CESIS
Electronic Working Paper Series

Paper No. 16

Technological Diversity and Jacob's Externality Hypothesis Revised[1]

OLOF EJERMO

(JIBS)

September 2004

The Royal Institute of technology
Centre of Excellence for studies in Science and Innovation

[1] Status of the paper: Accepted for publication in Growth and Change

www.infra.kth.se/cesis/research/workpap.htm
Corresponding author: olof.ejermo@ihh.hj.se

Technological Diversity and Jacobs' Externality Hypothesis Revisited*

Olof Ejermo†
olof.ejermo@jibs.hj.se

10th September 2004

Abstract

Recent empirical evidence strongly supports Jacobs' (1969) externality hypothesis, that urban diversity provides a more favourable environment for economic development. In order to correctly gauge Jacob's hypothesis, economic development should be understood as a result of innovations. Furthermore, it is argued that a relevant diversity-measure should take into account the degree of diversity between the inherent classes (e.g. pharmaceuticals are closer to chemicals than to forestry). These ideas are tested using regionally classified Swedish patent application data as a measure of innovativeness. Patent data are also used to reflect technological diversity. The results show that the number of patent applications in Swedish regions, are highly and positively dependent on regional technological specialization, quite the opposite to Jacobs' prediction. The paper raises general questions about earlier empirical results. It is concluded that the size of regions is important is an important factor to consider, since this in itself may affect patenting intensity and technological diversity.

JEL Classification: O31, R12, H41, O40

Keywords: Specialization, diversity, patenting, Sweden, regions

*The first draft of this paper was presented at my visit at the Centre for Research on Innovation and Internationalization (CESPRI) at Bocconi University in Milan in 2002. I thank Stefano Breschi, Francesco Lissoni and Fabio Montobbio for useful discussions and suggestions. Financial support from the Swedish Foundation for International Cooperation Research and Higher Education (STINT) that made this visit possible is also gratefully acknowledged.

†Jönköping International Business School, P.O. Box 1026, SE-551 11 Jönköping, Sweden

1 Introduction

Specialization and diversity of regional economic environments have been seen as partially opposing forces in agglomeration economies (Marshall, 1920; Jacobs, 1969). Recent empirical evidence strongly supports Jacobs' (1969) externality hypothesis, that urban diversity provides a more favourable environment for economic development. This paper claims that such an interpretation may be misleading. It is argued that in an empirical application of the concept, diversity should consider two dimensions. First, there is a "classification diversity" that arises because activities are grouped under different headings. That is, an environment is diverse if two activities are in different classes. Most contributions in the literature are content with a numerical measurement taking this into account. However, the classes of which any measure necessarily has to be composed are in themselves in varying degreees interrelated. For instance, a region producing chemicals and pharmaceuticals has clearly much more (technologically) related activities than a region producing paper and dairy products, a fact which is not captured by previous measures. This leads us to the second type of diversity which is labelled "technological diversity" in the paper. In response to these ideas, a measure of technological coherence earlier is adapted for regional use. This measure has previously been used used on the firm level (Teece et al., 1994 and Breschi, et al. 2003).

The measure is based on the following principles. All patents filed at the European Patent Office (EPO) are assigned a main and a supplementary class. Hinze et al. (1997) used this information to construct a table of closeness among patent technologies based on a divison on 30 patent classes. This table is adopted as a means to obtain numerical values of diversity in Swedish regions, after removing classification diversity, i.e. the diversity that accrues due to table of technological closeness. The remaining part is labelled technological diversity since it is mainly related to the technological coherence of the patents in a region. The resulting measure is used to re-evaluate the dependence of regional economic development, interpreted as development of new patents, on regional diversity.

The paper is organized as follows. Section 2 reviews theories on agglomeration economies relevant for the present contribution. Section 3 makes a critical examination of empirical tests of the theories. Section 4 argues in favour of our measure of diversity based on patents and go on to describe its construction. Section 5 tests how patent diversity affects development of new patents for Swedish regions in the 1990s period. Section 6 concludes.

2 Theories about agglomeration

Historically, agglomeration economies have played an integral part in urban and regional economics.[1] Identification of effects follows a long tradition. Marshall (1920, pp. 271) stressed that there were three main reasons why industries should agglomerate into industrial districts. First, subsidiary trades (i.e. specialised intermediaries) grow up in the neighbourhood of an industry, is facilitated by cost-efficiently and extensively used specialised machinery. Second, labour becomes specialized and moves between nearby-firms, thus reducing hiring costs and uncertainty, to the benefit of both labourers and employers. Third, ideas about inventions and improvements in machinery, in processes and the general organization, spread swiftly through the local district. According to Henderson et al. (1995, p. 1068), the effects described by Marshall may best be described as static localization externalities. One can visualize a local economy in equilibrium, with no dynamic development of products and where input-output relationships between industries are stable. Glaeser et al. (1992, p. 1127) form the MAR acronym by adding the contributions by Arrow (1962) and Romer (1986) to the discussion. These additions are motivated from a dynamic-localization externalities perspective. Arrow (1962) formalized the dynamic influence that learning has on unit-costs for the competitiveness of firms. Romer's (1986) paper, brought increasing returns to scale on the firm level into the picture by formalizing the impact of dynamic knowledge accumulation. All MAR-effects give support to the idea that specialization is what matters for the firm. Thus, these *localization economies* lead to lower average costs of an industry as more firms in an industry co-locate. As stressed by Hall (1959, quoted in Quigley, 1998), these economies arise because of fractional needs. Since there are many (potential) buyers, also fractional needs may be catered by the individual firm. *Urbanization economies* on the other hand (cf. Ohlin, 1933; Hoover, 1937), refer to advantages of the size and density of the local economy. This includes variety of specialized services, which having different industries located near each other entail. Consumers enjoy similar advantages.

Building on these works is the contribution of Storper (1995) with his recognition of the concept untraded dependencies. Untraded dependencies arise, not

[1] This discussion draws in part from Pettersson (2001). It is stressed that only agglomeration theories considered relevant for the present discussion are covered.

only through input-output linkages, but also due to conventions, rules, practices and institutions, which combine to produce 'worlds of production' (Storper and Salais, 1997). The empirical studies of Pinch and Henry (1999a; 1999b among other contributions) of the British 'Motor Sport Valley' act as illustrations of the Storper type of interdependencies. In this area of South-East England, there is fierce competition between suppliers to car-producers, and between car-producers themselves, in the racing-car industry. Yet, suppliers may occassionally provide subtle information about more promising directions of development for the different manufacturers based on their contacts with competitors. Also, new developments can seldom be kept hidden for long, since they are often partially visible on the racing circuit and hence noticable through learning-by-observing processes. Labour is relatively mobile between firms, and key engineers are paid large incomes to remain within the firm. Start-ups and spinoffs (as well as deaths) of firms are frequent, and seem to be "accepted" as a manifestation of an entreprenurship culture. There is a competitive environment, yet most actors seem aware about the benefits for the sport as a whole that competitors do not lag too far behind (Pinch and Henry, 1999a). Thus, there is also informal collaboration going on.

A different idea is attributed to Jacobs (1969). She argues that more diverse environments, such as those existing in cities, provide a better breeding-ground for *new ideas*, because in this way cross-fertilization of ideas from different areas are facilitated. The diversity idea is used differently in new spatial agglomeration theory (cf. Fujita et al., 1999 and Fujita and Thisse, 2001). The fact that diversity should enter the utility function of the consumer was modelled in the contributions by Spence (1976) and Dixit and Stiglitz (1977). As such, the very fact that consumers enjoy higher utility from a diverse set of consumer goods was explicitly modelled in a trade-off between diversity and output quantity by the later contribution. This result has been translated into spatial regional economics in two versions. In the first, the core-periphery model (Krugman, 1991, repeated in Fujita and Thisse, 2001), consumers' desire for variety is applied to a two-region case. concentration of people to one region, with ensuing production concentration implies a circular causation process. In the second version, higher specialization of individual intermediate producers entail higher production quantities for the final producer.[2] This specialization of the individual producer entails a higher diversity in the aggregate, which renders the

[2] See Fujita and Thisse, 2001, ch. 4 for a formalization.

end-producer more productive.

Given the advantages of residing in cities, suggested above, why does not all economic activity concentrate in large cities? The countervailing force comes from congestion, which impairs a cost penalty for the producer locating in urban areas (cf. Mills, 1967; Henderson, 1974 for a formalization). Price of land plays the role of an allocation mechanism, implying that only activities in sufficient demand of urban space will locate where the density of economic activities is high. If innovative activities can meet these conditions, an immediate link to spatial product life cycle theory is evident. This theory is exemplified by the formalized model of Duranton and Puga (2001), where diversified cities act as nurseries for new products. In the work by Andersson and Johansson (19xx) and Johansson and Karlsson (19xx), this nursing arises for several reasons. Urban settlement have a higher concentration of "alert customers", people with preference for trying out new goods.[3] This provides an impetus on the consumption side. On the supply side, production of these new goods require more educated workers, which are also concentrated in cities. It is also argued that knowledge-intensive firms are more concentrated to these regions, which give a comparative advantage for the innovative firm. Once firms have found a suitable production system for the new good, they decentralize production to avoid congestion costs, and engage in mass production. The comparative advantage of cities therefore reside in their abilitiies to develop and test new products, whereas the comparative advantage of more peripheral areas are in routinized production. In this way, the established fact that specialized and diverse cities co-exist can be explained, but due to the presence of congestion, we would expect more innovations to take place in the larger cities. Furthermore, while innovations are developed, static localization and urbanization economies are interwoven in a complicated dynamic game within urban settlements, so that cities can seldom be identified as completely diverse or specialized.

The above discussion suggests that the stage of the product life cycle and the size of the urban settlement together play a crucial role for the location of production, and as suggested by Jacobs, the development of ideas. We can conclude from this theoretical discussion that while diversity/specialization arguments emphasize agglomeration in general, Jacobs' discussion always emphasizes the development of ideas, resulting from diversity. The Marshallian externality literature is more complex, but "production" is generally considered to be the

[3] It can quite simply be argued that cities with a larger mass of people will contain more people with a preference for trying out new goods.

most affected factor. On the one hand, industries concentrate into specialized districts, so that a good indicator of the merits of specialization seems to be measure 'industrial development'. On the other hand, ideas are according to Marshall (1920), "in the air" in those districts, suggesting that specialization is useful even in this case. Therefore, it can be concluded that a test of Jacobs' externalities should focus on an empirical application of the concept "ideas", whereas Marshallian externalities should more likely be associated with production, with a weaker test based on "ideas". In addition, spatial product life cycle theory suggests that the extent to which production is sensitive to specialization is related to the stage of the product's life cycle.

3 Empirical Evidence

We now critically examine how the empirical literature has dealt with Marshallian and Jacobs' externalities in the light of the previous discussion. To start out, it is by now an established fact that innovations tend to cluster in space. Moreover, this concentration seems to be much stronger when industries are 'high-tech'. For instance, Audretsch and Feldman (1996) studied geographic concentration of innovative activity, measured by innovation counts from the Small Business Administration database, to compare this with the concentration of production. The starting point of their study was that innovations should, other things equal, follow the general structure of production activity. If new economic knowledge is a more important externality than other agglomeration externalities, the implication would be that these industries would concentrate more. Audretsch and Feldman (1996) are able to confirm this hypothesis, and also find that concentration of innovation is higher in R&D-intensive industries and in industries for which university R&D and skilled labour is important. A similar conclusion was reached by Kelly and Hageman (1999), who use a dartboard approach (cf. Ellison and Glaeser, 1997). That is, the distribution of patenting is studied with regard to how it differs from the distribution implied by the geographical structure of production. They find that the concentration of innovative activities in space is often quite different from the one of production. An interesting conclusion of their paper is that "tests of Marshallian externalities that focus on the growth of output (investigating whether areas with high wages or high concentrations of an industry experience the highest growth rates) may miss the most important aspect of these externalities, which is their effect on the innovation process." (Kelly and Hageman 1999, p. 50).

The empirical problem of evaluing Marshallian and/or Jacobs' externalities can from the perspective of this paper mainly be evaluated by means of two parameters. First, what is a relevant measure of specialization and diversity respectively, and secondly what is the outcome, i.e. the dependent variable to be tested. The empirical problem of evaluing Marshallian and/or Jacobs' externalities can from the perspective of this paper be approached in the following way. First, a relevant measure of specialization has to be selected. In what follows, it is assumed that diversity increases as this measure decreases. Second, an outcome or performance variable has to be selected.

Most researchers recognise the importance of controlling for 'normal' agglomeration effects, that is, other things equal, higher population density should increase economic concentration. Glaeser et al. (1992) is the starting point for the recent empirical evidence. This paper measures Marshallian, localization externalities in industries by using location quotients, the share of local industry production in a specific industry in relation to the average national share. Diversity is simultaneously measured by taking the share of local employment of the top five industries (other than the one in question) of total employment. These measures are, after controlling for population, gauged with regard to their effects on employment development over the period 1956-1987 in US cities. They find that more diverse industries have higher growth in employment. The fact that the investigated industries may be considered mature, should however have an effect on the outcome. Mature industries may have a slower growth of employment and distort true results. These observations led Henderson et al. (1995) to refine the setup by considering a later and shorter time-period, 1970-1987 thereby including all cities.[4] In addition, initial and regional conditions are included as control variables. Still, their focus is on industry growth, measured by employment. Specialization is measured by location quotients, and diversity by the Hirschmann-Herfindahl index. This measure[5] has the advantage that all industries are taken into account. On the other hand, its drawback is that diversity is measured symmetrically. That is, diversity comes from the fact that employment falls into different industry classes, not from how different these industries are. Henderson et al. (1995) divide between perceived

[4] In fact, the cities are American Metropolitan Statistical Areas (MSAs).

[5] In Henderson et al. (1995) the Hirschman-Herfindahl index for city i for industry k is written $HHI_{ik} = \sum_{j \notin k} s_{ij}^2$, where s_{ij} is the share of employment in city i in industry j.

mature industries (machinery, electrical machinery, primary metals, transportation and instruments) and new (computers, electronic components and medical instruments) and conclude that Marshall-Arrow-Romer effects are important for traditional industries, Jacobs externalities seem important for attracting new, but MAR-externalities are important for retaining them.

Apart from the continued use of employment growth, these papers use location quotients to account for the importance of specialization. It should be well-known, however, that this measure is very sensitive to the size of the region.[6] Nonetheless, it carries on to perhaps the most well-known piece in this literature, Feldman and Audretsch (1999). They count US innovations, as found in about 100 technology, engineering and trade journals for 1982 (the Small Business Administration data base) for industries, grouped together into six groups based on their scientific similarity. This similarity is taken from a survey of managers (cf. Levin et al. 1987) where company leaders rank the importance of different academic disciplines. Concentration of the scientific base of the industry is also measured by the location quotient. Even though Feldman and Audretsch in this way partially avoid problems of using the industry classification, the latter is still in the background when combining the six groups. A higher location quotient is interpreted to validate localization economies; lower to validate urbanization. Feldman and Audretsch (1999) find that diversity promotes the number of innovations; concentration acts in the opposite direction. Nearby concentration of industries from the same scientific base also increases the amount of innovations. Thus, Feldman and Audretsch (1999) check whether the number of innovations and not employment grows. A problem is that their results cannot be replicated easily, because data on innovation counts exist only for one year and it is unknown whether their results are specific to the US. While their approach avoids many pitfalls, as mentioned, they rely on the existing industry classification, at least indirectly.

The setup of Paci and Usai (1999) is quite similar to that of Feldman and Audretsch (1999), but Italian local labour markets is used as the regional unit of analysis. Location quotients are used to check for the presence of Marshallian externalities, together with an inverse Gini coefficient to estimate the importance of diversity. They use the same scientific group division as Feldman

[6] The location quotient of industry i in region r is written $LQ_{ir} = (E_{ir}/E_r)/(E_i/E)$, where it measures the share of local employment in industry i in the region r in relation to that of the whole of the nation. Clearly, the quotient rises if relatively more people works in the regional industry, but also if the region becomes smaller (E_r goes down) as would happen if a factory is located to a small region.

and Audretsch (1999). Moreover, they include a high-tech dummy to specify whether certain sectors are affected differently by localization or urbanization economies (as in Henderson et al., 1995). Also, spatial econometric techniques are used to take into account the possibility of cross-bordering effects. The results seem, in contrast to Feldman and Audretsch (1999) to show concentration to be more important for innovation development in Italy than in the US. The authors attribute this result to the structure of Italy's economy, which is largely composed of small- and medium-sized firms in traditional industries. This is further corroborated by the fact that the high-tech dummy for diversity is significant, meaning that high-tech sectors are more dependent on diversity, consistent with earlier results. Another interesting finding is that the effects do not transcend regional borders. Thus, local labour systems seem indeed to be largely self-contained units in this case.

van Oort (2002) examines the importance of diversity and specialization in a spatial econometrics framework for innovation intensity in the Netherlands. Labour costs of R&D across Dutch municipalities is used to proxy innovation activity. Concentration is measured by the location quotient of employment by a particular industry in a municipality, lack of diversity is measured by the Gini-coefficient of the distribution of employment by sector in a municipality. Also, competition is measured as establishments per worker in a municipality. The author finds it relevant to divide the country into national zoning areas and core-periphery regions. An interesting empirical fact reside in the fact that innovation intensity, R&D labour costs per employee, tend to be lower in population dense, i.e. core locations, in the Netherlands.[7] The econometric results show that innovation tends to cluster in urban areas, interpreted as needing ensurance of proximity to reservoires of high-skilled workers. Including spatially dependent lag variables does not change the results, which is interpreted as a limited reach of R&D spillovers and that innovation intensity patterns are caused by 'hot-spots' of large, dominants firms causing a spatial concentration of innovation intenstiy. Similar to earlier results, reported on above, innovation intensity seems to be fostered by diversity and an absence of concentration. Also, lack of competition seems to be innovation-inducing.

Acs, Fitzroy and Smith (2002) estimate panel equations for 5 high technology clusters in 36 MSAs for the four years 1988-1991. Employment was estimated

[7] The core, the Randstad region contains the major cities Amsterdam, Rotterdam, The Hague and Utrecht. It contains the provinces North Holland, South Holland, Utrecht and Flevoland.

to be dependent on real wages, total amount of innovations in the region and R&D expenditures from industry and university, respectively, in the specific high technology sector. Spillovers between industries and from one university sector to a specific industry cluster were measured by simply adding-up research expenditures of all other sectors, except for the sector studied. Thus, there is no attempt to control for technological differences and/or the technological proximity between high-tech sectors. While the authors control for possible selection-bias and endogeneity of employment to wages, it is not surprising they don't find evidence of spillovers between sectors. Rather they find a strong effect of research expenditures on employment, thus claimed to support the MAR-hypothesis.

Harrison et al. (1996) and Kelley and Helper (1999) study microdata to examine whether adoption of new production processes, by individual enterprises in machine-making industries, are promoted by specialized or diverse employment in the region. Diversity was found to promote adoption.

The only paper to investigate effects of something that comes close to technological diversity of regions, inferred by the patent classification, seems to be Autant-Bernard (2001). She investigates the production of patents in French administrative regions (departements) as a function of research and human capital, while also considering the impact of these variables from neighbouring regions. The technological profile index, inspired by the one used by Jaffe (1986)[8], takes into account how much is patented in different groups, but not the technological relatedness of different groups. The results point to a positive effect from technological proximity. That is, when the region and/or its neighbours technological specialization increase, patenting is positively affected. It is hypothesized that technological and geographical proximity seem to be determined simultaneously. The measure of technological proximity, however markedly differs from the one used in this paper. As will be seen, the measure used here in addition considers the relatedness between classes.

To sum up the empirical evidence, we note that as a general rule industry codes are used to classify firms. This industry classification is used to construct a diversity index (usually the Hirschmann-Herfindahl). We also note that specialization is measured by localization quotients despite its small-region problems. Diversity measures do not capture technological diversity as outlined in the introduction. Some papers use employment growth to test for Jacobs'

[8] See equation (1) below.

externalities, although it seems more apt to test localization economies.

4 The Use of Patents for Measuring Diversity

This section describes and explains how a measure of regional diversity based on patent data can be constructed. Given that *innovations* seem to be of special importance for clustering of activities, and therefore agglomerations, a measure trying to evaluate the effects of specialization or diversity should avoid the industry classification. This was already noted by Jacobs (1969, p. 61):

> "These are useful categories for some types of economic analysis, but insofar as they are relevant at all to understanding how old work leads to new, they interfere with our understanding."

She concludes:

> "The point is that when new work is added to older work, the addition often cuts ruthlessly across categories of work, no matter how one may analyze the categories" (ibid, p. 62)

Desrochers (2001, p. 375) adds that:

> "virtually all functional processes and material continually traverse "industrial branches" and that firms producing widely different outputs often use related production technologies. Another problem with industrial classification data is that they hide the multi-product nature of virtually all firms of any significance."

Patents on the other hand have their own classification codes. Within their own limits, databases provide full coverage of the units of observations. That is, patent data provide whole populations of data, and not just samples. Desrochers (2001) argues that the patent classification system would provide a better indicator of diversity, quoting Griliches (1990, p. 1666), because it "is based primarily on technological and functional principles and is rarely related to economists' notions of products or well-defined industries". Furthermore, more information can be conveyed from inventive actors, rather than just their industry code. We get more information on the technological profiles of regions from studying patents.

How are patents classified? Patents use the International Patent Classification (IPC) system which is "...based on its need to ease the search for prior art." (Griliches, 1990, p. 1666). Thus, patents are classified mainly due to ease of finding them for the patent examiner and others wishing to obtain knowledge about them. We should however note the disadvantages of patents as an observable. Some of the deficiencies include: 1) not all innovations are patentable; 2) not all patentable innovations are patented; 3) there are strong biases in the propensity to patent depending on the industry of origin, the size of the firm and the type of invention; 4) there are important reliability problems in patent data; 5) some patents prove to have an economic value, but the vast majority do not; 6) many patents are of a purely defensive nature; 7) patent requirements have evolved drastically over time and geographical space (see Griliches 1990; Desrochers 1998 and Kleinknecht 2002 for discussions).

Can patents be used to proxy innovations? Surely, patents have been used to proxy both inventive and innovative activity. In principle, a patent is not always used commercially for strategic reasons, but there is a strong inclination for doing so since the investment costs are high.[9] Secondly, a precondition for patenting is that it should be "industrially applicable" (Michel and Bettels, 2001).[10] In addition, recent evidence by Acs et al. (2002) suggest that patents are in fact very good proxies for "true" innovations. On balance, I conclude that patents can be used to proxy innovative activity and that their classification circumvent many of the problems of using industry classifications. Patent documents give information on where inventor(s) live, the applying organization, and we know what patents are cited etc. The regional location can for instance be obtained with high precision. We now turn to the diversity issue, i.e. how we can measure the degree of diversity.

4.1 The Technological Closeness of IPC-classes

In the literature, methods to relate patent classes have principally followed three routes, depending on what has been the research focus. One strand of analysis adopts the Yale concordance table, first used by Putnam and Evenson (1994), to assign patents to industries by sector of use and sector of production.[11] For

[9] Precise figures are hard to obtain, since this seems to vary by technology, application costs and practices by different law firms assisting the inventor(s).

[10] Although Michel and Bettels (2001) indicate that this requirement may to be too restrictive.

[11] Andersson and Ejermo (2004) use an improved version of this table to estimate knowledge production functions for sectors of Swedish regions.

our purpose, as has already been discussed, the industry classification is not a viable option. The second route uses the citation of patents to others to infer "spillovers" between classes.[12] Our analysis is not interested in the spillover dimension, but instead in the relatedness of each pair of technologies in use.

Therefore, we use the third route, which takes advantage of how close two IPC-classes are by examining the main and supplementary classification of patents. Verspagen (1997) claim that the main classification of a patent classification can be used to infer "spillovers" to the supplementary class of the patent. In this way, he forms a spillover matrix. However, in the work by the Frauenhofer Institute of Karlsruhe jointly with CESPRI of Milan and described in Hinze et al. (1997), no support could be found for the spillover hypothesis between the classes, because this is not the procedure by which patent examiners are instructed to add the supplementary codes. Rather, supplementary class information reflect further claims of the patents. Therefore, Breschi et al. (2002, 2003) conclude that this is a relevant measure of how technologically close classes are. We now describe this method in detail, because we will use it to construct measures of regional diversity.

Hinze et al. (1997) create a table which relates classes based on how often two classes occur jointly in main and supplementary classes. Dividing the material into 30 classes of the three countries USA, Germany and Japan over the years 1982-1993 such a classification proves to be robust both over the countries, and over time, from data of the European Patent Office (EPO). In their use of this material, Breschi et al. (2002 and 2003), devise a measure based on technological likeness. The degree to which there is co-occurrence in the classification of two classes, is denoted w_{ij}, where i and j are two of 30 classes and w_{ij} is the raw count of the number of co-occurrences. In principle, a measure could use only this relation, but it makes more sense also to use *indirect* relations. This can be done using the cosine index (cf. Jaffe 1986) which jointly relates the similarities, not only between groups i and j, but also at the same time takes into account their similarity to other classes:

$$q_{ij} = \frac{\sum_{k=1}^{n} w_{ik} w_{jk}}{\sqrt{\sum_{k=1}^{n} w_{ik}^2 \sum_{k=1}^{n} w_{jk}^2}} \tag{1}$$

[12] Such analyses have for instance been done by Verspagen (1997) and Verspagen and Loo (1999).

If $i = j$, the similarity q_{ij} becomes 1 and when $i \neq j, 0 \leq q_{ij} \leq 1$. The result using EPO patents from six major OECD countries from the EPO patent database is taken from Breschi et al. (2002) and is reproduced in Table 3 of Appendix 7.1. These relations between classes are used also in this paper. The underlying assumption is that the stability reported on by Hinze et al. (1997) carries on to Swedish patenting. Higher amounts of patenting in different classes can be the result of higher technological opportunities (cf. Dosi 1988) and higher propensities to patent (Scherer 1983, Breschi et al. 2000). Therefore, it is to our advantage that this information is 'built into' (1).

We now take into account patenting in different technologies and their mutual relationship through (1) in a single compound measure that reflects diversity.

4.2 Weighted Average Relatedness Measures

Teece et al. (1994) used two measures to operationalise firms' activities relative strength in a network (also used in Breschi et al., 2002, 2003). We apply one of them for measuring regional diversity and adjust the notation. Before describing the data we note that there are 81 regions used in the empirical analysis. Thus, each region ($r = 1, ..., 81$) may have some activity in one of patent technology fields $i = 1, ..., 30$ and a relation between two classes i and j denoted by q_{ij}. The *weighted-average-relatedness* (WAR) of region r:s technological activities, outside a given class i, is written:

$$WAR_{ir} = \frac{\sum\limits_{j \neq i} q_{ij} p_j}{\sum\limits_{j \neq i} p_j} \tag{2}$$

where q_{ij} is the weight of activities in each technology field j with respect to technology i for the region. In the case of patents, p_j is the share of patents in class j. Teece et al. (1994, p. 14-15) write:

> "The index measures the degree to which technology field i is linked to all of the other activities of firm k, both in terms of "technical distance" and relative weight. For each firm, an average value of WAR_i can thus be calculated to get an index of global technological coherence. Further average values can be computed for specific firm categories, such as firms with the same size (in our case, size is given by the number of patents held by the firm), a similar diversi-

fication range (the number of technological field in which the firm is found to be active)"

The measure's main weakness is, however, its dependence on the individual region's (or firm's in the case of Teece et al., 1994) diversification range. The more technological fields the firm adds to its portfolio, the more "weak links" between those fields and field i (low q_{ij} values) will be added to the index.

A remedy to this problem is to use the *weighted-average-relatedness of neighbours* ($WARN$). This measure takes into account only those links that belong to the so-called maximum spanning tree, i.e. those $(n-1)$ links which are strictly necessary for creating a connected graph between a unit's technological activities, and at the same time show the strongest connections. The $(n-1)$ links refer to the fact that regions may have no patents in several classes and the number of links necessary to form therefore differs. Originally, Teece et al. (1994) used WAR and $WARN$ to measure systematic diversification of firms' employment across different activities. The striking result was that diversification was systematic, so that activities were added to the firm's portfolio of activities reelated to those previously undertaken. Similarly, the $WARN$ (and WAR) measure was used by Breschi et al. (2003) to examine whether companies in the U.S., Japan, France, Germany, Italy and the United Kingdom filing applications to the EPO over the period 1982-1993 period diversified in a systematic way. They found that diversification into new technologies was more likely when the existing portfolio of technologies of companies was technologically related to the new technology. Thus, in Breschi et al.'s (2003) analysis, higher values of $WARN$ reflect higher technological coherence for firms. For the present paper, a higher value is assumed to measure higher connection between the patenting activities in the region, interpreted as higher specialization towards similar technologies. Smaller values show that patenting activity in regions are more diverse.

Mathematically we can write the $WARN$ measure as follows. Denoting for region r, class i

$$WARN_{ir} = \frac{\sum_{j \neq i} q_{ij} \lambda_{rij} p_{rj}}{\sum_{j \neq i} \lambda_{rij} p_{rj}} \tag{3}$$

where $\lambda_{rij} = 1$ if the link between i and j belongs to the maximum spanning tree that relates region r's activities, and $\lambda_{rij} = 0$ otherwise. Examining $WARN_{ir}$ we see that it is constrained between 0 and 1. Appendix 7.2 gives a practic numerical example of how this spanning tree was calculated. Equation (3) can be modified using the number of patents as weights instead, since the following equalities hold:

$$WARN_{ir} = \frac{\sum\limits_{j \neq i} q_{ij}\lambda_{rij}p_{rj}}{\sum\limits_{j \neq i} \lambda_{rij}p_{rj}} = \frac{\sum\limits_{j \neq i} q_{ij}\lambda_{ij}\frac{P_{rj}}{\sum\limits_{j} P_{rj}}}{\sum\limits_{j \neq i} \lambda_{ij}\frac{P_{rj}}{\sum\limits_{j} P_{rj}}} = \frac{\sum\limits_{j \neq i} q_{ij}\lambda_{ij}P_{rj}}{\sum\limits_{j \neq i} \lambda_{ij}P_{rj}} \tag{4}$$

where P_{rj} is the number of patents in class j for region r. where c_r Taking the average over the classes for each region gives us a measure for the overall technological relatedness of patenting in a region.

$$WARN_r = \frac{\sum\limits_{i} WARN_{ir}}{n_r} = \frac{\sum\limits_{i}\left(\frac{\sum\limits_{j \neq i} q_{ij}\lambda_{rij}P_{rj}}{\sum\limits_{j \neq i} \lambda_{rij}P_{rj}}\right)}{n_r} \tag{5}$$

where n_r denotes the number of classes in which region r has patents. Going back to examine the components of equation (4), we see that for any component $q_{ij}P_{rj}$ being added to the numerator, a component P_{rj} is added to the denominator. In a "network" with only patenting in one class, this would reduce to $q_{ij}P_{rj}/P_{rj} = q_{ij}$. In a two-class example the equation reduces to a weighted sum of the two q_{ij}'s. This means that there is no *a priori* reason to think that an increase in the number of classes would by itself change the size of $WARN_{ir}$. There is also no reason to think that the weights for the different link strengths, P_{rj}, would be different in a random network. A randomly simulated $WARN_{ir}$ would therefore have equal weights with the random link simply being the average value of *all* possible $E\left[q_{ij}\right]_{j \neq i} = \sum\limits_{j \neq i} q_{ij}/n_r \equiv \bar{q}$. Thus, the expected value of $WARN_{ir}$ and $WARN_r$ is simply:

$$E\left[WARN_{ir}\right] = E\left[WARN_r\right] = \bar{q} \tag{6}$$

$\bar{q}$ is about 0.0457 with the values for q_{ij} from Table 3 in use. $WARNDIFF_r$ is defined as the calculated value $WARN_r$ minus the non-random, expected part:

$$WARNDIFF_r = WARN_r - \bar{q} \tag{7}$$

Deducting $\bar{q}$, removes a non-random part from the diversity measure which has been imposed due to the construct of the technological similarity measure. In other words, there is always some technological closeness that arises because all the classes are related to some extent. To obtain the random part we remove this. The technological similarity of patenting in a region consists of two parts: 1. The frequency of patenting, and 2. the 'closeness' of these classes. These two can safely be assumed to be inseparable and jointly determined in a region. That is, the choice of how many patents to produce in a specific class is assumed to be a function of both how many patents there were previously, and how related they are technologically. The next section illustrates various features of $WARNDIFF_r$.

5 Diversity in Swedish Regions

We now go on to describe the acquired data set of patents for Swedish regions with emphasis on diversity. Patent data from EPO exist from 1977 to the present, following the establishment of the European Patent System. The system was gradually developed in the first years and can from inspection of the data, not be regarded to have been fully at work until 1982. The last year from which we use data is 1999, given that these are based on the priority date.[13] The setup considered here concerns testing the hypothesis that diversity in one period affects new innovations, proxied by the number of patent applications in a later period. The patent data consisting of EPO data with at least one Swedish inventor, was divided into two samples spanning equally long time periods, the first from 1982-1990 and the second from 1991-1999. $WARNDIFF_r$ was calculated for regions in the first period and the number of patents were summed together for the same regions in the second period.

Figure 1 shows the number of patent applications in 81 Swedish functional regions, i.e. local labour markets, 1991-1999. A functional region is determined by its economic rather than administrative properties, because they are delimited by their commuting patterns. A patent application is located to a specific region depending on the residence of the first inventor in the list of inventors. If the inventor for some reason could not be located, the location of number 2 was used and so on.[14]

[13] It takes years for data to go through the system of patent application after they are initially given a priority date. Hence, many patents filed after 1999 are not in the database.

[14] This method was preferred, since, as was also pointed out by an anonymous referee, using the location of the company may introduce bias because it may reflect the location of headquarters, rather than the actual location of the R&D-lab. Using the inventor's address should avoid this problem.

Figure 1: Patenting activity 1991-1999 in Swedish functional regions to the European Patent Office.

It can be observed that patenting activity[15], is highly concentrated to the three most populated regions: Stockholm, Göteborg and Malmö. Of course, this has to do with the relatively high population concentration in these regions, but this cannot be the only explanation; if we count the number of patents per employees, we find that these regions still score among the top five. The two top-scoring regions are Västerås and Ludvika. Västerås is ranked number 11 in terms of population, while Ludvika is quite small. Both are important production localities for the Swedish part of the power and automation technology company ABB.[16]

Figures 2 and 3 illustrate different properties of $WARNDIFF_r$. Figure 2 illustrates regional values. $WARN$-values could not be calculated for regions without patents and those with patenting in only one class. Compared to Figure

[15] Henceforth referring to patent applications.

[16] The top 20 regions in terms of population and top 20 in terms of patents per employee 1991-1999 are listed in Appendix 7.3, Table 5.

Figure 2: Technological specialisation measured by average $WARN$ in Swedish functional regions over the 1982-90 period.

1, we see that high values of $WARNDIFF_r$ tends to be where patenting is also extensive. This means that technological specialisation tends to go hand in hand with high levels of patent applications. The values range from -0.0147 for the Åre region[17] to 0.1533 for the Göteborg region. We can also examine $WARNDIFF_r$ with respect to the number of classes in which a region has patenting. Figure 3 illustrates this.

Figure 3 plots $WARNDIFF_r$, the adjusted values for each region in relation to the number of technology classes the region has. We see that there is a clear tendency for $WARNDIFF_r$ to go up as regions add on more classes. This therefore seems to indicate that as more classes are added, regions tend to patent in classes which are technologically close and in which they may build on related experience from other technologies. Breschi and Lissoni (2002) obtained rising $WARNDIFF_r$-values for firms, but with a much faster rise in its value. The lack of this rising speed of $WARNDIFF_r$ probably reflects the fact that some regions contain many firms, and larger regions (those most likely to have patent

[17] This negative value arises only because we deduct the expected value.

Figure 3: Calculated $WARNDIFF$-values in relation to the number of classes with patenting for Swedish functional regions, 1982-1990.

activities in many classes) have of course, a lot of them. Hence, $WARNDIFF_r$, tends to be diluted compared to an increase of classes based on a pure firm calculation of the value.

Other variables are likely to affect patent production concurrently. The next section tests whether innovation, as measured by the number of patents, is affected by diversity, while controlling for other variables.

5.1 A Test of the Effects of Diversity on New Patents

The adjusted $WARNDIFF$-measure captures technological diversity, which seems like a reasonable interpretation of Jacobs' externality hypothesis. With regard to the Marshallian externalities, mainly product effects are highlighted in theoretical discussions, but Marshall (1920, Book IV, p 225) also speaks about that "the mysteries of the trade become no mysteries; but are as it were in the air". This has widely been interpreted as "knowledge spillovers". Since suppliers of local intermediate goods and reserves of labour pool act as providers of knowledge, higher specialization is thought to induce higher amounts of "spillovers" and hence more innovations can build on these ideas. Therefore, while Marshallian externalities is a mixture of many effects, they can to some extent be thought to run counter to Jacobs' arguments.

To test formally the connection with innovation, we regress the number of patents applied for from 1991-1999 on diversity measured by $WARNDIFF$ from 1982-1990. The use of an earlier time period for the diversity variable takes into account the implicit assumption that development of innovations should be seen as dependent on previous diversity/specialization. From our introductory discussion of the importance of agglomeration, we conclude that there should be a general tendency for more densely populated regions to exhibit agglomeration effects. We therefore include a measure of employment to control for this effect, which should also reduce potential problems of heteroskedasticity. $EMP9199$ measures the sum of employment in the 1991-1999 period. In addition, there may be path dependence of regional innovative activity, i.e. regions with high activity in the past are more likely to patent more in the future.

The literature has identified that certain sectors have different propensities to patent. Regions may also exhibit path dependence in their innovation behaviour. We therefore use the number of patents in the 1982-1990 period in each region, as a proxy for 'opportunities' in the region of future innovative activity.

Our starting model is of the type:

$$PAT9199_r = f\left(\overset{-/+}{WARNDIFF_r}, \overset{+}{PAT8290_r}, \overset{+}{EMP9199_r}\right) \tag{8}$$

where $PAT9199_r$ is the number of patents applied for 1991-1999 in region r modeled as a function of $WARNDIFF_r$ and $PAT8290_r$, the number of applications 1982-1990. Above the variables is shown the expected relationship. If Jacobs' externalities are more inducive to future patents, $WARNDIFF_r$ should show a negative sign, since regions are less diverse in their technological activities when $WARNDIFF_r$ rises. If localization externalities are more important, in line with Marshall's suggestion i.e. specialization of economic activity, the sign should be positive.

A number of authors have commented on the appropriate technique for estimation when the dependent variable is discrete, and ≥ 0 as in our case (cf. Cameron and Trivedi, 1998; Long, 1997). The basic setup for such data is the Poisson regression. It is outlined by the following formulation:

$$\Pr[PAT_r = pat_r] = \frac{e^{-\lambda_r} \lambda_r{}^{pat_r}}{pat_r!}, \qquad pat_r = 0, 1, 2, 3, ... \tag{9}$$

where λ_r is in turn related to the set of regressors $\mathbf{x}_r$, normally in a linear fashion:

$$\ln \lambda_r = \boldsymbol{\beta}' \mathbf{x}_r \tag{10}$$

A problem that often arises for users of the Poisson model is overdispersion. Overdispersion refers to when the variance is higher than the expected value.[18] The average number of patent applications 1991-1999 in regions where $WARN$ could be calculated is about 182 and the variance 354,198. Thus the high variance observed suggest that there is a problem of overdispersion (see Long 1997, p. 220). Intuitively, ovderdispersion can be understood as heterogeneity (e.g. productivity) in the observed units. The negative binomial model does not impose the restriction of equality between expected value and variance. More generally, the process is then written:

$$\ln \tilde{\lambda}_r = \boldsymbol{\beta}' \mathbf{x}_r + \varepsilon \tag{11}$$

where e^{ε} is normally assumed to have a gamma distribution with mean 1 and variance α, which is estimated from the data. The probability distribution of this negative binomial model is:

$$\Pr[PAT_k = pat_k \mid \varepsilon] = \frac{e^{-\tilde{\lambda}_r} \lambda_r{}^{pat_r}}{pat_r!}, \qquad pat_r = 0, 1, 2, 3, ... \tag{12}$$

It can be tested if the estimate of α equals zero, in which case we retain the Poisson model. A value significantly higher indicates overdispersion. If $H_0 : \alpha = 0$ is rejected, it is taken as evidence in favour of the negative binomial model. The test is implemented by computing the likelihoods of the restricted model $L\left(\tilde{\theta}_r\right)$, i.e. $\alpha = 0$ which results in the Poisson model, and the likelihood from the unrestricted model, $L\left(\hat{\theta}_u\right)$. The likelihood-ratio statistic is given by

$$LR_{\alpha=0} = -2\left[L\left(\tilde{\theta}_r\right) - L\left(\hat{\theta}_u\right)\right] \tag{13}$$

[18] $E[pat_k \mid \mathbf{x}_k] = \mathrm{Var}[pat_k \mid \mathbf{x}_k] = \lambda_k = e^{\boldsymbol{\beta}' \mathbf{x}_k}$. See for example Hausman et al. (1981) or Greene (2000, pp. 880-886) for discussions on the properties of Poisson and negative binomial models.

Table 1 shows a correlation matrix along with some descriptive statistics of our data.

Table 1: Correlation matrix and descriptive data of variables.

Variable	PAT9199	WARNDIFF	PAT8290	EMP9199
PAT9199	1			
WARNDIFF	0.4982	1		
PAT8290	0.9961	0.5327	1	
EMP9199	0.9862	0.5533	0.9867	1
min	1	-0.0147	2	14,997
max	4,492	0.1533	2201	7,729,353
mean	182.4286	0.0641	103.5714	498,932.3429
std. dev.	595.1454	0.0372	294.8615	1,022,604.464

All variables show positive correlation with each other. Not surprisingly, we find a high positive correlation between employment size and the amount of patenting in regions, and of patenting in the two periods (close to unity). Also, $WARNDIFF$ displays substantial correlation with patents in the 1982-1990 period and employment 1991-1999. Because of high correlation between $PAT8290$ and $EMP9199$, we exclude them in two models. Also, there are two variants of estimations: a) the Poisson model and b) the negative binomial model. The likelihood-ratio statistic below each estimated negative binomial model indicate if the Poisson model was rejected based on the test indicated in 13. Each model has the number of patent applications 1991-1999 as the dependent variable. Model (1) excludes from the list of independent variables $EMP9199$, Model (2) excludes $PAT8290$ and Model (3) includes all variables, i.e. $WARNDIFF, PAT8290$ and $EMP9199$. The results are shown in Table 2 along with the marginal effects of parameters.

Likelihood ratio tests of the overdispersion parameter α strongly reject the null hypothesis in all cases, that is, the negative binomial model seems to be the most appropriate one. The results show that higher technological specialization, $WARNDIFF$, in a region raises the likelihood for regional innovations. This result is robust over both the Poisson and the Negative binomial model and across specifications, and is highly significant. Moreover, old patents ($PAT8290$) raise the likelihood for more patents in all models where it is included. Employment shows an incoherent pattern, which is most likely attributable to the multicollinearity problem. It is positive and strongly significant in two cases (2a and 2b), negative and strongly significant in one case (3a) and insignificant in one case (3b). The marginal effects are the parameter values which we are normally interested in, in a likelihood setting such as this one. The marginal effects evalued at variable means, are given in the lower half of the table. Increasing $WARNDIFF$ "marginally" means increasing it by one, which is a very high number in this case, since it is constrained betwen 0 and 1 (see Section 4.2). If we do so, the number of patents are expected to increase by 1272-1530 (all models) if we start out at the average $WARNDIFF$ (0.0641). The parameter value for $WARNDIFF$ is also similar among most types of models (about 24), except for model (2b) where it is very low (but still highly significant and positive). Increasing the number of patents in the base period, 1982-1990, has a "smaller" effect. A rise of one patent 1982-90, from the average of 104 patents in this period, is expected to raise the number of patents 1991-1999 by 0.0251-0.1129. The marginal effect of raising the number of people employed in the region by one person, has different effects depending on the model. In the full negative binomial model (3b), it is not significantly different from zero, so we abstain from drawing conclusions.

Table 2: Estimation results of regression analysis along with marginal effect estimates. Dependent varia
errors are in parenthesis. Models marked (a) indicate that the Poisson model was used, (b) the use o
model. Values rounded to four decimals, except standard errors and marginal effects which are rounde
correspond to significance at 10, 5 and 1 per cent levels

Variable	Model (1a)	Model (1b)	Model (2a)	Model (2b)	Mod
constant	2.4680 (0.03)***	2.3630 (0.25)***	2.3951 (0.03)***	2.3153 (0.24)***	2.5016 (0
WARNDIFF	24.2248 (0.31)***	23.2981 (4.09)***	24.5346 (0.31)***	6.46E-07 (2.48E-07)***	24.1085 (0
PAT8290	0.0012 (1.18E-05)***	0.0020 (7E-04)***	-	-	0.0018 (2E
EMP9199	-	-	3.33E-07 (3.41E-09)***	22.0958 (4.33)***	-1.72E-07 (5.37E
nobs	70	70	70	70	
pseudo R^2	0.8990	0.1319	0.8971	0.13	
LR-test. α=0	-	3514.19	-	3592.90	
			Marginal effects of coefficients at variable mean		
WARNDIFF	1521.448 (19.11)***	1347.82 (273.44)***	1530.2510 (19.16)***	1272.424 (280.42)***	1517.3210 (19
PAT8290	0.0727 (1.4E-03)***	0.1129 (0.04)***	-	-	0.1104 (0
EMP9199	-		2.08E-05 (4.16E-07)***	3.72E-05 (1.48E-05)***	-1.08E-05 (3.39E

The most important result from these investigations is of course, that technological specialization seems to foster higher innovative output measured by the number of patents. It should be noted that the regional unit of analysis, local labour market regions, may in Sweden contain only few patent innovators in those cases where the region is relatively small. In these cases the specialization measure largely reflects the diversification ranges of individual innovators. As shown by previous contributions using $WARN$ (Teece et al., 1994 and Breschi, et al. 2003), firms tend systematically to diversify their patenting activities into technological areas close to those they previously engaged in. If this is the case here, as seems highly likely, this seems to drive the results. In larger regions, with presence of more and larger firms, these are likely to engage in more types of technological activities, raising $WARNDIFF$ of the region if their patenting behaviour can be attributed greater technological coherence.

6 Conclusions and Extensions

This paper has examined the importance of diversity for regional innovativeness, along the lines suggested by Jacobs (1969) and Desrochers (1998). Diversity, as measured over Swedish local labour market regions, has been used in a novel way to examine the diversity of Swedish regions. The measure used takes into account technological diversity using both the relatedness of technologies in patenting and the number of patents in individual classes. The results indicate that higher technological specialization improves patenting productivity among Swedish regions. This runs counter to the findings of earlier contributions which do not take technological diversity into account. Even though there are limits to patents as a measure of innovation and many patent have only indirect economic value, they are useful as indicators of technological activity. The present study highlights some additional important differences to the studies in the literature. Earlier contributions tend to analyze regional diversity using administrative units such as American states. The present study uses local labour market regions, which is well motivated if we want to analyze the economic outcome of policy-decisions, since policy-changes on a regional level are largely confined within the borders of the local labour market region. In this case, the regional unit is therefore smaller, and hence the number of innovations and the diversification range tends to be sensitive to individual firms' diversification range. As regions get larger, we may conjecture that firms in Swedish regions increase their range of activities in a systematic way to technological areas close to those already existing in the firm's portfolio. Hence, the average relatedness tend to go up for the individual firms in the larger regions, as do the weighted average relatedness of the whole region. Although the empirical results are only for one country, technological diversity should be seriously considered as an explanatory factor for a researcher evaluating the importance of diversity.

Innovations have here been proxied for by patent applications. A possible way to remedy Jacob's ideas in this setting, is to examine applications from firms that have not patented before. If many applications are based on old knowledge, it is not so strange that the technological trajectory follows established paths. However, innovation of a new patentee may follow a different route, and possibly draw on a more diverse set of regional forces. This may be particularly important in the Swedish case, with a dominance of R&D in multinationals with in many cases century-long tradition of continuous innovation in related technologies. In those cases, patent applications may reflect incremental rather than radically new ideas. This also points to the possibility to give weight to patents by the number of times they have been cited to reflect the "size" of new ideas.

The main conclusions/recommendations from this paper are that technological diversity needs to be addressed to better examine the diversity hypothesis; that regions reflect the technological range of individual firms and hence a division on firm-level data coupled, with care taken to the size of the regions under analysis should be considered in future research.

References

Acs, Z., L. Anselin, and A. Varga (2002): "Patents and Innovation Counts as Measures of Regional Production of New Knowledge," *Research Policy*, 31, 1069–1085.

Acs, Z., F. FitzRoy, and I. Smith (2002): "High-Technology Employment and R&D in Cities: Heterogeneity vs Specialization," *Annals of Regional Science*, 36(3), 373–386.

Andersson, M., and O. Ejermo (2004): "Sectoral Knowledge Production in Swedish Functional Regions 1993-1999," in *Forthcoming in Knowledge Spillovers and Knowledge Management*, ed. by C. Karlsson, P. Flensburg, and S.-Å. Hörte. Edward Elgar: Cheltenham.

Arrow, K. J. (1962): "The Economic Implications of Learning-by-Doing.," *Review of Economic Studies*, 29(1), 155–173.

Audretsch, D., and M. Feldman (1996): "R&D Spillovers and the Geography of Innovation and Production," *American Economic Review*, 86(3), 630–640.

Autant-Bernard, C. (2001): "The Geography of Knowledge Spillovers and Technological Proximity," *Economics of Innovation and New Technology*, 10, 237–254.

Breschi, S., F. Lissoni, and F. Malerba (2002): "The Empirical Assessment of Firms' Technological "Coherence": Data and Methodology," in *The Economics and Management of Technological Diversification*, ed. by J. Cantwell, A. Gambardella, and O. Granstrand, Routledge Studies in the Modern World Economy. London: Routledge.

——— (2003): "Knowledge-Relatedness in Firm Technological Diversification," *Research Policy*, 32, 69–87.

Breschi, S., F. Malerba, and L. Orsenigo (2000): "Technological Regimes and Schumpeterian Patterns of Innovation," *The Economic Journal*, 110(463), 388–410.

Cameron, A., and P. Trivedi (1998): *Regression Analysis of Count Data*. Cambridge: Cambridge University Press.

Desrochers, P. (1998): "On the Abuse of Patents as Economic Indicators," *Quarterly Journal of Austrian Economics*, 1(4), 51–74.

——— (2001): "Local Diversity, Human Creativity, and Technological Innovation," *Growth and Change*, 32, 369–394.

Dixit, A., and J. E. Stiglitz (1977): "Monopolistic Competition and Optimum Product Diversity," *American Economic Review*, 67(3), 297–308.

Dosi, G. (1988): "Sources, Procedures and Microeconomic Effects of Innovation," *Journal of Economic Literature*, 26, 1120–1126.

Duranton, G., and D. Puga (2001): "Nursery Cities: Urban Diversity, Process Innovation, and the Life-Cycle of Products," *American Economic Review*, 91(5), 1454–1477.

Ellison, G., and E. Glaeser (1997): "Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach," *Journal of Political Economy*, 10, 889–927.

Feldman, M. P., and D. B. Audretsch (1999): "Innovation in Cities: Science-Based Diversity, Specialization and Localized Competition," *European Economic Review*, 43, 409–429.

Glaeser, E. L., H. D. Kallal, J. A. Scheinkman, and A. Shleifer (1992): "Growth in Cities," *Journal of Political Economy*, 100(6), 1126–1152.

Greene, W. (2000): *Econometric Analysis*. Prentice Hall: New Jersey.

Griliches, Z. (1990): "Patent Statistics as Economic Indicators: A Survey.," *Journal of Economic Literature*, 28, 1661–1707.

Hall, M. (1959): *Made in New York*. Cambridge MA: Harvard University Press.

Harrison, B., M. R. Kelley, and J. Gant (1996): "Specialization versus Diversity in Local Economies: The Implications for Innovative Private-Sector Behavior," *Cityscape*, 2(2), 61–93.

Hausman, J., B. Hall, and Z. Griliches (1981): "Econometric Models for Count Data with an Application to the Patents-R&D Relationship," Working Paper 17, NBER.

HENDERSON, J. (1974): "The Sizes and Types of Cities," *American Economic Review*, 64(4), 640–656.

HENDERSON, V., A. KUNCORO, AND M. TURNER (1995): "Industrial Development in Cities," *Journal of Political Economy*, 103(5), 1067–1090.

HINZE, S., T. REISS, AND U. SCHMOCH (1997): "Statistical Analysis on the Distance Between Fields of Technology," Innovation Systems and European Integration (ISE), Targeted Socio-Economic Research Program, 4th Framework Program of the European Commission (DGXII), Frauenhofer-Institute Systems and Innovation Research (ISI), Karlsruhe Germany.

HOOVER, E. M. (1937): *Location Theory and the Shoe and Leather Industries.* Cambridge Mass.: Harvard University Press.

JACOBS, J. (1969): *The Economy of Cities.* New York: Vintage.

JAFFE, A. (1986): "Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents Profits and Market Value," *American Economic Review*, 76, 984–1001.

KELLEY, M., AND S. HELPER (1999): "Firm Size Capabilities, Regional Agglomeration and the Adoption of New Technology," *Economics of Innovation and New Technology*, 8(1-2), 79–103.

KELLY, M., AND A. HAGEMAN (1999): "Marshallian Externalities in Innovation," *Journal of Economic Growth*, 4, 39–54.

KLEINKNECHT, A., K. VAN MONTFORT, AND E. BROUWER (2002): "The Non-Trivial Choice Between Innovation Indicators," *Economics of Innovation and New Technology*, 11(2), 109–121.

LEVIN, R., A. KLEVORICK, R. NELSON, AND S. WINTER (1987): "Appropriating the Returns from Industrial Research and Development," *Brookings Papers on Economic Activity*, 1987(3), 783–820.

LONG, J. (1997): *Regression Models for Categorical and Limited Dependent Variables.* California: Thousand Oaks.

MARSHALL, A. (1920): *Principles of Economics.* London: MacMillan.

Michel, J., and B. Bettels (2001): "Patent Citation Analysis - a Closer Look at the Basic Input Data from Patent Search Reports," *Scientometrics*, 51(1), 185–201.

Mills, E. (1967): "An Aggregative Model of Resource Allocation in a Metropolitan Area," *American Economic Review Papers and Proceedings*, 57, 197–210.

Ohlin, B. (1933): *Interregional and International Trade*. Cambridge, Mass.: Harvard University Press.

Paci, R., and S. Usai (1999): "Externalities, Knowledge Spillovers and the Spatial Distribution of Innovation," *GeoJournal*, 49, 381–390.

Pettersson, L. (2001): "Location, Housing and Premises in a Dynamic Perspective," Ph.D. thesis, Jönköping International Business School, Ch. 5, coauthored with J. Klaesson.

Pinch, S., and N. Henry (1999a): "Discursive Aspects of Technological Innovation: The Case of the British Motor-Sport Industry," *Environment and Planning A*, 31, 665–682.

——— (1999b): "Paul Krugman's Geographical Economics, Industrial Clustering and the British Motor Sport Industry," *Regional Studies*, 33, 815–827.

Porter, M. (1990): *The Competitive Advantage of Nations*. New York: Free Press.

Putnam, J., and R. E. Evenson (1994): "Inter-Sectoral Technology Flows: Estimates from a Patent Concordance with an Application to Italy," Mimeo.

Quigley, J. M. (1998): "Urban Diversity and Economic Growth," *Journal of Economic Perspectives*, 12(2), 127–138.

Romer, P. M. (1986): "Increasing Returns and Long Run Growth," *Journal of Political Economy*, 94(5), 1002–1037.

Scherer, F. (1983): "The Propensity to Patent," *International Journal of Industrial Organization*, 1, 107–128.

Spence, M. (1976): "Product Selection, Fixed Costs, and Monopolistic Competition," *Review of Economic Studies*, 43(2), 217–235.

Storper, M. (1995): "The Resurgence of Regional Economies Ten Years Later: The Region as a Nexus of Untraded Interdependencies," *European Urban and Regional Studies*, 2, 191–221.

Storper, M., and R. Salais (1997): *Worlds of Production: The Action Framework of the Economy*. Harvard University Press, Cambridge MA.

Teece, D. J., R. Rumelt, G. Dosi, and S. Winter (1994): "Understanding Corporate Coherence - Theory and Evidence," *Journal of Economic Behavior and Organization*, 23, 1–30.

van Oort, F. (2002): "Innovation and Agglomeration Economies in the Netherlands," *Tijdschrift voor Economische en Sociale Geografie*, 93(3), 344–360.

Verspagen, B. (1997): "Measuring Intersectoral Technology Spillovers: Estimates from the European and US Patent Office Databases," *Economic Systems Research*, 9, 47–65.

Verspagen, B., and I. de Loo (1999): "Technology Spillover Between Sectors and over Time," *Technological Forecasting and Social Change*, 60, 215–235.

7 Appendix

7.1 Similarity of Technological Fields

Table 3: Relatedness of 30 patent technology classes, used as a measure of technological similarity, q_{ij}, in the main text. Source: Adapted from Breschi et al. (2003).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	1																			
2	0.056	1																		
3	0.06	0.154	1																	
4	0.033	0.094	0.143	1																
5	0.116	0.038	0.037	0.054	1															
6	0.068	0.098	0.067	0.038	0.102	1														
7	0.083	0.05	0.087	0.112	0.043	0.062	1													
8	0.022	0.016	0.011	0.024	0.008	0.037	0.08	1												
9	0.01	0.008	0.003	0.004	0.006	0.048	0.101	0.03	1											
10	0.057	0.019	0.004	0.003	0.024	0.095	0.018	0.053	0.148	1										
11	0.005	0.004	0.002	0.004	0.003	0.018	0.097	0.068	0.755	0.089	1									
12	0.01	0.007	0.009	0.014	0.006	0.018	0.257	0.042	0.477	0.059	0.479	1								
13	0.101	0.016	0.005	0.004	0.079	0.04	0.023	0.021	0.043	0.086	0.026	0.019	1							
14	0.009	0.002	0.001	0.003	0.003	0.008	0.025	0.03	0.141	0.043	0.161	0.232	0.017	1						
15	0.027	0.015	0.002	0.003	0.012	0.071	0.033	0.032	0.411	0.187	0.214	0.173	0.092	0.119	1					
16	0.031	0.008	0.006	0.013	0.018	0.022	0.095	0.061	0.152	0.083	0.073	0.082	0.165	0.081	0.171	1				
17	0.12	0.027	0.009	0.007	0.128	0.062	0.031	0.035	0.034	0.159	0.016	0.015	0.191	0.02	0.068	0.102	1			
18	0.045	0.021	0.004	0.007	0.018	0.053	0.03	0.05	0.036	0.268	0.018	0.021	0.097	0.023	0.113	0.098	0.189	1		
19	0.078	0.006	0.008	0.007	0.019	0.01	0.06	0.019	0.009	0.012	0.004	0.012	0.124	0.017	0.033	0.104	0.037	0.036	1	
20	0.019	0.003	0.002	0.003	0.01	0.011	0.026	0.025	0.045	0.048	0.021	0.038	0.151	0.034	0.097	0.373	0.067	0.048	0.133	1
21	0.053	0.008	0.005	0.009	0.029	0.026	0.049	0.024	0.007	0.024	0.004	0.008	0.083	0.012	0.024	0.043	0.072	0.09	0.042	0.021
22	0.047	0.005	0.008	0.007	0.009	0.006	0.052	0.019	0.004	0.005	0.003	0.008	0.028	0.003	0.011	0.042	0.019	0.013	0.084	0.058
23	0.057	0.012	0.008	0.009	0.01	0.013	0.055	0.034	0.004	0.018	0.002	0.007	0.031	0.006	0.012	0.04	0.047	0.069	0.067	0.036
24	0.035	0.027	0.015	0.034	0.021	0.056	0.06	0.036	0.012	0.038	0.007	0.011	0.021	0.037	0.023	0.066	0.102	0.116	0.026	0.022
25	0.012	0.005	0.003	0.006	0.004	0.006	0.036	0.021	0.021	0.015	0.022	0.05	0.012	0.144	0.038	0.062	0.018	0.026	0.026	0.031
26	0.051	0.011	0.015	0.01	0.007	0.013	0.055	0.014	0.003	0.019	0.002	0.006	0.013	0.003	0.005	0.015	0.031	0.038	0.037	0.021
27	0.092	0.025	0.019	0.019	0.03	0.045	0.082	0.063	0.012	0.014	0.009	0.018	0.064	0.007	0.038	0.045	0.046	0.016	0.04	0.04
28	0.019	0.008	0.018	0.011	0.009	0.023	0.075	0.007	0.01	0.013	0.006	0.014	0.02	0.004	0.011	0.025	0.03	0.021	0.019	0.013
29	0.035	0.034	0.011	0.022	0.011	0.021	0.056	0.073	0.007	0.03	0.007	0.009	0.02	0.03	0.016	0.045	0.077	0.084	0.058	0.026
30	0.026	0.011	0.009	0.007	0.007	0.008	0.041	0.01	0.007	0.025	0.003	0.008	0.045	0.005	0.022	0.041	0.053	0.049	0.042	0.044

Table 4: Technology class names. Source: Breschi et al. (2002).

1.	Electrical engineering	16.	Chemical Engineering
2.	Audiovisual techn.	17.	Surface techn.
3.	Telecommunication	18.	Materials Processing
4.	Information techn.	19.	Thermal Processes
5.	Semiconductors	20.	Environmental techn.
6.	Optics	21.	Machine tools
7.	Control techn.	22.	Engines
8.	Medical techn.	23.	Mechanical Elements
9.	Organic Chem.	24.	Handling
10.	Polymers	25.	Food Processing
11.	Pharmaceutics	26.	Transport
12.	Biotechn.	27.	Nuclear Engineering
13.	Materials	28.	Space techn.
14.	Food Chem.	29.	Consumer Goods
15.	Basic Materials Chem.	30.	Civil Engineering

7.2 Calculations of maximum spanning trees

Two figures conveniently illustrate the *Kruskal algorithm* of finding the maximum spanning tree. A spanning tree connects all nodes, without unnecessary connections. This means that for a network like in Figure 4 with five nodes ($n = 5$), only four need to be selected to form the spanning tree. Inside the nodes are denoted, (in this case) the number of patents in a specific class. The numbered link between them denotes the strength of the link.

The Kruskal algorithm takes as its first link simply the largest value of the valued links, 0.5 between nodes 1 and 2. The second link is the second largest link, *unless* no new component is added to the tree. It may be that there are more than one tree currently being formed, in which connecting two earlier separete trees is ok (example will follow). If the second link is not valid, the third largest link is used unless condition is not fulfilled. If two links are equally strong (and both valid), it doesn't matter which is chosen. In the figure above, 0.4 is chosen in the second step between nodes 1 and 2. The third link should be 0.3 between nodes 2 and 3. But this line is not valid, since no new node is added. Therefore, we check the fourth largest number, 0.2 between nodes 4 and 5, which is allowed. The fifth largest value is 0.1, which is allowed because two earlier disjoint tree-parts are linked together. The formed maximum spanning tree is shown as Figure 5:

Figure 4: A network with corresponding strength of the links. Source: Constructed by the author with help of the freeware visone, http://www.visone.de/

Figure 5: The finished maximum spanning tree, with emphasized lines showing the chosen links.

It remains to calculate the $WARN$-value for our tree. Since $p_j = P_j / \sum_j P_j$, (3) may be rewritten as

$$WARN_{ir} = \frac{\sum_{j \neq i} q_{ij} \lambda_{ij} \frac{P_j}{\sum_j P_j}}{\sum_{j \neq i} \lambda_{ij} \frac{P_j}{\sum_j P_j}} = \frac{\sum_{j \neq i} q_{ij} \lambda_{ij} P_j}{\sum_{j \neq i} \lambda_{ij} P_j} \tag{14}$$

which highlights the importance of the strength of the nodes, and turned out to be more convenient for the procedure of calculating different $WARN$. Taking the average of the values for the $WARN_{ir}$ belonging to the maximum spanning tree, gives for our example above the value 0.2931. To calculate the values for Swedish regions, an algorithm was written in SAS implementing the outlined Kruskal procedure.

7.3 Regions ranked by Population and Patents per Capita

Table 5: Top 20 regions ranked in terms of total population (1991-1999) and patents per total number of employees. A star indicates that the region does better than is motivated by its population size. Note: The names refer to the regions and not only the cities themselves. Source: Own calculations based on data from EPO and Statistics Sweden.

Regions ranked by population	**Regions ranked by patents per employee**	**Patents per 1,000 employees 1991-1999**
Stockholm	Västerås*	0.7591
Göteborg	Ludvika*	0.6639
Malmö	Malmö	0.5824
Helsingborg	Stockholm	0.5812
Uppsala	Göteborg	0.5314
Linköping	Uppsala	0.4926
Örebro	Gävle*	0.4401
Uddevalla	Karlskoga*	0.3945
Skövde	Helsingborg	0.3834
Västerås	Gnosjö*	0.3759
Norrköping	Fagersta*	0.3518
Kristianstad	Karlstad*	0.3367
Borås	Linköping	0.3290
Karlstad	Åmål*	0.3074
Gävle	Lycksele*	0.2782
Sundsvall	Luleå*	0.2681
Falun	Sundsvall	0.2631
Luleå	Hudiksvall*	0.2630
Jönköping	Karlshamn*	0.2619
Umeå	Eskilstuna*	0.2434