CESPRI
Centro Studi sui Processi di Internazionalizzazione
Università Commerciale "Luigi Bocconi"
via R. Sarfatti 25 - 20136 Milano
tel. 02 5836.3397 - fax 02 5836.3399
http://www.cespri.uni-bocconi.it/

Catherine Beaudry – Stefano Breschi

DOES 'CLUSTERING' REALLY HELP FIRMS' INNOVATIVE ACTIVITIES?

WP n. 111 **July 2000**

Does 'clustering' really help firms' innovative activities?

Catherine Beaudry [a] Stefano Breschi [b*]

[a] *Manchester Business School, University of Manchester, United Kingdom*
[b] *Università C. Cattaneo, Castellanza (VA) and CESPRI - Università L. Bocconi, Milan, Italy*

February 2000

Abstract

This paper examines whether location in strong industrial clusters translates into a higher probability of innovating. A firm-level analysis of the UK and Italy is performed. Innovative activities of firms (European patent data for 1988-98) are related to regional employment, other cluster-specific and firm-specific variables. Clustering *per se* does not explain all of a firm's innovative performance. While location in a cluster densely populated by innovative firms positively affects the likelihood of innovating, strong disadvantages arise from the presence of non-innovative firms, both in own and other sectors industries. Firms' innovative persistence and the regional knowledge stock are also important.

Keywords: Clusters, Innovation, Knowledge, Spillovers

JEL classification: L10, O30, O40, R12

This research has been funded by European Community Contract No. SOE1 – CT97 – 1058 (DG12 – SOLS), *Industrial Districts and Localised Technological Knowledge: The Dynamics of Clustered SME Networking*, which is part of the *Targeted Socio-Economic Research* (TSER) Programme organised by DGXII (Science, Research, and Development). We wish to warmly thank Cristiano Antonelli and Peter Swann for their helpful suggestions and comments, as well as Franco Malerba, who kindly allowed us to use the EPO-Cespri database. Helpful comments from other European partners in this research programme, notably Naresh Pandit, Gary Cook, and the earlier work of Rui Baptista developing econometric models of innovation in clusters are also gratefully acknowledged. None of these, however, are responsible for any remaining errors.

---

* Corresponding author: Stefano Breschi. Address: Via S. Mansueto 4, 21053 Milano. Telephone (+39)-02-

58363365: Fax (+39)-02-58363399. E-mail: stefano.breschi@uni-bocconi.it

1. Introduction

The late 1980s, and even more so the 1990s, have witnessed a resurgence of interest in the economics of industrial clustering. While initially fuelled by the popularisation of a number of success stories of regional and urban boom or revitalisation (Porter, 1998), this new wave of studies has increasingly focused upon a number of theoretical as well as methodological issues. In particular, international trade, growth and industrial economists have resurrected Marshall's (1920) traditional list of agglomeration forces rediscovering what urban and regional economists have long taken for granted, and applied more sophisticated and rigorous (mainstream) modelling tools.

The observation that innovative activities are strongly geographically agglomerated both in Europe (Cäniels, 1999; Breschi, 1999) and the US (Jaffe, 1989; Feldman, 1994; Audretsch and Feldman, 1996) has thus led many researchers to investigate the likely causes of this phenomenon. The literature on national and regional systems of innovation has stressed the role of public authorities as providers of resources and institutions that spur or support firms' innovative effort, and provided some evidence on the role of specialised suppliers and skilled labour as sources from which originates a continuous flow of incremental innovations (Nelson, 1993; Howells, 1999). The force that has attracted most research efforts and dominated much of the debate among economists is, however, related to the notion of knowledge externalities and spillovers. Spurred by the theoretical advances in new growth economics (Romer, 1986; Lucas, 1988), industrial economists have identified the source of agglomeration of innovative activities in the existence and effects of geographically localised knowledge externalities. Starting from the result that firms' innovative efforts do not proceed in isolation, but are supported by external sources of knowledge (Kline and Rosenberg, 1986; Dosi, 1988), it has been argued that firms which are located close to these sources will enjoy relative advantages over more distant firms and consequently tend to have a higher innovative performance. The reason for this advantage must be searched in the properties of knowledge relevant to firms' innovative activities (Winter, 1987). As a matter of fact, this knowledge is largely tacit, uncodified and informal, and both the transmission and the acquisition of knowledge is therefore affected by the geographical proximity of agents. A parallel empirical literature has also flourished, trying to measure knowledge spillovers, assess the extent of their localisation, and evaluating their impact on regional innovative performance (Jaffe *et al.*, 1993; Audretsch and Feldman, 1999; Kelly and Hageman, 1999).

Despite this widespread enthusiasm, a few authors have nevertheless guarded against the risk of over-emphasising the concept of knowledge spillovers and the importance of unbounded increasing returns. After all, the production of goods and innovations is not concentrated in a single location. Any meaningful model of economic geography must therefore recognise the tension between forces that tend to promote geographical concentration (centripetal) and those that tend to oppose it (centrifugal) (Krugman, 1998). Moreover, the emphasis upon pure knowledge externalities also leads to something that looks too much like the tautological argument that agglomeration takes place because of agglomeration economies. The workings of a larger set of effects should be considered in models of economic geography.

Building upon this stream of literature, the purpose of this paper is then to examine empirically in what sense and to what extent 'clustering' is really beneficial to firms' innovative activities. A geographical cluster is here defined as a strong collection of related companies located in a relatively small geographical area. The paper follows the methodology used in a number of earlier studies (Swann and Prevezer 1996; Swann *et al.*, 1998, Baptista and Swann, 1998). These have modelled firms' entry, growth rates and innovation as a function of the strength of the cluster in which they are located, and explored whether strong clusters tend to attract a disproportionate number of new start-up firms, and are responsible for a disproportionate share of innovations. These earlier studies reached a number of important conclusions, notably that firms located in clusters that were strong in their own industry tended to grow faster, and to introduce a greater number of innovations than more isolated firms. On the contrary, firms in clusters that were strong in other industries did not grow faster, and presented lower propensities to introduce innovations, thus indicating the existence of some kind of congestion effects related to the spatial concentration of many firms in a diversity of industries. The conclusions of these studies, however, relate only to a restricted group of high-technology industries, and to the UK and US; they may not be appropriate in a wider context and in other countries. The purpose of this study is then to explore the same hypotheses, and to assess the extent to which they carry over to other industries and other countries.

The objective of this paper is therefore to determine whether firms located in strong clusters (using regional employment as a measure of the cluster strength) are more likely to innovate than other firms. This will be pursued by performing a firm-level analysis of patent counts for the manufacturing sectors in Italy and the UK.

The paper is organised as follows. Section 2 reviews the main theoretical arguments as to why firms located in clusters may be more innovative. Section 3 provides a description of the data sources and the empirical approach adopted in this study, while Section 4 summarises the econometric approach used. Section 5 reports estimation results. Finally, section 6 discusses the significance and interpretation of our results.

2. Theoretical background

This section reviews the theoretical arguments found in the literature which support the hypothesis that location in a strong industrial clusters enhances firms' innovative performance. Factors limiting the spatial clustering of innovative firms are also considered, as well as sources of industry and national variation in spatial patterns of innovation.

2.1 Forces leading to the agglomeration of innovative activities

In general terms, the benefits that lead to clustering can be divided into demand and supply sides (Swann, 1998). On the demand side, firms may cluster to take advantage of strong local demand, particularly that deriving from related industries. Moreover, consumer search costs and demonstration effects arising from the observation of successful firms at a particular location might also be important determinants of agglomeration. In the realm of innovation, it has also been pointed out that customers represent important sources of new ideas and that a continuous flow of incremental innovations is generated through the localised user-supplier interaction (von Hippel, 1988; Lundvall, 1993).

The most relevant benefits from clustering for firms' innovative activities, however, are likely to originate on the supply side. A localised industry can support a greater number of specialised local suppliers of industry-specific intermediate inputs and services, thus obtaining a greater variety at a lower cost. Furthermore, a localised industry attracts and creates a pool of specialised workers with similar skills, which benefits both the workers and their employers, smoothing the effects of the business cycle through the effects of increasing numbers (David and Rosenbloom, 1990).

The most fundamental supply side externality, however, is related to knowledge, and particularly to knowledge spillovers. In the pursuit of innovation, firms are not isolated, but they are embedded in networks of relationships with users, suppliers, competitors, universities, public research centres, and other public and private institutions (Kline and Rosenberg, 1986; Nelson, 1993). These agents are repositories of scientific, technical as well as market knowledge, which is a vital input for the success of a firm's innovative effort by reducing the degree of technical and commercial uncertainty and facilitating problem-solving tasks. These networks of innovators are often highly localised, because geographical and cultural proximity helps to build the codes of communication, social norms, institutions and trust, which are essential to generate smooth and effective flows and exchanges of knowledge among agents (Storper, 1992; Freeman, 1991).[1]

In addition, geographical proximity also plays a fundamental role in mediating the diffusion and acquisition of knowledge spillovers, i.e. positive externalities of technical and scientific discoveries on the productivity of firms which neither made the discovery themselves nor licensed its use from the holder of intellectual property rights. The classic story states that by being near to universities, where leading-edge research is carried out, and to other

---

[1] Alternatively, one can also argue that proximity matters because there is a high probability that *initial* business contacts will be established in the local environment, and that these initial contacts will develop into a strong local network.

innovative firms, employees of local firms will hear of important discoveries first and thus be able to utilise them before others are even aware of their existence. Furthermore, the reason geographical distance is likely to reduce the ability to receive relevant knowledge has to be found in the nature of knowledge that leaks out of universities and firms. In this respect, a distinction is normally made between *information*, which can be easily codified and has a singular meaning and interpretation, and *knowledge*, which is more vague, difficult to codify, and often only serendipitously recognised. While the former can be transmitted at low (marginal) cost over long distances, the latter is best transmitted via face-to-face interactions and through frequent and repeated contacts, all of which are most efficiently managed within local proximity (Audretsch, 1998). This implies that inventors and firms operating in a specific industry and located close to each other will be more innovative than those who are geographically isolated because of the greater likelihood of sharing tacit knowledge and having access to the scientific discoveries at local universities and public research centres. A vast empirical literature has developed in recent years to test this hypothesis (Jaffe, 1989; Jaffe *et al.*, 1993; Feldman, 1994; Audretsch and Feldman, 1996, 1999; Mansfield, 1995). The most important result of this empirical effort has been to show the existence of a statistically significant effect on a firm's innovative record of being near great universities and other sources of scientific discovery.

A further supply-side advantage from clustering is related to the importance of learning-by-doing and learning-by-using (Malerba, 1992). The larger the industrial size of a cluster, the higher the opportunities and the stimuli to experiment with new techniques, because of the *technological bottlenecks* arising in existing productions and the interdependent nature of technical improvements (*technological convergence*). This means that workers can learn how to produce, use, as well as improve goods and machinery by solving production problems, meeting customers' needs, and overcoming technical unbalances (Pred, 1966).[2] Moreover, location in a large cluster will also increase the likelihood of acquiring information on new techniques and thus stimulate learning-by-imitating and learning-by-adopting.

Finally, any advantage from clustering is likely to become cumulative and self-reinforcing (Arthur, 1990). The reason is that firms' current technological efforts strongly depend and build upon previous scientific advances and technical achievements. Innovation is a highly cumulative activity. This implies that firms located in regions which have accumulated high levels of innovative success and possess a relevant stock of knowledge will be relatively advantaged in the next round of innovations compared to other firms. In other words, what matters more for firms' innovative activities is not only (and not much) the current scale or size of an industry in a given area, but the *accumulated* stock of knowledge. Regions that first emerge as centres of innovative activity in a certain industry tend to maintain their advantage over time. Again, this is not an entirely new idea. As Thompson (1962, p.260) wrote some time ago: "The proposition that the production centers of an industry are also its research centers seems plausible, at least on first blush. The reasoning here is that new products and techniques are largely spawned by persons who work for or are otherwise closely associated with the industry most closely linked with the particular idea or device. To the extent that this is

[2] A related theory argues that regions where an industry is experiencing new investment or rapid growth in production will tend to generate higher levels of inventive activity, because of higher opportunities or pressures in such regions to experiment with new techniques (Pred, 1971).

so, a substantial, persistent, and even cumulative advantage would accrue to any region which gained a head start in a particular industry; perhaps technological differentials tend to widen – the rich grow richer".

2.2 *Is clustering always beneficial to firms' innovative activities?*

A common element of these explanations for the clustering of innovations is that the geographic location of inventive activities in an industry is a function of the location of production in that specific industry as well as in related industries. Out of the enthusiastic chorus for localised knowledge spillovers and increasing returns, a few authors have recently pointed out the disadvantages involved by locating in a geographical cluster.[3] Many of these disadvantages are likely to be generic rather than sector-specific. Thus, a company located in the centre of a metropolitan area faces higher property rentals and labour costs than a more isolated firm.

Some congestion effects or external diseconomies are likely to arise on the demand side. Increased local competition in output markets may result in lower profit margins, thereby reducing the amount of resources devoted to R&D. More importantly, greater physical isolation from other producers also entails more limited obligations and weaker relational ties, which under certain circumstances may induce higher flexibility and responsiveness to technical and organisational changes (Suarez-Villa and Walrod, 1997). In other words, location in a

---

[3] One could also question the relative importance of 'pure' versus 'pecuniary' knowledge externalities, as an agglomeration force. For a critical review of the recent literature on localised knowledge spillovers, see Breschi and Lissoni (2000).

cluster may lead to technological lock-ins and resistance to innovation to the extent that firms' locational choices reflect the search for cheap material inputs and other non innovation-related motives.

On the supply side, it has been emphasised that networks of innovators are not necessarily bounded by geographical proximity (Echeverri-Carroll and Brennan, 1999). The innovation process is likely to involve both local and global influences, and the relative importance of these influences varies accordingly to the stage of the innovation process. Global networks tend to be relevant for the conception and the commercialisation of new products, whereas the development work involves mainly localised sources of knowledge. In particular, the most dynamic and innovative firms look for knowledge embodied in engineers and scientists, wherever they are available, and not necessarily constrained in this by geographical barriers. Moreover, local knowledge sources are probably less important for firms located in lower-order regions, i.e. areas with a relatively small accumulation of knowledge. For these firms, local universities are viewed as suppliers of skilled workforce, rather than loci of innovations, sources of product ideas or spillover effects. In order to sustain high rates of innovation they must develop linkages with actors (universities and other high-tech firms) located in higher-order regions (Lyons, 1995).

From a different perspective, Baptista and Swann (1998) and Beaudry (2000) have convincingly shown that whereas strong cluster employment in a firm's own sector significantly improves innovative capabilities, higher levels of employment in other industries can entail negative effects on a firm's innovative performance, thus providing evidence of mild congestion effects on the supply side.

In general, it must be pointed out that the concentration of firms and production in a certain location *per se* is not a sufficient condition to determine high rates of innovative activity. As Lamoreaux and Sokoloff (1997) have quite persuasively shown, inventive activity will tend to concentrate in locations where invention rates *had long been high* and where a *market for technology* has evolved more fully, irrespective of the share of industry production. In this respect, an important role is played by those 'bridging' institutions that provide information about technological opportunities and mediate relations among inventors, suppliers, and those who would commercially develop or exploit new technologies. This implies also that industries may move across regional and national borders without a corresponding relocation of inventive activity. Inventive activity is more 'sticky' than production, possibly because the richness of *generic technological know-how* in higher-order regions serves as an effective substitute for specific knowledge and allows to find new applications across a wide range of industries.

2.3 Sectoral and national specificities

The strength of advantages and disadvantages from clustering, as well as the balance between these forces, is likely to vary among industrial sectors. In this respect, there are two distinct approaches to discuss.

First, as the recent literature on new growth economics has recognised, two types of localised knowledge externalities are at work (Glaeser *et al.*, 1992). On the one hand, some of the advantages from clustering arise from industry specialisation. This happens whenever knowledge externalities exist, but are limited to firms within the same industry. This type of effect has been termed Marshall-Arrow-Romer (MAR) externalities or *localisation* *economies*. The implication of MAR externalities is that concentration of an industry in a location will induce higher levels of knowledge spillovers and therefore facilitate innovation. On the other hand, some external economies arise from diversity or variety between complementary industries. Firms in a certain industry can benefit from innovative ideas, skills, know-how and human capital originating from different, but somehow related industries. This type of effect has been termed Jacobs or *urbanisation economies* (Jacobs, 1969). The implication of this hypothesis is that regions that exhibit a broad and diversified industrial base will also promote firms' innovative activities. At the empirical level, the relative importance of specialisation versus diversity for firms' innovative activities has been recently addressed by Audretsch and Feldman (1999). They reach the conclusion that diversity matters more than specialisation.[4]

A second line of research is associated to the concept of *technological regime*. According to this approach, industries widely differ in terms of sectoral patterns of innovative activity and underlying properties of technologies used. In particular, the conditions of

[4] In our view, the empirical verification of the specialisation hypothesis has incurred some problems of misspecification and generated some confusion. The hypothesis is usually verified by using a relative index – such as the location quotient defined as the share of employment of industry i in region j divided by the share of employment of the same industry in the country – *without* controlling for the size of the industry i itself in region j. This methodology neglects the fact that a relatively small region can be highly specialised in industry i, to the extent that a large share of workers in region j are employed in industry i, even though the absolute size of the industry is much smaller than in another region, which is not specialised in that industry. We suspect that the lack of any significant effect between specialisation and innovative performance may be due to the failure to control for this effect.

technological opportunity, appropriability and cumulativeness of technical advances, and the properties of the knowledge base, differ across industrial sectors, thus determining different patterns of innovation (Breschi, Malerba and Orsenigo, 2000). Moreover, the basic features defining a technological regime are also likely to have a spatial dimension and thereby have consequences for the geographical distribution of innovative activities (Breschi, 2000). If technological opportunities affect the rate of innovation, then the spatial location of innovators will be affected by the location of such opportunities (universities, public research centres, users, suppliers) and by the nature of the relevant knowledge base. The latter determines how information about scientific discoveries and new technologies is transmitted between agents and therefore defines the spatial boundaries within which this transmission can effectively take place. Broadly speaking, the more the knowledge base is tacit and non-codifiable, the higher the spatial concentration of innovators one can expect. This type of knowledge is better transmitted through informal means and interpersonal contacts, whose effectiveness sharply decreases with the geographical distance between agents.[5] Conversely, the more codified, simple and independent is the relevant knowledge base in a sector, the less important is the role of geographical distance in mediating knowledge flows. Appropriability and cumulativeness conditions also affect the spatial distribution of innovators. Industries with a higher level of appropriability and cumulativeness at the firm level will also be associated with strong

[5] The fraction of knowledge base that is tacit and non-codifiable is especially high for industries and technologies that are in the early stages of their life-cycle, when knowledge is still highly complex and ever-changing. However, the importance of tacit know-how can be high also in relatively mature industries (e.g. mechanical engineering), where the innovation process involves idiosyncratic capabilities to 'design' products that fit customers' specific requirements.

selective pressures, since in these circumstances, technological leaders are more likely to innovate further and keep their competitive advantage. This implies that a relatively high level of sectoral concentration can be expected and therefore, with a lower number of firms, geographical concentration of innovative activities is also more likely to emerge.

Finally, one can expect that the industrial structure, institutional system and history of industrial development of each country will affect in fundamental ways the spatial distribution of production and innovation, and the actual workings of agglomeration economies. Thus, for example, the post-war history of industrialisation explains an important part of the current spatial distribution of innovative activities in Italy. Similarly, industrial policies adopted in the UK during the 1970s and the 1980s are largely responsible for the relocation of most innovative activities in the South-East.

3. Data and empirical approach

This paper combines three sources of data: patent data, company data, and regional employment data. These three sources of data, the specific issues associated with their merger, and the variables built from them are examined in detail below.

3.1 Data sources

The first set of data used in this study is the EPO-CESPRI[6] database which provides information on patent applications to the European Patent Office (EPO) since 1978 to 1998 of

[6] Center for Research on Internationalisation, Bocconi University, Milan.

firms from two countries: Italy and the UK.[7] For each patent document, the EPO-CESPRI database contains information on: (i) the name and the address of the patenting firm; (ii) the date of filing to the EPO; (iii) the technological field which was assigned by patent examiners.[8] For this study, the patenting firm's address reported in the patent document has been used to locate it in space. As far as the spatial unit observation is concerned, *Nomenclature of Statistical Territorial Units* (NUTS) level 3 regions have been adopted here. According to the definition provided by the European Office of Statistics (Eurostat), this level referred in 1991 to 65 *counties* for the UK and 95 *provinces* for Italy. It is recognised that the use of the applicant's address to locate patents in space is not entirely satisfactory due to the widely diffused practice of firms' headquarters to patent inventions which have been originally developed by divisions and manufacture establishments located in different regions.[9]

---

[7] The EPO-CESPRI database also contains data for Germany, France, Japan and United States. These countries have not been considered in the present analysis because of time and budget constraints, but are in the pipeline for future research. Please note also that the processing of UK patent data to 1998 was not completed in time for this study and the data stops in 1994 for this country.

[8] The EPO-CESPRI database has been constructed at the level of individual firms and institutions. Firms that are part of business groups have been treated in the present analysis as individual companies. In case of co-patenting, each co-patentee has been credited the patent. Individual inventors have been excluded from the dataset. Overall, the EPO-CESPRI database includes 39,582 patents and 7,121 firms for the UK, and 25,058 patents and 6,265 firms for Italy.

[9] Only in few cases the applicant name reported in patent documents refers to manufacture divisions or establishments where the invention has originated. An alternative approach, which has been followed in the empirical literature, is to use the inventors' address, instead of applicant's address. Despite the specific problems associated with this approach, we decided to adopt the applicant's address because the

Particularly, this approach can lead to an over-estimation of the volume of innovative activities carried out in large metropolitan areas within each country, where most headquarters are located. While this problem is not easy to solve, there is a number of reasons that can help to mitigate the resulting bias.[10] First of all, misattributions of patents to the company headquarters cluster rather than another cluster are likely to be most serious only in the case of larger firms (which, however, are a minority in this database) and in certain industries, where multi-plant firms are important. Second, it has been shown that many large firms tend to locate their R&D facilities close to company headquarters and do not disperse them throughout the corporation (Howells, 1984, 1990). This implies that as long as a greater proportion of patents can be effectively considered as flowing from basic and applied research activities (i.e. from R&D laboratories), then the extent of the distortion is likely to be further lessened. Third, it can also be argued that any potential innovation has to pass through the company headquarters before it is patented (e.g. through internal mobility of researchers), so that some kind of knowledge spillover is likely to benefit the company headquarters even if the invention has been originally developed elsewhere.

---

central question of our study is whether firms located in strong clusters are more likely to innovate than firms outside these regions.

[10] If headquarters tend to be located in stronger clusters than other divisions and establishments of the company, then any misattribution of patents to the company headquarters cluster can lead to an upward bias in the effect of cluster employment on the probability of patenting. However, for the reasons given in the text, it is probable that this bias is not particularly large.

The advantages and limitations of patent indicators are well known so we will not review them here.[11] We just point out that, although not perfect, patents represent an extremely valuable source of data for the spatial analysis of innovative activities. First, by containing the address of the inventing firm, they permit to map the spatial structure of technological activities at a level of geographical detail that no other indicator to date has been able to provide. Second, patents represent a very homogeneous measure of technological novelty, are available for long time series, and provide very detailed data at the firm level, which make them suitable for comparing the innovative activities of firms located in clusters of different countries.

The second set of data used in this paper is company information. Two commercial company databases were used to extract company economic data: Dun and Bradstreet's OneSource *UK vol. 1 and 2* for the UK and Bureau Van Dijk's *AIDA* for Italy.[12] Three categories of company information were considered for this study: firm size, primary sector of

---

[11] For a recent discussion on the use of patents as economic indicator see Griliches (1991).

[12] It is important to note that both databases include a sample of all manufacturing companies active in each country. *AIDA* provides balance sheet data of all Italian companies with an annual turnover higher than 2 million Euros, and of a significant proportion of companies with an annual turnover higher than 1.5 million Euros. Overall, the release of *AIDA* used for this study (28, June 1998) contained economic information for 48,216 manufacturing firms. OneSource *UK vol. 1 & 2* (release of September 1996) on the other hand, provides very detailed data on 360,000 UK companies, but applies a more complicated cut off point to choose which firms to include. In total, *UK vol. 1 & 2* provide information on 60,306 manufacturing firms.

activity and region.[13] Firm size was measured as the number of employees in year 1996.[14] Each company was assigned, according to its main activity, to one industry sector, each corresponding to a two-digit *UK Standard Industry Classification* (SIC) (1980 Rev.) industry for the UK, and to a two-digit *Nomenclature of Economic Activities in the European Community* (NACE) (Rev. 1) industry for Italy. For the present study, we considered 15 manufacturing sectors for the UK and 17 manufacturing sectors for Italy. Note that it was necessary to aggregate up to the two-digit level since a large proportion of Italian companies in *AIDA* were classified at this level. The list of industries considered in this study and a correspondence between the two industry classifications is reported in the appendix (see Table A1). Finally, for each company, the NUTS 3 level region in which its headquarters are located was also identified.

The third set of data used in this paper is the employment by NUTS 3 level region for the UK and Italy. These data are provided by the *Central Statistical Office* (CSO) for the UK and by *Istituto Nazionale di Statistica* (ISTAT) for Italy, and refer to the 1991 census. For the present study, regional employment data at the two-digit UK SIC (1980 Rev.) level for the UK and at the two-digit NACE (Rev. 1) level for Italy were used.

---

[13] In addition to these, a fourth category of information, only available for the UK, refers to the status of the company, and whether it files consolidated accounts.

[14] Of course, the current set-up of data can imply a simultaneous relationship between patents and company employment. Also note that by choosing companies for which firm employment was available, the sample is reduced by less than half for the UK.

3.2 *Databases merger*

In addition to the specific problems associated with each database, some important issues appeared while merging the three databases. In a first instance, we had to match the EPO-CESPRI database containing the names of patenting companies with the business databases OneSource *UK* and *AIDA*, containing economic information on companies. We successfully found economic data for 1,091 patenting firms for the UK, and 2,616 patenting firms for Italy. This corresponds to 7,905 patent applications for the UK and 12,142 for Italy.[15]

Our final sample includes 23,872 and 37,724 manufacturing firms for the UK and Italy respectively. Please note that the vast majority of these firms did not patent (see Table 1). The proportion of firms with no patenting activity is higher in the UK (95.4%) than in Italy (93.1%). Moreover, Italy also shows (in this sample) a higher proportion of firms with one or two patents compared with the UK.

[Place Table 1 approximately here]

[15] In percentage terms, we found economic data for 27.4% of all firms that obtained patents in the period 1988-94 for the UK, and 52.7% of all firms that obtained patents in the period 1988-98 for Italy. This corresponds to 42.0% of all patent applications for the UK and 65.6% for Italy, over the same periods of time. The merger of the two databases was carried out manually and presented several difficulties that partly explain the relatively low matching ratio. First, there was no common pattern in naming companies across databases. Second, the number of firms included in company databases is limited and consequently many small innovative companies are simply not reported in these databases. Third, patenting firms that have ceased to exist are not reported either and therefore cannot be matched.

Another important issue concerns the technology classification of patents. Indeed, all patent documents are assigned by patent examiners of the EPO according to one main classification code of the *International Patent Classification* (IPC).[16] It is important to note that these technology codes do not correspond directly to any UK SIC or NACE industry codes. For this reason, in this paper we sum up all patents of each company into a single value as a first approximation and ignore the distinction of patents into technological fields as well as the possible correspondence between technology fields and industry codes.[17] In what follows, however, we will include as an explanatory variable the extent of technological diversification of each cluster. For this purpose, we have used a classification of IPC codes into 30 technology fields proposed by Fraunhofer Gesellschaft-ISI (Karlsruhe).[18]

3.3 *Variable definitions and descriptive statistics*

The variables used for the subsequent econometric analysis are described in Table A2 in appendix. At this stage, it is important to focus on the variables that represent regional (cluster) industrial strength. Following Baptista and Swann (1998), Swann and Prevezer (1996), Swann (1998) and Beaudry *et al.* (2000), this was measured by the logarithm of sector employment in a firm's own industry (OWNEMP) and employment in other industries

---

[16] The IPC is an internationally agreed, non-overlapping and comprehensive patent classification system. Currently, the IPC (6th ed.) refers to almost 60,000 individual codes (12-digits) and it may be used at different hierarchical levels (WIPO, 1994).

[17] It is extremely difficult to evaluate the technological fields that should be counted as being related to the main sector of activity of a firm, to a secondary sector of activity or not related at all.

[18] A version of the classification is reported in OECD (1994).

(OTHEMP). The main rationale for doing this is that, if the arguments of cluster-specific knowledge spillovers (i.e. localisation and agglomeration externalities) are true, then there must be a positive relationship between a firm's innovation output and the scale of industries in the cluster or region. A relative measure, such as the proportion of sector employment in the region's total employment, would not serve our purpose as well, by neglecting the fact that a given region might represent a strong cluster in a certain industry, even if this industry accounts for a negligible share of the region's overall range of activities.

At the same time, it is quite reasonable to argue that not all employees generate equal spillovers. In particular, one can assume that the employees of innovative companies are likely to generate higher levels of, as well as better quality, knowledge spillovers than the employees of non-innovative companies. Following this argument, we distinguished between employment of innovative firms (defined as those firms that obtained patents in the period 1978-98) (OWNINN) and employment of non-innovative firms (OWNNOINN) in a firm's own industry (see Table A2). Similarly, we distinguished between employment of innovative firms (OTHINN) and employment of non-innovative firms (OTHNOINN) in other industries.

4. Econometric approach

The dependent variable in the model is the total number of patents produced by firm n, active in industry i and located in cluster c, over the period 1988-94 for the UK and 1988-98 for Italy ($INNOV_n$). Because this is a limited dependent count variable, where the large majority of observations is zero (see Table 1), a simple ordinary least-squares regression analysis would yield biased results. For this reason, this study focuses on the family of linear exponential models, such as the Poisson (Hausman *et al.*, 1984; Crepon and Duguet, 1997) and the negative binomial regression models, that are more appropriate for count data (Greene, 1997, Hausman *et al.*, 1984). [19]

The specification of the model follows quite closely that used by Baptista and Swann (1998) and Swann and Prevezer (1996). The right-hand side of the model includes, besides the variables measuring regional (cluster) industrial strength, a variable related to firm size (CIEEMP)[20]. Even tough the empirical evidence on the impact of firm size on innovation performance has so far been inconclusive (Cohen, 1995), we included this variable in our model in order to avoid possible misspecifications.

The model was estimated both with reference to all industries pooled together and for each industry separately considered. In the pooled analysis, we added constant fixed effects through industry dummies. Tests on the significance of these effects showed their importance.

---

[19] A major drawback of the Poisson model is that the conditional mean is assumed to be equal to the conditional variance, so that any cross-sectional heterogeneity is ruled out. However, this restriction is normally violated in most economic phenomena, resulting in problems of overdispersion, i.e. the variance exceeds the mean (conditional on covariates). The negative binomial model provides a generalisation that permits to solve the problem, by introducing an individual unobserved effect into the conditional mean (Greene, 1997).

[20] A potential problem arises when companies file consolidated accounts. Indeed, when a holding company files consolidated accounts, and its subsidiaries appear in the database alongside the parent company, double counting of employees occurs. This problem was especially serious for the UK. For this country, dummy variables for holding companies and consolidated accounts were therefore introduced to test the extent of the problem of double counting of employees and their influence was judged minimal.

Thus, the simplest model to be estimated has the following form:

$$\begin{aligned} INNOV_n &= \alpha + \beta_1 \ln e_n + \beta_2 \ln\left(\sum_i e_{ic}\right) + \beta_3 \ln\left(\sum_{j \neq i} e_{jc}\right) + \sum_{i=1}^{I-1} \gamma_i D_i \\ &= \alpha + \beta \ln(CIEEMP_n) + \beta_2 \ln(OWNEMP_{ic}) + \beta_3 \ln(OTHEMP_{jc}) + \sum_{i=1}^{I-1} \gamma_i D_i \end{aligned} \tag{1}$$

where $CIEEMP_n$ is employment of firm n, active in industry i and located in cluster c, $OWNEMP_{ic}$ is cluster (c) employment in firm n's own industry i, $OTHEMP_{jc}$ is cluster (c) employment in all other industries j, and D_i are industry fixed effects.

Five variables are added subsequently to the benchmark specification of the model. A first group of firm-specific variables aims to capture some of the individual heterogeneity that may be present in the data. The age of the company (AGE) was included to assess whether older or more recently established firms are better equipped to conduct innovative activities. As with firm size, existing empirical evidence on the effect of age on innovation performance is somewhat inconclusive. At best, one can argue that the relationship between company age and innovation is likely to vary accordingly with the characteristics of the technological environment and the stage of the industry life cycle (Acs and Audretsch, 1990; Klepper, 1996). A more robust result emerged in the recent empirical literature is that persistence and cumulativeness are key features of firms' innovative activities (Geroski *et al.*, 1997; Malerba, Orsenigo and Peretto, 1997). Firms that have successfully innovated in the past are thus more likely to innovate again in the future compared to less successful or non-innovative firms. To model this effect, we included a dummy variable that is set to one if a firm has previously patented, and zero if it has not (PATPREV).

A second set of cluster-specific variables is also included to capture other effects on firms' innovative performance. An indicator of employment dispersion (EMPHERF) across industries within each region was included in order to measure industry variety in a region. This measure is simply the Herfindahl index of employment in all two-digit manufacturing sectors within each region. It should not be neglected, however, that this measure captures in a very imperfect way the notion of agglomeration or Jacob's externalities (Jacobs, 1969). On the one hand, it rules out any complementarity between industries, assuming that all sectors are equally close to each other. On the other hand, it is also likely that at the level of industry aggregation considered in this study (two-digit), most agglomeration externalities take place within and not across industries. Similarly, a measure of technological variety within each region was included in the model as the Herfindahl index of patents across 30 technology fields (PATHERF). The same *caveats* apply, of course, to this indicator.

Finally, a last variable was included to control for the innovative size of each cluster. A firm's probability of patenting is likely to be a function of the accumulated stock of knowledge as well as the level of (localised) knowledge spillovers that benefit its innovative activities. In order to capture this effect we included in the model the total number of patents produced within a region in the period 1978-87 (STOCKP).

5. Estimation results

5.1 Pooled results

Equation (1) has been estimated[21] using both Poisson and negative binomial regression models, which are most appropriate to handle count data (Hausman *et al.*, 1984; Crepon and Duguet, 1997; Greene, 1997). The two methods give qualitatively similar results. In what follows, we therefore report estimates only for the negative binomial model. Estimates of equation (1), for all industries pooled together, are reported in Table 2.

[Place Table 2 approximately here]

As expected, firm-specific variables have a highly significant explanatory power and their sign is strongly robust. Throughout all specifications, the coefficient on firm size (CIEEMP) is positive and statistically significant, thus indicating that (in our sample) large firms produce on average a larger number of patented innovations than small and medium sized enterprises.[22] Similarly, the coefficient on the variable PATPREV is positive (well above one) and statistically significant in all specifications both for the UK and Italy, therefore confirming the highly cumulative and persistent nature of innovative activities. Firms with a higher stock of

[21] All estimations were obtained using Proc Genmod in SAS 6.12.

[22] We carried out estimations of the model including firm age. Generally speaking, the coefficient of this variable is negative and statistically significant both for Italy and the UK, thus implying that (in our sample) relatively younger firms have a higher likelihood of innovating than firms established far in the past. Moreover, the introduction of this variable does not affect the sign and significance of all other variables. However, since the information on age is missing for more than 2,500 firms of the sample for Italy, we present only results without this variable. Results including *age* are available upon request.

knowledge tend to generate a higher number of innovations than firms that have not patented in the past.

Concerning cluster employment variables, results for the United Kingdom confirm previous findings by Baptista and Swann (1998). The coefficient of the variable measuring own sector employment (OWNEMP) is always positive and statistically significant, thereby suggesting, that firms located in clusters that are strong in their own industry tend to have a better innovative performance than more isolated firms. In contrast, the effect of employment in other industries (OTHEMP) is always negative and significant. A possible interpretation of this result is that congestion effects are likely to offset any spillover advantage deriving from the presence of other industries.[23] It is interesting to note that this result is fairly robust with respect to the introduction in the model of the accumulated stock of patents (STOCKP). Indeed, the negative effect of employment in other industries is amplified, implying that congestion effects are stronger than previously thought. Furthermore, this interpretation seems to be supported also by the lack of any significant impact of industrial diversity indicator (EMPHERF).

Results are somewhat more puzzling and less unequivocal in the case of Italy. The coefficient of own-sector employment (OWNEMP) has a positive sign, but is not statistically significant before the introduction in the model of the stock of patents (STOCKP), and even

[23] In interpreting these results, one should also keep in mind that the two-digit level of aggregation used here is probably inadequate to draw conclusions on cross-sectoral effects. As a matter of fact, it is likely that the presence of certain industries should advantage innovative activities, whereas the presence of others only leads to congestion (Baptista and Swann, 1998).

then becomes negative. Once we control for these effects, the coefficient on this variable turns negative and statistically significant. Similarly, no congestion effects arising from the presence of other industries (OTHEMP) are observed for Italy prior to the introduction in the model of the stock of patents. However, after controlling for the technological size of each cluster, the coefficient on this variable becomes negative and its magnitude increases, therefore supporting the hypothesis that negative effects are likely to offset any positive spillover deriving from the presence of other industries within a cluster. These results are apparently in contrast with that found for the UK and with previous findings by Baptista and Swann (1998), who found moderately large positive effect of own sector employment on the probability of a firm to innovate. On the one hand, one could argue that in Italy the location of innovative activities follows less closely production activities than in the case of the UK. On the other hand, the fact that the signs of both employment variables change after controlling for the stock of patents points out the importance of the accumulated stock of knowledge within a cluster and suggests that the specification adopted is probably not adequate to capture this effect, especially in Italy.

In order to investigate further the effect arising from the location in a cluster of a large set of innovative firms, we estimated a more complex model distinguishing between cluster employment of innovative companies and cluster employment of non-innovative companies in a firm's own industry and in other industries. Results are reported in Table 3[24], and show that in both countries cluster employment of innovative firms in a firm's own industry (OWNINN)

---

[24] Once again, the results for the Poisson and negative binomial regressions are very similar and only the latter are presented here. They are however available upon request.

affects in a positive and statistically significant way firms' innovative performance, while a negative and statistically significant effect is associated with a strong presence of non-innovative firms (OWNNOINN). This result is fairly robust throughout all the specifications, particularly in the case of Italy, suggesting that intra-sectoral positive externalities are likely to flow locally only from innovative firms, whereas the presence in a cluster of non-innovative firms is associated with negative (congestion) effects. In other words, the benefits from clustering with other firms in the same industry are not generic, instead they arise only in clusters that are already densely populated by innovative firms and have a large accumulated stock of knowledge. The lack of evidence of any significant impact of own-sector employment (OWNEMP) in the case of Italy can thus be interpreted as a consequence of a misspecification problem in the basic model. Imposing the restriction that the coefficients of OWNINN matches that of OWNNOINN in the case of Italy is clearly wrong.

[Place Table 3 approximately here]

Concerning employment in other industries, results show instead a persistent difference between the two countries examined here. In the case of the UK, while greater employment in innovative firms in other industries positively affects firms' innovative activities, a negative impact is associated with the location in a cluster of non-innovative companies in other industries. Moreover, once we control for the technological size of each cluster, both coefficients on other industries employment are negative and statistically significant. This result thus reinforces previous evidence on the existence of rather strong congestion effects arising from the co-location of a diversity of industries within a cluster. In contrast, in the case of Italy, any negative impact deriving from the co-location in a cluster of other industries seems to derive in this case from the presence of other innovative companies (OTHINN), whereas the presence of other non-innovative companies (OTHNOINN) apparently encourages firms' innovative activities. Even though the current set up of model does not allow to draw any definitive conclusion, we hazard the following interpretation for this result. In Italy, what is really important for firms' innovative activities is to be co-located with firms in 'related' and 'complementary' other industries, no matter whether they are highly innovative or not. In other words, it is possible that localised networks and users-producers interaction are much more important engines of innovation in the case of Italy than for the UK. A natural extension of this research is therefore to test this hypothesis.

Further important differences between the two countries considered here are found in the coefficients of the sectoral fixed effects. In the case of the UK, the probability of firms' patenting is significantly higher in chemicals, instruments and motor vehicles. In the case of Italy, a positive and statistically significant coefficient can be observed for a larger number of sectors, i.e. chemicals, plastics, mechanical engineering, electrical engineering, and instruments. In the next section, we present regression estimates for each industrial sector.

5.2 *Results by industry*

The significance of the sectoral dummy variables suggests to separate the sample by industry. Table 4 reports results of negative binomial regression[25] analysis. For the sake of simplicity, results have been summarised indicating the sign and significance of each relevant coefficient.

---

[25] Ibid. [24]

[Place Table 4 approximately here]

Once again, a consistent result throughout the analysis is the strong significance of firm-specific variables. Both company employment (CIEEMP) and previous innovative performance (PATPREV) have a positive and statistically significant coefficient in most sectors considered here. This supports the idea that an important explanatory factor of firms' current innovative performance is the extent to which firms persistently innovate and grow in size.

Concerning the impact of cluster-specific variables, our results show that for most sectors, firms' innovative activities are facilitated by the co-location of other innovative companies (OWNINN) in a firm's own industry, while a strong presence of non-innovative firms (OWNOINN) hinders firms' innovative performance, especially in the case of Italy. In contrast, a strong presence of firms in other industries (OTHEMP) seems to negatively affect firms' innovative performance, in several sectors. Our industry results point out in any case some important differences between the two countries. Generally speaking, negative effects associated with a strong presence of non-innovative companies in a firm's own industry are higher for Italy than for the UK, whereas congestion effects arising from the co-location of a diversity of industries seems to be more relevant for the UK than for Italy. Finally, results confirm that firms located in clusters with a large accumulated stock of knowledge (STOCKP), are also more likely to innovate than firms located elsewhere.

6. Conclusions

The central question of this paper was whether location in a strong industrial cluster really facilitates firms' innovative activities. The main result emerging from a firm-level analysis of patent counts for Italy and the UK is that clustering *per se* is not a sufficient condition to explain a firm's innovative performance. Whereas location in a cluster densely populated by innovative companies in a firm's own industry positively affects the likelihood of innovating, quite strong disadvantages arise from the presence of non-innovative firms in a firm's own industry. Innovative persistence at the firm-level and the accumulated stock of knowledge within a cluster are additional explanatory forces of innovative performance to simple clustering. We interpret these results as evidence that positive knowledge externalities are likely to flow locally only from innovative companies. In addition, clusters having a higher number of innovative companies and a larger stock of knowledge accumulated in the past are more likely to have a better innovative performance, irrespective of their overall share of employment in an industry. While these results are sufficiently similar for the two countries examined here, some important differences emerge with respect to the role played by employment in other industries. While in the case of the UK, results show the existence of strong congestion effects associated with the location in a cluster of diversified industries, no such evidence is found for Italy. The hypothesis left for future research is that, in the case of Italy, localised networks of 'related' and 'complementary' industries are more important engines of innovation than in the UK.

Appendix

Table A1 - Correspondence between the two-digit UK SIC (1980 Rev.) codes and the two-digit NACE (Rev. 1) codes

UK SIC 1980	United Kingdom	NACE Rev. 1	Italy
22	Metal manufacturing	27	Metal manufacturing
24	Manufacture of non-metallic mineral products	26	Manufacture of non-metallic mineral products
25	Chemical industry	24	Chemical industry
31	Manufacture of metal goods NEC	28	Manufacture of metal goods excluding machinery
32	Mechanical engineering	29	Mechanical engineering
33	Manufacture of office machinery and EDP equipment	30	Manufacture of office machinery and EDP equipment
34	Electrical and electronic engineering	31	Electrical and electronic engineering
		32	Telecommunication equipment
35	Manufacture of motor vehicles and parts	34	Manufacture of motor vehicles and parts
36	Manufacture of other transport equipment	35	Manufacture of other transport equipment
37	Instrument engineering	33	Instrument engineering
43	Textile industry	17	Textile industry
46	Timber and wooden furniture industries	20	Timber industry (excluding furniture)
47	Manufacture of paper and paper products;	21	Manufacture of paper and paper products
		22	Printing and publishing
48	Processing of rubber and plastics	25	Processing of rubber and plastics
49	Other manufacturing industries	36	Other manufacturing industries

Table A2 - Variable definitions and descriptive statistics

Variable	Definition	Descriptive statistics[a] UK	Italy	Data sources
INNOV	Total number of patents produced by a firm over the period 1988-94 for the UK, and 1988-98 for Italy	0.331 (9.411)	0.321 (6.696)	European Patent Office[b]
CIEEMP	Firm employment (1996)	229.136 (1873.19)	63.77 (480.40)	Dun's and Bradstreet OneSource UK vol. 1 and 2 and AIDA Bureau van Dijk
AGE	Number of years since a firm's incorporation to 1996	23.217 (20.833)	19.889 (17.774)	Dun's and Bradstreet OneSource UK vol. 1 and 2 and AIDA Bureau van Dijk
OWNEMP	Total regional employment (NUTS 3 level) in a firm's own industry (1991)	4107.55 (6476.56)	2891.86 (6273.44)	Census of Employment (CSO for the UK and ISTAT for Italy)
OWNINN	Total regional employment of innovative firms in a firm's own industry (1991)	829.21 (2456.95)	728.45 (3172.25)	Census of Employment (CSO for the UK and ISTAT for Italy)
OWNNOINN	Total regional employment of non-innovative firms in a firm's own industry (1991)	3278.33 (5075.17)	2163.40 (4192.26)	Census of Employment (CSO for the UK and ISTAT for Italy)
OTHEMP	Total regional employment (NUTS 3 level) in all other industries (1991)	71233.81 (65254.47)	58276.24 (75569.95)	Census of Employment (CSO for the UK and ISTAT for Italy)
OTHINN	Total regional employment of innovative firms in all other industries (1991)	16803.91 (23058.67)	15047.61 (31228.93)	Census of Employment (CSO for the UK and ISTAT for Italy)
OTHNOINN	Total regional employment of non-innovative firms in all other industries (1991)	54429.89 (45936.43)	43228.63 (48287.59)	Census of Employment (CSO for the UK and ISTAT for Italy)
STOCKP	Total regional number of patents (NUTS 3 level) over the period 1978-87	323.01 (1116.19)	69.75 (229.63)	European Patent Office
PATPREV[d]	Dummy variable set to one if a firm had innovated previously in the period 1978-87	2.661	3.728	European Patent Office
PATHERF	Patent Herfindahl index measuring the diversity of each region (NUTS 3 level) for 30 technological classes in the period 1978-87[c]	0.148 (0.174)	0.260 (0.264)	European Patent Office
EMPHERF	Employment Herfindahl index measuring the diversity of each region (NUTS 3 level) at the 2-digit industry level (1991)	0.107 (0.035)	0.117 (0.037)	Census of Employment (CSO for the UK and ISTAT for Italy)
D_i	Industry fixed effects for each 2-digit UK SIC and NACE industrial sector	-	-	Census of Employment (CSO for the UK and ISTAT for Italy)

[a] The numbers reported refer to the mean and standard deviation (in brackets). The number of observations is, respectively, 23,872 for the UK, and 37,724 for Italy.
[b] European patent data have been processed by Cespri, Bocconi University, Milan.
[c] We adopted the aggregation of patents into 30 technological classes proposed by Fhg-ISI, Karlsruhe.
[d] Proportion of all firms in the sample for which the dummy variable is equal to one.

References

Acs, Z., Audretsch, D.B., 1990. Innovation and small firms. MIT Press, Cambridge MA.

Arthur, W.B., 1990. Silicon Valley locational clusters: do increasing returns imply monopoly?. Mathematical Social Sciences 19, 235-251.

Audretsch, D.B., 1998. Agglomeration and the location of innovative activity. Oxford Review of Economic Policy 14(2), 18-29.

Audretsch, D.B, Feldman, M.P., 1996. R&D spillovers and the geography of innovation and production. American Economic Review 86, 630-40.

Audretsch, D.B, Feldman, M.P., 1999. Innovation in cities: Science-based diversity, specialisation, and localised competition. European Economic Review 43, 409-29.

Baptista, R., Swann, G.M.P., 1998. Do firms in clusters innovate more?. Research Policy 27, 525-40.

Beaudry, C., 2000. Entry, growth and patenting in industrial clusters: A study of the aerospace industry in the UK, Working Paper, Manchester Business School.

Beaudry, C., Breschi, S., Swann, G.M.P., 2000. Clusters, Innovation and Growth: A Comparative Study of European Countries. In: Lipsey, R., Mucchielli, J.L. (Eds.) Proceedings of the 7th Sorbonne Conference on the Multinational Firm. Harwood Academic Publishers, page numbers not yet available.

Breschi, S., 1999. Spatial patterns of innovation. In: Gambardella, A., Malerba, F. (Eds.), The organisation of economic innovation in Europe. Cambridge University Press, Cambridge, pp. 71-102.

Breschi, S., 2000. The geography of innovations: a cross-sector analysis. Regional Studies 34(3), 213-229.

Breschi, S., Lissoni, F., 2000. Geographical boundaries of sectoral systems. WP Cespri, Università L. Bocconi, Milan.

Breschi, S., Malerba, F., Orsenigo, L., 2000. Technological regimes and Schumpeterian patterns of innovation. Economic Journal 110, 1-23.

Cäniels, M., 1999. Knowledge spillovers and economic growth: regional growth differentials across Europe. Edward Elgar, Cheltenham.

Cohen, W.M., 1995. Empirical studies of innovative activity and performance. In: Stoneman, P. (Ed.), Handbook of the Economics of Innovation and Technical Change. Blackwell, Oxford, pp. 182-264.

Crepon, B., Duguet, E., 1997. Estimating the innovation function from patent numbers: GMM on count panel data. Journal of Applied Econometrics 12, 243-263.

David, P.A., Rosenbloom, J.L., 1990. Marshallian factor market externalities and the dynamics of industrial localisation. Journal of Urban Economics 28, 349-370.

Dosi, G., 1988. The nature of the innovative process. In: Dosi, G., Freeman, C., Nelson, R., Silverberg, G., Soete, L. (Eds.), Technical change and economic theory. Pinter Publishers, London, pp. 221-38.

Echeverri-Carroll, E.L., Brennan, W., 1999. Are innovation networks bounded by proximity?. In: Fischer, M., Suarez-Villa, L., Steiner, M. (Eds.), Innovation, Networks and Localities. Springer Verlag, Berlin, pp. 28-49.

Feldman, M.P., 1994. The geography of innovation. Kluwer Academic Publishers, Dordrecht.

Freeman, C., 1991. Networks of innovators: a synthesis of research issues. Research Policy 20, 499-514.

Geroski, P., Van Reenen, J., Walters, C.F., 1997. How persistently do firms innovate?. Research Policy 26, 33-48.

Glaeser, E. L., Kallal, H.D., Scheinkman, J., Shleifer, A., 1992. Growth in cities. Journal of Political Economy 100, 1126-1152.

Greene, W.H., 1997. Econometric analysis 3rd edition, Prentice Hall, New York.

Griliches, Z., 1991. Patent statistics as economic indicators: a survey. Journal of Economic Literature 28, 1661-1707.

Hausman, J., Hall, B., Griliches, Z., 1984. Econometric models for count data with an application to the patents-R&D relationship. Econometrica 52, 909-938.

Hippel von, E., 1988. The sources of innovation. Oxford University Press, New York.

Howells, J., 1984. The location of research and development: some observations and evidence from Britain. Regional Studies 18, 13-29.

Howells, J., 1990. The location and organisation of research and development: new horizons. Research Policy 19, 133-46.

Howells, J., 1999. Regional systems of innovation?. In: Archibugi, D., Howells, J., Michie, J. (Eds.), Innovation policy in a global economy. Cambridge University Press, Cambridge, pp. 67-93.

Jacobs, J., 1969. The economy of cities. Penguin, London.

Jaffe, A.B., 1989. Real effects of academic research. American Economic Review 79, 957-970.

Jaffe, A.B., Trajtenberg, M., Henderson, R., 1993. Geographic localization of knowledge spillovers as evidenced by patent citations. Quarterly Journal of Economics 108, 577-598.

Kelly, M., Hageman, A., 1999. Marshallian externalities in innovation. Journal of Economic Growth 4, 39-54.

Klepper, S., 1996. Entry, exit, growth and innovation over the product life-cycle. American Economic Review 86(3), 562-83.

Kline, S.J., Rosenberg, N., 1986. An overview of innovation. In: Landau, R., Rosenberg, N. (Eds.), The positive sum strategy. National Academy Press, New York, pp. 275-305.

Krugman, P., 1998. What's new about the new economic geography?. Oxford Review of Economic Policy 14(2), 7-17.

Lamoreaux, N.R., Sokoloff, K.L., 1997. Location and technological change in the American glass industry during the late nineteenth and early twentieth centuries. NBER Working Paper No. 5938.

Lucas, R., 1988. On the mechanics of economic development. Journal of Monetary Economics 22, 3-42.

Lundvall, B.A., 1993. National systems of innovation. Pinter Publishers, London.

Lyons, D., 1995. Agglomeration economies among high technology firms in advanced production areas: the case of Denver/Boulder. Regional Studies 29(3), 265-78.

Malerba, F., 1992. Learning by firms and incremental technical change. Economic Journal 102, 845-59.

Malerba, F., Orsenigo, L., Peretto, P., 1997. Persistence of innovative activities, sectoral patterns of innovation and international technological specialisation. International Journal of Industrial Organisation 15(6), 801-26.

Mansfield, E., 1995. Academic research underlying industrial innovations: sources, characteristics, and financing. Review of Economics and Statistics February, 55-65.

Marshall, A., 1920. Principles of economics. Macmillan, London.

Nelson, R.R., 1993. National systems of innovation: a comparative study. Oxford University Press, New York.

OECD, 1994. The measurement of scientific and technological activities. Using patent data as science and technology indicators. Patent Manual 1994, Paris. (http://www.oecd.org//dsti/sti/stat-ana/prod/e_94-114.pdf)

Porter, M.E., 1998. Clusters and the new economics of competition. Harvard Business Review November/December, 77-90.

Pred, A., 1966. The spatial dynamics of US urban-industrial growth, 1800-1914. Interpretative and theoretical essays. MIT Press, Cambridge MA.

Pred, A., 1971. Urban growth and the circulation of information: the United States system of cities, 1790-1840. Harvard University Press, Cambridge MA.

Romer, P., 1986. Increasing returns and long run growth. Journal of Political Economy 94, 1002-1037.

Storper, M., 1992. The limits to globalisation: technology districts and international trade. Economic Geography 68(1), 60-93.

Suarez-Villa, L., Walrod, W., 1997. Operational strategy, R&D and intra-metropolitan clustering in a polycentric structure: the advanced electronics industries of the Los Angeles basin. Urban Studies 34(9), 1343-80.

Swann, G. M. P., 1998. Towards a model of clustering in high technology industries. In: Swann, G.M.P., Prevezer, M., Stout, D. (Eds.), The dynamics of industrial clusters: international comparisons in computing and biotechnology. Oxford University Press, Oxford, pp. 52-76.

Swann, G.M.P., Prevezer, M., 1996. A comparison of the dynamics of industrial clustering in computing and biotechnology. Research Policy 25, 1139-1157.

Swann, G.M.P., Prevezer, M., Stout, D., 1998. The dynamics of industrial clusters: international comparisons in computing and biotechnology. Oxford University Press, Oxford.

Thompson, W.R., 1962. Locational differences in inventive effort and their determinants. In: Nelson, R.R. (Ed.), The rate and direction of inventive activity. Princeton University Press, Princeton, pp. 253-72.

Winter, S.G., 1987. Knowledge as strategic asset. In: Teece, D.J. (Ed.), The competitive challenge: strategies for industrial innovation and renewal. Ballinger, Cambridge MA, pp. 159-84.

World Intellectual Property Organization, 1994. *International Patent Classification, 6th edition (1994)*, volume 9 Guide, survey of classes and summary of main groups, Geneva.

Table 1 – Number of patents by firm, United Kingdom and Italy

United Kingdom			Italy		
Patents per firm	*No. of firms*	*Proportion (%)*	*Patents per firm*	*No. of firms*	*Proportion (%)*
0	22781	95.430	0	35108	93.065
1	556	2.329	1	1524	4.040
2	199	0.834	2	407	1.079
3	87	0.364	3	193	0.512
4	52	0.218	4	107	0.284
5	46	0.193	5	70	0.186
6	24	0.101	6	53	0.140
7	21	0.088	7	47	0.125
8	8	0.034	8	22	0.058
9	8	0.034	9	19	0.050
10-19	40	0.168	10-19	87	0.231
20-29	14	0.059	20-29	31	0.082
30-39	9	0.038	30-39	16	0.042
40-49	3	0.013	40-49	8	0.021
50-99	14	0.059	50-99	18	0.048
100-499	7	0.029	100-499	13	0.034
500-1035	3	0.013	500-967	1	0.003

Table 2 – Impact of clustering on firms' innovative activities

(negative binomial regression, pooled analysis)

	United Kingdom		*Italy*	
Parameter	[1]	[2]	[1]	[2]
INTERCEPT	-6.2698 ***	-2.4889 ***	-7.8497 ***	-4.6231 ***
ln(CIEEMP)	0.7965 ***	0.7596 ***	0.8724 ***	0.8631 ***
ln(OWNEMP)	0.1175 ***	0.1453 ***	0.0081	-0.0687 ***
ln(OTHEMP)	-0.0965 ***	-0.6197 ***	0.1545 ***	-0.1821 ***
PATPREV	3.0928 ***	3.0867 ***	1.8894 ***	1.8751 ***
PATHERF	-3.6298 ***	0.6664	-0.5158 ***	-0.0242
EMPHERF	0.7001	-0.0919	2.5190 ***	1.8194 ***
ln(STOCKP)		0.3482 ***		0.2776 ***
DTextile	-1.0345 ***	-0.9772 ***	-1.3219 ***	-1.2172 ***
DTimber	-1.1129 ***	-1.2459 ***	-0.4184 **	-0.4339 **
DPaper	-1.2751 ***	-1.4895 ***	-0.1805	-0.2044
DPrint			-1.2724 ***	-1.2846 ***
DChemical	1.0555 ***	0.8378 ***	0.8749 ***	0.9183 ***
DPlastic	-0.9300 ***	-1.1312 ***	0.4032 ***	0.4260 ***
DNon-metallic	-1.4356 ***	-1.4729 ***	-0.7804 ***	-0.7072 ***
DMetalManuf	-0.9313 ***	-1.0343 ***	-1.1290 ***	-1.1143 ***
DMetalGoods	-0.6154 ***	-0.7406 ***	0.1458 *	0.2336 ***
DMechanical	-0.6341 ***	-0.7918 ***	1.0430 ***	1.1030 ***
DOffice	-0.9639 ***	-0.9983 ***	0.4782 ***	0.3943 ***
DTeleco			0.6706 ***	0.6863 ***
DElectrical	-0.1591	-0.4451 ***	1.3257 ***	1.3535 ***
DInstruments	0.7811 ***	0.5875 ***	0.8006 ***	0.7675 ***
DVehicles	0.4809 ***	0.3989 ***	0.0152	0.0428
DOtherTransports	-0.1099	-0.3351 **	-0.7414 ***	-0.7288 ***
α	0.000170	0.000158	0.000553	0.000488
Observations used	23872	23872	37724	37724
Log-Likelihood	-11585	-11315	-20992	-20861
Deviance/DF	0.9717	0.9491	1.1136	1.1067
Pearson χ^2/DF	5.2202	4.9250	3.2769	3.2185

Notes: *, ** and *** indicate, respectively, significance at the 10%, 5% and 1% levels.
For the UK, dummy variables for company status and the type of accounts it files are also included, but not shown here.

Table 3 – Clustering of innovative companies and firms' innovative activities, United Kingdom and Italy (negative binomial regression, pooled analysis)

	United Kingdom				*Italy*			
Parameter	[1]	[2]	[3]	[4]	[1]	[2]	[3]	[4]
INTERCEPT	-2.4466 ***	-4.1734 ***	-0.4416	-2.1505 ***	-8.0076 ***	-7.5943 ***	-6.1955 ***	-6.0465 ***
ln(CIEEMP)	0.9743 ***	0.7954 ***	0.9499 ***	0.7550 ***	1.0882 ***	0.8383 ***	1.0805 ***	0.8313 ***
ln(OWNINN)	0.3756 ***	0.1262 ***	0.3593 ***	0.1238 ***	0.2886 ***	0.1966 ***	0.2569 ***	0.1694 ***
ln(OWNNOINN)	-0.2719 ***	0.0225 *	-0.2582 ***	0.0043	-0.3092 ***	-0.1430 ***	-0.3435 ***	-0.1729 ***
ln(OTHINN)	0.1713 ***	0.2288 ***	-0.0499 **	-0.0573 **	-0.0350 ***	-0.0149	-0.1362 ***	-0.1084 ***
ln(OTHNOINN)	-0.5020 ***	-0.4396 ***	-0.5972 ***	-0.5453 ***	0.2318 ***	0.1552 ***	0.1199 ***	0.0554 **
PATPREV		2.9858 ***		2.9766 ***		1.8038 ***		1.8009 ***
PATHERF		-1.2358 **		1.0636 **		-0.5220 ***		-0.1756
EMPHERF		-5.0145 ***		-1.7803		3.3146 ***		3.0650 ***
ln(STOCKP)			0.2496 ***	0.3338 ***			0.2070 ***	0.1949 ***
DTextile	-1.3810 ***	-0.9849 ***	-1.3686 ***	-0.9811 ***	-1.2981 ***	-1.2218 ***	-1.2698 ***	-1.1796 ***
DTimber	-1.3047 ***	-1.1385 ***	-1.3647 ***	-1.1767 ***	-0.2051	-0.1599	-0.2595	-0.2025
DPaper	-1.5499 ***	-1.3676 ***	-1.6194 ***	-1.4544 ***	-0.4235 ***	-0.0827	-0.4368 ***	-0.1197
DPrint					-1.3021 ***	-1.1094 ***	-1.3528 ***	-1.1469 ***
DChemical	0.7564 ***	0.8865 ***	0.6317 ***	0.7317 ***	0.6084 ***	0.5914 ***	0.6191 ***	0.6085 ***
DPlastic	-0.8324 ***	-1.0570 ***	-0.9563 ***	-1.2648 ***	0.2457 ***	0.2968 ***	0.2371 ***	0.3066 ***
DNon-metallic	-1.4667 ***	-1.5165 ***	-1.5099 ***	-1.5964 ***	-1.0218 ***	-0.7182 ***	-0.9964 ***	-0.6840 ***
DMetalManuf	-0.9335 ***	-1.1147 ***	-0.9762 ***	-1.1753 ***	-1.5173 ***	-1.2270 ***	-1.4983 ***	-1.2043 ***
DMetalGoods	-0.8991 ***	-0.7836 ***	-0.9336 ***	-0.8009 ***	0.0959	0.0729	0.1497 *	0.1313 *
DMechanical	-0.7486 ***	-0.8324 ***	-0.8598 ***	-0.9080 ***	0.7595 ***	0.7416 ***	0.8006 ***	0.7834 ***
DOffice	-1.0804 ***	-1.0636 ***	-1.1849 ***	-1.0754 ***	-0.3165 ***	0.1022	-0.4553 ***	-0.0041
DTeleco					0.3090 ***	0.5205 ***	0.3075 ***	0.5261 ***
DElectrical	-0.3700 **	-0.3525 **	-0.5297 ***	-0.5367 ***	0.8709 ***	1.0226 ***	0.8676 ***	1.0323 ***
DInstruments	0.6898 ***	0.6003 ***	0.5825 ***	0.4290 ***	0.4840 ***	0.6160 ***	0.4510 ***	0.5950 ***
DVehicles	-0.1079	0.3632 **	-0.1198	0.2853 *	-0.6072 ***	-0.4554 ***	-0.6397 ***	-0.4569 ***
DOtherTransport	-0.3950 **	-0.2458	-0.4995 ***	-0.4634 ***	-1.5640 ***	-1.0740 ***	-1.5969 ***	-1.0794 ***
α	0.024537	0.000162	0.023020	0.000142	0.006146	0.000429	0.005832	0.000425
Observations used	23872	23872	23872	23872	37724	37724	37724	37724
Log-Likelihood	-13367	-11479	-13356	-11278	-21986	-20775	-21967	-20697
Deviance/DF	1.1210	0.9627	1.1201	0.9460	1.1663	1.1021	1.1653	1.0980
Pearson χ^2/DF	8.9108	5.1883	8.2345	4.8337	4.6404	3.1643	4.4327	3.1336

Notes: *, ** and *** indicate, respectively, significance at the 10%, 5% and 1% levels.
For the UK, dummy variables for company status and the type of accounts it files are also included, but not shown here.

Table 4 – Impact of cluster variables on firms' innovative performance (negative binomial regression, by industry analysis)

	United Kingdom																
	Textile	Timber	Paper	Printing	Chemical	Plastics	Non-met	Met man	Met good	Mechanical	Electrical	Office	Telecom	Instr.	Vehicles	Transport	Other
CIEEMP	+		+		+	+	+	+	+	+	+	+		+	+	+	+
ln(OWNEMP)			-				+			-	-	-		+			
ln(OTHEMP)		-			-	-	-					-		-	+	+	+
PATPREV	+	+	+		+	+	+	+	+	+	+	+		+	+	+	+
ln(STOCKP)					+	+	+	+			+	+			-		
ln(CIEEMP)	+		+		+	+	+	+	+	+	+			+	+	+	+
ln(OWNINN)	+	+	+		+	+	+		+					+	+	+	+
ln(OWNNOINN)		+	-		+	+	+		-		-				-		-
ln(OTHINN)	+				+	-	-		-	-						-	
ln(OTHNOINN)	-	-			-	-	-		+					-	+	+	
PATPREV	+	+	+		+	+	+	+	+	+	+			+	+	+	+
ln(STOCKP)			+		+	+	+	+			+			-	-		
Observations	823	1321	3711	-	1426	1563	734	694	1939	5108	2705	495	-	899	637	569	1248
% of zeros	96.23	98.64	98.98	-	92.71	94.50	94.69	95.82	97.01	94.44	93.16	96.77	-	89.43	92.15	94.73	97.76

	Italy																
	Textile	Timber	Paper	Printing	Chemical	Plastics	Non-met	Met man	Met good	Mechanical	Electrical	Office	Telecom	Instr.	Vehicles	Transport	Other
ln(CIEEMP)	+	+	+		+	+	+	+	+	+	+	+	+	+	+	+	+
ln(OWNEMP)			-			-	-		+			+	-	-			-
ln(OTHEMP)	-		+	-		-		+	-	-	-	-	-		-		
PATPREV	+	+	+	+	+	+	+	+	+	+		+	+	+	+	+	+
ln(STOCKP)	+		-		+	+	+	-	+	+	+	+	+	+	+	+	+
ln(CIEEMP)	+	+	+		+	+	+	+	+	+	+	+	+	+	+	+	+
ln(OWNINN)	+	+	+	+	+	+	+	+	+	+		+	+	+	+	+	+
ln(OWNNOINN)	-	-	-	-	-	-	-	-		-	-	+	-	-	-	-	-
ln(OTHINN)	+	-			-		+		-			-	-		+		+
ln(OTHNOINN)	-		+	+	+			+	-		-	-				+	-

PATPREV	+		+	+	+	+	+	+	+	+		+	+	+	+	+	
ln(STOCKP)			-		+	+		-				+	+		-		
Observations	3916	990	1059	1881	2495	2464	2625	1419	5837	6121	2431	240	934	1093	813	525	2881
% of zeros	98.42	97.37	96.03	98.88	88.34	90.91	97.45	95.70	94.77	86.23	93.01	93.33	88.87	87.56	88.92	93.33	95.73

Note: +: positive sign and significant at the 10% level; -: negative sign and significant at the 10% level. PATHERF and EMPHERF are also included in the model.