to the models of the real business cycle) where dynamics are essentially driven by small uncorrelated shocks.

\textsuperscript{4} Models as different in inspiration as those of Nelson and Winter (1982) and Ericson and Pakes (1992) show that these two alternative patterns of technological change can be interpreted as two faces of the stochastic process, which drives technological accumulation at the firm level and thereby drives the dynamics of the industry.
More generally, understanding whether innovative activities are persistent or not at the firm level constitutes an important piece of evidence for founding and improving current theories of industrial dynamics and evolution, where some forms of dynamic increasing returns play a major role in determining degrees of concentration and its stability over time, rates of entry and exit, and so forth (Klepper 1996, Jovanovic 1982, Hopenhayn 1992, Nelson and Winter 1982, Dosi et al. 1995).

However, little is known about the relative empirical relevance of persistence in innovative activities. Recently, a few studies have begun to provide some (and somewhat contrasting) empirical results. Malerba and Orsenigo (1999) examined the patterns of innovative entry, exit and survival, using European Patent Office data for six countries. They find that innovative activities are characterised by high degrees of turbulence. The population of innovators changes substantially over time through processes of entry and exit. A large fraction of new innovators is composed by occasional innovators. These also constitute a significant part of the whole population of innovators, but a much lower share of the total number of patents at any given time. Only a fraction of entrants survive and succeed in remaining innovative after their first patent. When they do, however, their technological performance improves consistently in the following years. Moreover, large innovators tend to remain large for long periods of time. These results would seem to suggest that although turbulence is a pervasive and quantitatively important phenomenon, innovative activities are nevertheless generated - in general - by a relatively stable core of large (both in terms of patents and employees) and persistent innovators, that account for a very large share of total patents. Around this core, one observes a large turbulent fringe of small, occasional innovators, that often patent only once and cease to patent thereafter.

Conversely, Geroski, Van Reenen and Walters (1997) estimate a Proportional Hazards Model of the probability that the spell of time in which a firm innovates will end at any particular moment. They find little evidence of persistence at the firm level. Very few firms innovate persistently and this happens only after a threshold level (5 patents or 3 major innovations), which only few firms ever reach.

Cefis (2003a) uses instead non-parametric techniques (Transition Probability Matrices) on a sample of patents applied for by 577 UK firms. She finds little persistence in general, but simultaneously strong persistence among the greatest and the smallest

---

5 It is worth noting the close analogy between these results and the evidence on industrial demography that shows that entrants in a given industry in any particular year tend to be small firms and do not survive for long. When they do, however, either they tend to grow at very fast rates or they have an initial larger size. See, for a survey, Audretsch (1997). At this stage of the analysis, we cannot pursue further the issue of the extent to which and how the two phenomena are related.
innovators. Moreover, substantial heterogeneity in the degree of persistence across sectors and across firms’ size is observed.

In other words, there is a strong persistence in remaining in the states in which firms do not apply for a patent every year, or in the state in which firms apply for many patents (at least 6). "Great" innovators (firms that apply for at least six patents per year) are very few in number (2.37% on average), but they account for the large majority of patents requested (77.85% on average). These results suggest that innovative activities, at least those that are captured by patents, are persistent.

3.4 Does persistence in innovation translate in persistence in firm’s profitability?

In order to make a further step in analysing the relationship between innovation and performance, we should ask whether the persistence in firm innovative activities is translated into persistence in firm’s performances. Considering profit margins as a proxy of firm’s profitability, the question could be rephrased as: are firms who persistently innovate able to earn profits that are persistently above the average?

Substantial research effort has been devoted to the examination of profit persistence. Recent literature has been addressing the question whether industrial profit rates eventually converge to a common rate. Several empirical studies have shown that firms display persistent differences in profitability (Cubbin and Geroski, 1987; Mueller, 1990; Glen, Lee, Singh, 2001), and more in general in firm performance indicators like output shares and labour productivity (Geroski, 1998 and Cantner and Kruger, 2003). That is, profits do not seem to converge to a common rate of return. Moreover, evidence seems to indicate that the adjustment of profits to their firm-specific “permanent” values is rather quick, although a significant variability is observed across different countries (see for example, Geroski and Jacquemin, 1988; Drucopoulus and Lianos, 1993; Maruyama and Odagiri, 2002). However, it is hard to say to what extent the observed persistence in profitability differentials reflects the persistence of differential “efficiency” levels, which are not eroded away by the competitive process.

Profits are generally modelled as a function of past innovations (for example Geroski et al., 1993; Roberts, 2001) considering innovation as a shock that has an impact (large or small, permanent or transient) on the firm profitability. On the contrary, Cefis’s purpose (2003b) in analysing whether there is persistence in the joint distributions of patent applications and profit margins is to provide empirical evidence on the relations between the processes that generate profits and innovation. Her results show that there is high
persistence in a firm’s relative position with respect to profits and innovation jointly considered. That is, firms that are systematic innovators and earn profits above the average have a high probability to keep innovating and earning profits above the average, just as firms that are occasional innovators and earn profit below the average have a high probability to remain in the initial situation. The persistence in innovation seems to foster the persistence in profits, especially in the long run. Moreover, the modification of the firm’s idiosyncratic characteristics (given especially by its technological and organisational capabilities) is a very gradual process: we can observe mobility in both dimensions only in the long run.

Finally, the mobility of a firm’s relative position with respect to the average profitability appears not to be correlated with the firm’s relative position in the innovation dimension in the short run. The high persistence that characterises the firm profitability is not influenced by firm innovative characteristics. However, firm’s relative position in the innovation dimension matters in the long run: the probability to earn profits above the average, in the long run, is higher if a firm starts as a systematic innovator than as an occasional innovator.

These results suggest that the relation that needs to be studied is the persistence in profits as a function of the persistence in innovations. That is, the attention should be focused on treating both profits and innovation as processes, and in the case of innovation as a cumulative process.

3.5 Innovation: a matter of life or death.

Recent papers (Malerba and Orsenigo 1995, 1996 and 1999; Breschi, Malerba and Orsenigo, 2000, Malerba, Orsenigo and Peretto, 1997) have provided empirical evidence that different technological classes exhibit different patterns of innovative activities, one resembling the creative destruction model, the other the creative accumulation model, with little variation across countries. It was then shown that the patterns of innovative activities are linked to the nature of the relevant technological regime, defined by opportunity and appropriability conditions, degrees of cumulativeness of technological advances, and the nature of the knowledge base (for further discussions of the notion of technological regimes, see, e.g., Nelson and Winter 1982, Winter 1984, Dosi 1988, and Malerba and Orsenigo 1996). These results suggest that it might be misleading to derive strong conclusions from aggregate data. Rather, we expect innovative activities to be systematically characterised by different degrees of persistence and profitability across technologies and industries, as well
as across countries.

These intersectoral differences do not concern only persistence in innovative activities, but in general they concern the whole role of innovation in firms’ performance. Indeed in some sectors (or maybe it would be better to use the words: in some “technological regimes”) innovation affects firm’s performance in such a way that it becomes a matter of life or death.

A quite recent stream of the literature has investigated the role of innovation in shaping firm’s survival. In the last two decades the empirical literature in industrial economics has highlighted the high degree of turbulence, due to entry and exit dynamics, that characterised most industries (Caves, 1998). A large proportion of new firms exit the industry within a few years after entry as well as established firms are subject to shakeouts. The growth and survival of firms will depend on their ability to successfully adapt their strategies to changing environments. In such environments, innovation creates a variety of competitive positions and enhances a firm’s potential to succeed in the market. This effect is important for new enterprises as well as established firms. Innovation may increase the chance of survival of new firms by allowing successful niche strategies. At the same time, innovation is necessary for well-established firms to deal with new and emerging or ‘disruptive’ technologies (Christensen, 1997).

A number of studies have looked empirically at the factors that influence the probability of firms to survive in the market. At the firm level, these factors have been traditionally identified in the size and age of the firm, both increasing survival probability (Dunne and Hughes, 1994). At the industry level, the characteristics of demand, such as market size and growth rates (Mata and Portugal, 1994), the characteristics of technology (Audretsch, 1995) and the life cycle (Agarwal and Audretsch, 2001) have been found to be important determinants of the survival probability. These studies, however, focus either on structural features of the firm or on differences in the external environment. Only few empirical studies have looked at the role of innovation within the firm in shaping the survival probability of it.

Among them, Cefis and Marsili (2004) have explored how innovation influences the survival probability of manufacturing firms. Confirming previous research, they find that the demography of firms is influenced by firm size and age. The firms most likely to exit and disappear from the market are small and young firms. The effect of size and age is shaped, however, by the extent firms do engage in innovative activities.

In general, their results show that the ability to innovate increase survival probabilities
for all firms and across most industrial sectors. They label this as the ‘innovation premium’ associated with survival. In particular, this premium is highest for small and young firms: indeed, for those firms that are in the greatest danger of failure. In fact, small, young firms that innovate have a 23 per cent greater chance of surviving than those that do not innovate.

### 3.6 Innovation and Firms’ Growth

While the evidence discussed so far would seem to support the view that innovation does indeed have an impact – and possibly a permanent one - on firms’ performance, a rather different picture is suggested when one looks at firms’ growth. Indeed, it proves quite difficult to find any systematic – let alone persistent – effect of innovation on firms’ growth.

There are different issues at stake here. First, the absence of correlation between indicators of innovation and growth might be simply due to the fact that innovation does not have any significant impact on firms’ growth because the latter is driven by other factors, e.g. rates of growth of demand, advertising, price competition, etc. While growth is certainly influenced by a host of different factors other than innovation, this conclusion is not particularly convincing though, given the relevance usually attributed to innovation as a fundamental competitive tool in the management literature, in thousands of case studies and in survey data.

A second interpretation might be that innovation does not translate into growth because other firms are innovating too, in other words imitation immediately erodes away differentials in competitiveness across firms. This explanation sounds more convincing than the previous one. However, the effects of differential innovation should be visible in the data, at least for short periods of time. And while imitation is certainly a crucial phenomenon, still plenty of evidence suggests that firms are capable of privately appropriating – at least partially – the benefits from their own innovative activities through a variety of means, which range from patents (only in few industries like chemicals and pharmaceuticals), to secrecy, lead times, learning curves, control of complementary assets, etc. (Klevorick et al. 1995).

Third, innovation may be considered as a largely random and unpredictable phenomenon. Thus, innovation events would appear as unsystematic shocks hitting growth rates, but only temporarily and rapidly fading away. Moreover, as innovation is purely random, all firms would have the same probability of innovating and the ensuing competitive advantages would be randomly distributed across firms.
This interpretation is consonant with the now enormous literature on the so-called Gibrat Law, or Law of Proportionate Effects, which assumes that firm's size follows a random walk and hence that firm's growth is erratic (for a recent discussion see Sutton, 1997). As a consequence there is no convergence in firms’ size within or across industries, and no stable or predictable differences in growth exist either in the short or in the long run. Rather, growth is driven by small idiosyncratic shocks. This is indeed what one would expect if growth were driven by innovation, but the latter were erratic.

We cannot discuss here the tangled discussion about Gibrat law. For the purposes of this chapter it is just worth mentioning a few issues and problems.

First, the question whether innovative activities are erratic at the firm and at the industry level is closely related (but not identical) to the discussion about the persistence of innovation discussed in Section 3.3. More broadly, one should consider that innovation is the outcome and it is driven by a variety of learning processes, including learning-by-doing and by-using, formal R&S, imitation, “spillovers” from other firms and industries, etc. Thus, the question might and should be reformulated by asking if these learning processes are persistent and systematic. Here again, availability of suitable indicators and measures constitutes a major problem for empirical analysis. And whereas a substantial body of evidence at the firm level or through case studies and survey data indicates that learning – in its various forms – is often characterised by cumulativeness, still it is often hard to obtain similar results in econometric works (see for instance Geroski and Mazzucato 2002).

Second, the evidence on Gibrat Law is mixed. Trying to summarise heroically, conventional wisdom appears to be that Gibrat Law holds for large firms, but smaller (and younger) companies tend to grow faster and with a higher variance. Thus, firms’ size would be characterised by regression to the mean. Moreover, several further violations of Gibrat Law have been identified in the literature, in particular as it concerns the decrease of the variance of growth rates with size and age, and the existence of autocorrelation between growth rates. Finally, Gibrat Law (but also the reversion to the mean interpretation) might also result from the failure to properly control for firms’ heterogeneity (Cefis et al, 2005).

Having said this, it remains a major puzzle why innovation appears to bear systematic effects on many measures of firms’ performance (notably productivity, profit and persistence of supra-normal profits), but not on growth.
4. Discussion and conclusions

This chapter has just begun to scratch the surface of the enormously complex relation between innovation and performance. A more thorough discussion would need much more detailed analysis of available empirical evidence (including measurement issues), theoretical models, and the links between the two. Moreover, other measures of performance should be considered and discussed: performance is a multifaceted concept, which can hardly be captured by simple and univocal indicators.

The main message of this chapter is however quite simple, and at the same time challenging: in the essence, we know that innovation benefits countries and firms. But how and how much remains remarkably obscure. In particular, while at the micro level we can find evidence of the existence of the relationship between innovation and some indicators of performance, as we move up to the sector- and country-level things become fuzzier. It appears that some sort of compensation effect is at work that crowd out firm-level gains from innovation. The key question is whether this is a mere consequence of the unavailability of suitable indicators, or rather this is a structural characteristic of the market economy.

In this chapter, it was suggested that a microeconomic perspective is necessary to begin to understand these phenomena and to unveil if and through which mechanisms innovation at the firm level translates into sectoral performance and countries performances. As mentioned in the Introduction, at the macro level a disproportionate emphasis is attributed to spillover effects as the main mechanism through which innovation impacts on countries' growth. Remarkably, this notion is now customarily employed in most of the literature concerning regional growth, economic geography, and regional economics, despite some recent contributions that start to challenge this view (Zucker and Darby, 1998; Breschi and Lissoni, 2001).

In order to unpack the concept of spillover, it would probably be useful to go below the firm level, in order to understand how innovative activities at the level of particular research teams or projects impact on the behaviour and the performance of the firm as a whole. In this respect, the literature in innovation management and organisation sciences provides quite useful insights in terms, for example, of the communication patterns underpinning successful projects, the balance between central authority and empowerment, the relationship between explorative and exploitative R&D. It is at this level of analysis which we can explore what makes innovating firms permanently different from non-innovating firms.
Moreover, most likely, the mechanisms linking innovation and performance are variegated and differ across technologies, firms, industries, countries – and over time. Black boxing such mechanisms into generic terms such as ‘spillovers’ is likely to obfuscate our understanding and merely create yet another indicator of our ignorance.
References


Dosi, G., K. Pavitt e L. Soete (1990), The Economics of Technical Change and International Trade, Brighton, Wheatsheaf.


Freeman C., (1982), The Economics of Industrial Innovation, Francis Pinter, London


Klevorick, A. Levin, R., Nelson, R. and Winter S. (1995), On the Sources and Significance of Inter-Industry Differences in technological Opportunities, Research Policy, 24, 185-


