

THE SPATIAL DIMENSION OF KNOWLEDGE SPILLOVERS IN EUROPE : EVIDENCE FROM FIRM PATENTING DATA

by
Bart Verspagen*
and
Wilfred Schoenmakers**

Paper prepared for the AEA Conference on INTELLECTUAL PROPERTY
ECONOMETRICS, 19-20 April 2000, Alicante, Spain

FIRST DRAFT

* Corresponding author, ECIS and MERIT, bart.verspagen@merit.unimaas.nl

** MERIT



ECIS - Eindhoven Center for Innovation Studies
PO Box 513, 5600 MB Eindhoven, the Netherlands



MERIT - Maastricht Economic Research Institute on Innovation and Technology
PO Box 616, 6200 MD Maastricht, the Netherlands

1. Introduction

Technology spillovers are important because they have a large impact on economic growth. In the interpretation of, e.g., Grossman and Helpman (1991), technology spillovers increase the efficiency of the research process within an individual firm. The notion that is behind this is that technological knowledge is a non-rival good, i.e., can be shared without reducing its value. The paradoxical situation is that this characteristic of knowledge, while beneficial at the aggregate level, leads to a lack of incentives to produce knowledge at the micro level. The reason is that firms that have the prospect that the knowledge they develop will be imitated by other firms at lower costs will decide not to invest in research.

One of the institutions that has been developed to remedy this incentive problem is the patent system (see David, 1993, for a discussion of the context of this system). The patent system ‘solves’ the incentive paradox by granting a temporary monopoly to part of the knowledge that has been developed by the patent applicant, while leaving other aspects of this knowledge for public use. For example, a patent prohibits (exact) copying of the knowledge described in it, but it does not prohibit the possibility of building further on this knowledge to develop a new piece of knowledge (which can then on its turn be patented). It is this process of cumulative inventions that Grossman and Helpman use to model the research sector in their model of endogenous economic growth, and without which, in their view, growth ceases in the long run.

This paper follows in a tradition (e.g., Jaffe, 1986, Jaffe, Trajtenberg *et al.*, 1993, Sjöholm, 1996, Verspagen, 1997, Los, 1999) that attempts to use this aspect of the patenting system to measure and explain knowledge spillovers. Various methods have been proposed to implement the measuring aspect of this, such as counting patent citations (Jaffe *et al.*, 1993), or defining technological relatedness in various ways (Jaffe, 1986). The current paper will use patent citations for measuring knowledge spillovers. The pros and cons of this choice, as well as the details of the measurement process will be described in more detail below.

The fact that externalities are related to public aspects of knowledge does not mean that all parties are equally well able to use the spillovers. The two most extensively discussed factors that impact on the efficiency with which knowledge spillovers are received include the absorptive capability of the receiver, and geographical distance. The first of these, absorptive capacity, relates to the notion that the receiving party must have specific competencies to make use of the ‘received’ knowledge. Cohen and Levinthal (1989) show how these competencies are often (broadly) the same as the competencies that are required to actively develop knowledge. In other words, receiving and using spillovers on the one hand, and actively creating knowledge on the other hand, are two processes that can hardly be separated in practical terms.

This notion fits neatly to the idea of using patents in some form as a measure of spillovers. Patents are, in the first place, an outcome of an inventive process. Hence using, for example, patent citations as indicators of spillovers, illustrates nicely how spillovers are interwoven into the inventive process itself. This also links up closely to the specific interpretation of knowledge spillovers offered by Grossman and Helpman (1991), as discussed above. However, this does not imply that the types of knowledge spillovers that have been discussed so far are the only ones that are relevant. For example, Griliches (1979) discusses the concept of ‘rent spillovers’, which are closely linked to knowledge

embodied in traded capital or intermediary goods. This type of spillover will be largely left out of the analysis here.

The geographical dimension of absorbing knowledge spillovers results from the distinction between tacit and codifiable knowledge. Tacit knowledge is embodied in people, and cannot be transferred by other ways than personal interaction. Codifiable knowledge on the other hand can be put down in written instructions, and can thus be studied independently of personal contacts. In this interpretation, the new information and communication technologies are mainly related to codifiable knowledge. Thus, their increased efficiency in transferring information or codifiable knowledge over large distances would not necessarily have a large impact on the transfer of tacit knowledge.

If knowledge is mainly tacit, geographical distance is important in transmitting it. In other words, tacit knowledge travels easily over small distances, but far less easily over longer distances (Caniëls, 1999). Also, the transmission of tacit knowledge is easier in a highly dense network of researchers working on similar topics (Saxenian, 1994, Oerlemans, Meeus *et al.*, 2000). This leads to the hypothesis that, controlling for factors such as technological relatedness, the intensity of spillovers increases with geographical proximity. This issue was first researched using (US) patent citation statistics by Jaffe *et al.* (1993). They found strong evidence for a positive proximity effect on spillovers. Their level of aggregation was the individual organization (firms and universities).

Maurseth and Verspagen (1999) and Maurseth and Verspagen (1999b) tested the proximity effect in the European context, using data on patent citations from the European Patent Office. Their level of aggregation was the region, i.e., patents were assigned to regions (on the basis of the inventor) rather than individual firms. One problem associated with this is the inclusion of intra-firm patents and patent citations in regional patent (citation) counts. Obviously, by definition, intra-firm citations do not indicate spillovers. Rather, they indicate cumulativeness of innovation processes at the firm level. In order to reduce the possible bias associated to this problem Maurseth and Verspagen excluded intra-regional citations (assuming that a large part of these citations in fact occurs at the intra-firm level), which are in fact a large part of the sample. Using only the inter-regional part of the database, they found, as did Jaffe *et al.*, a strong proximity effect on spillovers.

The main reason for Maurseth and Verspagen for not doing the analysis at the firm level, was that the assignment of patents to firms is a costly and lengthy procedure. The European Patent Office (EPO), which is the prime source of the data, does not standardize applicants' addresses, nor does it indicate mother-daughter relationships between applicants. Thus, each individual firm may appear in the EPO database under many different subsidiary names, or even (slightly) different varying names of the same legal entity.

This paper will attempt to extend the Maurseth and Verspagen analysis by drawing upon a database of firm level patenting for 27 large multinational firms. This database was constructed by standardizing applicants' names in the EPO data, and by merging subsidiaries onto the level of 'groups'. Although there are some conceptual problems related to the procedure that was used to perform this task, we are confident that the resulting database provides valuable insights into European patenting at the firm level. The research question that we pose is then whether or not the strong proximity effect on patent citations (as an indicator of technology spillovers) found by Maurseth and Verspagen, also holds in a micro context of between-firms citations.

The rest of this paper is structured as follows. We start, in Section 2, with a discussion of our data. This will introduce the way in which patents were assigned to firms. Also, we will discuss the problems that are still related to our procedure. Section 2 will also provide a brief background discussion of the advantages and disadvantages related to using patent statistics and patent citations as indicators of technology spillovers. Section 3 will provide a descriptive look at the data, and explain some of the choices that were made in terms of defining the sample period. Because geographical distance is the major theme of our analysis, we will provide some insight into the spread of patenting activity over European space. Also, some characteristics of the firms that are in the sample will be discussed. Section 4 presents the model that we will estimate, and discusses the results. Finally, Section 5 will summarize the results and provide some conclusions.

2. Data issues¹

This paper follows in a tradition that uses patent citations as an indicator of technology spillovers. As has been noted before (Maurseth and Verspagen, 1999), this indicator is far from perfect. In the first place, all disadvantages that apply to patent statistics as indicators of technology or innovation, apply to patent citations as well. These disadvantages have been illustrated well in papers by, e.g., Griliches (1990) and Basberg (1987). Nevertheless, most authors surveying these issues tend to conclude that patent statistics can be useful indicators.

Specifically regarding citations, in the European patent system, which is our primary data source, patent citations are mostly added by the patent examiners rather than the inventors or applicants. Hence, the fact that a certain patent cites another patent does not imply that the inventor was actually aware of the cited patent at the time of the inventive process. Moreover, the citations are mostly added for legal reasons, i.e., to indicate which parts of the described knowledge can be claimed by the applicants, and which parts have been claimed earlier.

Nevertheless, it seems reasonable to take patent citations as an indicator of technology spillovers, for at least two reasons. The first is that, especially in the context of large multinational firms (of which we take a sample in this paper), one may safely assume that inventors are aware of what competitors' R&D labs have patented. Many of these large firms have special departments that search patent databases and other technical literature, and disseminate relevant information within the own R&D department. The second reason is that patent citations, even when added by patent examiners, reveal technological relatedness at the very micro-level of the individual invention. The more related two inventions are, the higher the spillovers between them can be expected to be.

Our primary data source is the EPO database on patent applications. We select all patent applications², whether they are granted, have been rejected (or withdrawn), or are still under review. The EPO database we use contains data up to the beginning of 1998, including approximately 935,000 patent applications. Unfortunately, for the purpose of identifying within-firm patent applications, we cannot rely upon the information that EPO supplies in the "applicant name" field. In that field, one may find personal names or names of firms or organizations. In the case of firms, however, it may be the name

¹ The procedure of constructing group patent data described in this section was carried out solely by Wilfred Schoenmakers.

² We will use the term 'patents' loosely, i.e., also when we refer to patent applications.

of an independent firm, or the name of some form of a larger conglomerate or holding firm. The EPO database contains approximately 180,000 unique names in the “applicants” field.

Our sample of firms is limited to large multinational firms that appeared on the Fortune 500 list in 1997, supplemented by a few large firms from the Fortune lists in earlier years. Of these, we select a subsample of firms active in either chemicals (excluding pharmaceuticals), computers or electronics.³ For these firms, we make use of the Dun & Bradstreet Linkages database to construct a list of the subsidiaries. The Dun & Bradstreet Linkages database includes only full, i.e., one hundred percent, subsidiaries. We refer to this list as the “group”. The version of the Dun & Bradstreet Linkages database we used is from late 1998, and represents thus the mother-daughter relationships at that point in time. Of course these connections have not always been like they were in 1998. Thus, when we use the mother-daughter ties of 1998 to construct patent data for groups during the period 1983-1995 (see below), the resulting data do not reflect group patent data at the time of the patent applications involved, but rather at some ex post period in time (1998).

Obviously, this is a sub-optimal procedure. An obviously better procedure would be to use yearly data from the Dun & Bradstreet Linkages database, and to construct for every year under investigation a groups database as we did for 1998. Unfortunately, this involves too much work to be feasible in terms of the resources available for this project. There are, however, at least two reasons why we feel that the procedure used is acceptable for the present purposes. First, most multinational companies apply for the bulk of their patents under the multinational company name. In this area the changes are less pressing, i.e., mergers occur more often on a lower level. Second, the main aim of using group level data is to exclude intra-firm (i.e., intra-group) patents. Although one cannot rule out mistakes on this part that result from the present procedure, it seems reasonable to expect that the number of such mistakes is relatively small. Such a mistake would result if a firm would patent during the period under consideration through a certain subsidiary, and sell off this subsidiary before 1998 (our point of reference for linkages). Given the tendency to patent through group headquarters, one would expect such patents to be a relatively small number.

A further practical problem results from the fact that there is not a one-to-one correspondence between the subsidiary names in the Dun & Bradstreet Linkages database and the names in the “applicant” field of the EPO database. We made a pre-selection from the EPO database by searching for different parts of the names found in the Linkages data. The results from this pre-selection were then compared, usually on a one-to-one basis, to the group list from Linkages.

Some names that were found in the pre-selection from the EPO database, could not be identified using the D&B database. In many of these cases, the name found in the EPO database was partly identical to the name of the (subsidiary) firm we were looking for. This may, for example, happen if the applicant name is the name of a plant, rather than the legal entity it belongs to. In order to be able to learn more about these firms, we constructed a table with the applicant names and addresses from the EPO database. In this way we could compare not only the applicant’s name but also the address to the data found in Linkages. If the applicant was not found in the D&B database, but the applicant’s name was almost identical and the address was identical to other daughter firms of the same multinational, then we included the name in the group list.

A final note refers to the case when companies merge, and the original company name under which they applied for a patent might be lost. This could mean that we would not find these patents, although they belong to the multinational firm under investigation. Therefore we also looked at different parts of the company name, thereby eliminating as much as possible this bias.

3. A descriptive look at the data

We have data for 27 multinational firms with headquarters in Europe (EU), the USA or Japan. These 27 firms are a selection of a larger group of 40 firms for which we presently have patent data. The list of 27 firms resulted from leaving our firm with less than 1,000 patent applications since 1979. The main variable of our interest is the number of patent citations between (and within) these firms. Patent citations take place over a long time span. Figure 1 shows the average number of citations to patents of a certain year as a function of the ‘citation lag’. The latter is defined as the difference between the priority year⁴ of the citing patent and the cited patent (one corresponds two a citation within the same year, two for a citation a year later, etc.).

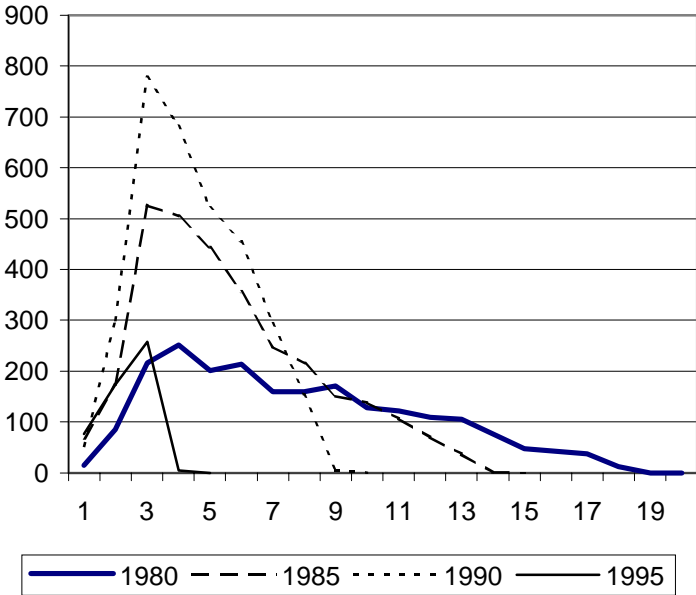


Figure 1. Citation profiles for patent applications in the sample of 27 firms

The figure shows that the number of citations peaks at 2 years (for 1985 and onwards) or 3 years (1980) after the priority date of the original application. The two earliest years in the figure show that citations may go on for a long while, even up to almost 20 years. The sudden drop in the lines for 1990 (lags 9 and 10) and 1995 (lags 4 and 5) is caused by the truncation of the sample towards the end of the period. Many patent applications for the most recent years (generally with priority dates up to one year before the application date) do not show ‘full citation potential’ yet because not all search reports (in which citations appear) have been completed yet.

³ The process of constructing the database described in this section is still under way. Other sectors are being added to the database, but these data were not yet available at the time of writing this paper.

⁴ Wherever we use dates in the analysis of patents or patent citations, these dates are priority dates.

Figure 1 shows how two potential biases may result from a non-careful choice of citation dates to be used. To see this, consider the case where one would simply count all citations present in the figure. The first problem this creates is related to the fact that, e.g., for a given citation lag, patents with priority dates in 1990 get cited more often than patents with priority dates in 1980. As will be shown below in more detail, this has to do with the fact that the number of patent applications at the European Patent Office shows a positive trend with regard to the priority year. Especially the early years (EPO started in 1979) show only few applications. The second problem is related to the fact that ‘earlier’ patent applications have a longer citation history. In case of Figure 1, patents with priority dates in 1980 have a citation history of almost 20 years, while patents with priority dates in 1995 have a history of only two years.

Thus, one would ideally like to consider a sample of patent citations in which both the maximum citation lag is fixed, and the number of patent applications is roughly constant over priority years. In this way, the two biases explained above can be minimized. Figure 2 shows how this can be achieved for the current sample.

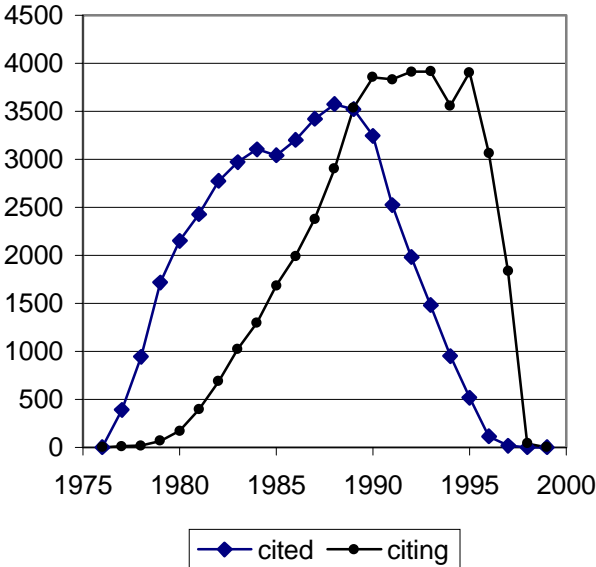
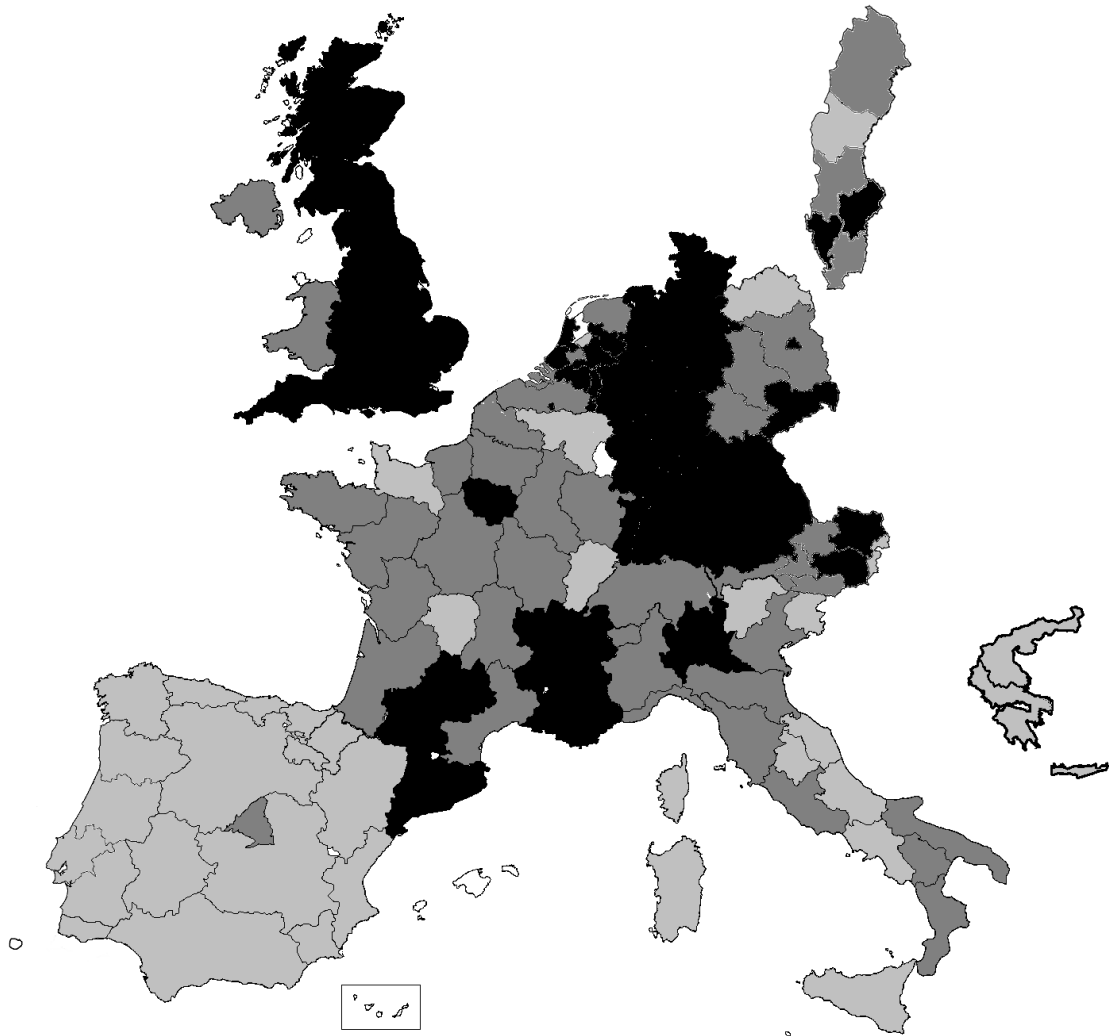


Figure 2. The number of intra- and interfirm patent citations in the sample of 27 firms

The rise in the number of cited patents is caused by the fact that the European Patent Office was set up 1979 (because patents are dated by priority date we do get, however, some observations before 1979). Patent applications after (roughly) 1990 do not get cited much, because they have not yet reached there full citation potential yet. Figure 1 shows that, e.g., for 1990 this is caused by a cut-off at lags roughly larger than 7, while for later years this cut-off point moves to the left (3 for 1995). For the period 1989 - 1995, the number of total citing patent applications is roughly constant. A citation lag of about 7 (in Figure 1) is about the maximum that can be applied (backward) to this period. This suggests using the period 1983 - 1989 as the period for cited patents, with the maximum citation lag equal to 7. Thus, we apply a moving citation window of six years to the period 1983 - 1989: cited patents in 1985 may include citations up to 1989 (including), cited patents in 1989 may include citations up to 1995 (including).



Map 1. Regional distribution of patents of the 27 firms in the sample, darker shades indicate more patents, 1983-1995

In order to capture the geographical dimension of the data, we specify 109 European regions, largely based on the NUTS regional classification applied by Eurostat. We use the same regional breakdown as in Maurseth and Verspagen (1999a), which includes both NUTS 1- and 2-digit regions. In the case of Germany and Austria, small urban regions (Bremen, Hamburg, and Vienna) were merged with neighbouring or surrounding regions to avoid too small geographical entities. Denmark, Ireland, Finland and Luxembourg are excluded because we do not have patents broken down by region for these countries. The assignment of patents to regions is one by using the postal code of the inventor address. The exact procedure used to do this is similar as in Caniëls (1999).

A patent usually has more than one inventor. The inventors are (legally) different from the applicants. Applicants are usually firms or other organizations, inventors are usually persons. In this paper, we select patents by their applicants (i.e., the multinational firms groups), but assign patents to

geographical regions by their inventors. The reason for this is that some firms always apply for a patent from their headquarters, even when a patent was invented in a different region. By nature of the research question (which basically deals with spillovers between inventors), we are interested in the geographical location of the inventor rather than the applicant. In order to cope with multiple inventors of a single patent, we measure citations at the level of individual inventors. For example, when patent A with 3 inventors in 3 different regions (A_1 , A_2 and A_3) cites patent B with 2 inventors in 2 different regions (B_1 and B_2), we count 6 (3×2) citations, i.e., all the permutations of regional combinations (A_1 - B_1 , A_1 - B_2 , A_2 - B_1 , etc.). Naturally, this implies that the total number of 'inventions' we count for a firm or a region is larger than the total number of patents for that firm or region. We refer to this method of counting as 'full-counting'.

Map 1 gives an indication of how many patents ('full-counted') are applied for from the various regions. The sample is divided into three groups (26 high-activity, 36 medium-activity and 37 low-activity regions). The five regions with most patents (all more than 10,000 patents over the complete sample period, i.e., 1983-1995) are four German regions and one Dutch region. There are 35 regions (i.e., almost all of the low-activity regions in the map) with less than 10 patents. Of these, there are 13 regions without any patents.

In general terms, the spatial pattern on Map 1 corresponds relatively well to the one found by Maurseth and Verspagen (1999), who use all patents instead of just a sample of a limited number of firms. The regions with high patenting activity in our firm sample are concentrated in the band United Kingdom - Netherlands - Germany, as well as a limited number of more isolated clusters outside this band. Most of these clusters are fairly well-known centers of technological activity, such as Gothenburg-Stockholm in Sweden, Paris and South France, North Italy and the region around Vienna. The Southern European countries show low patenting density, although the somewhat large concentration in patents in South Italy and the Spanish regions around Madrid (medium) and Barcelona (high) are somewhat surprising. In the Madrid region, three firms (Monsanto, BASF and General Electric) are responsible for the activity, in South Italy it is solely General Electric that is responsible for patenting. In the region around Barcelona, it are Henkel and Hewlett Packard that are responsible for the high concentration.

The firm sample that we use includes 27 firms, 11 of these are active in chemicals, the other 16 in electronics (including office machinery). Almost half (13) of these firms have headquarters in Europe, but five firms have headquarters in North-America, and nine firms are Japanese-based. Obviously, European-based firms may be expected to have larger numbers of inventors located in Europe (see, e.g., Patel and Pavitt, 1991, 1994). This is indeed a tendency that we observe in the sample. Nevertheless, all of the firms in the sample did have 'substantial' numbers of patents (full-counted) originating from Europe. Figure 3 gives an indication of the differences in number of patents between the firms in the sample.

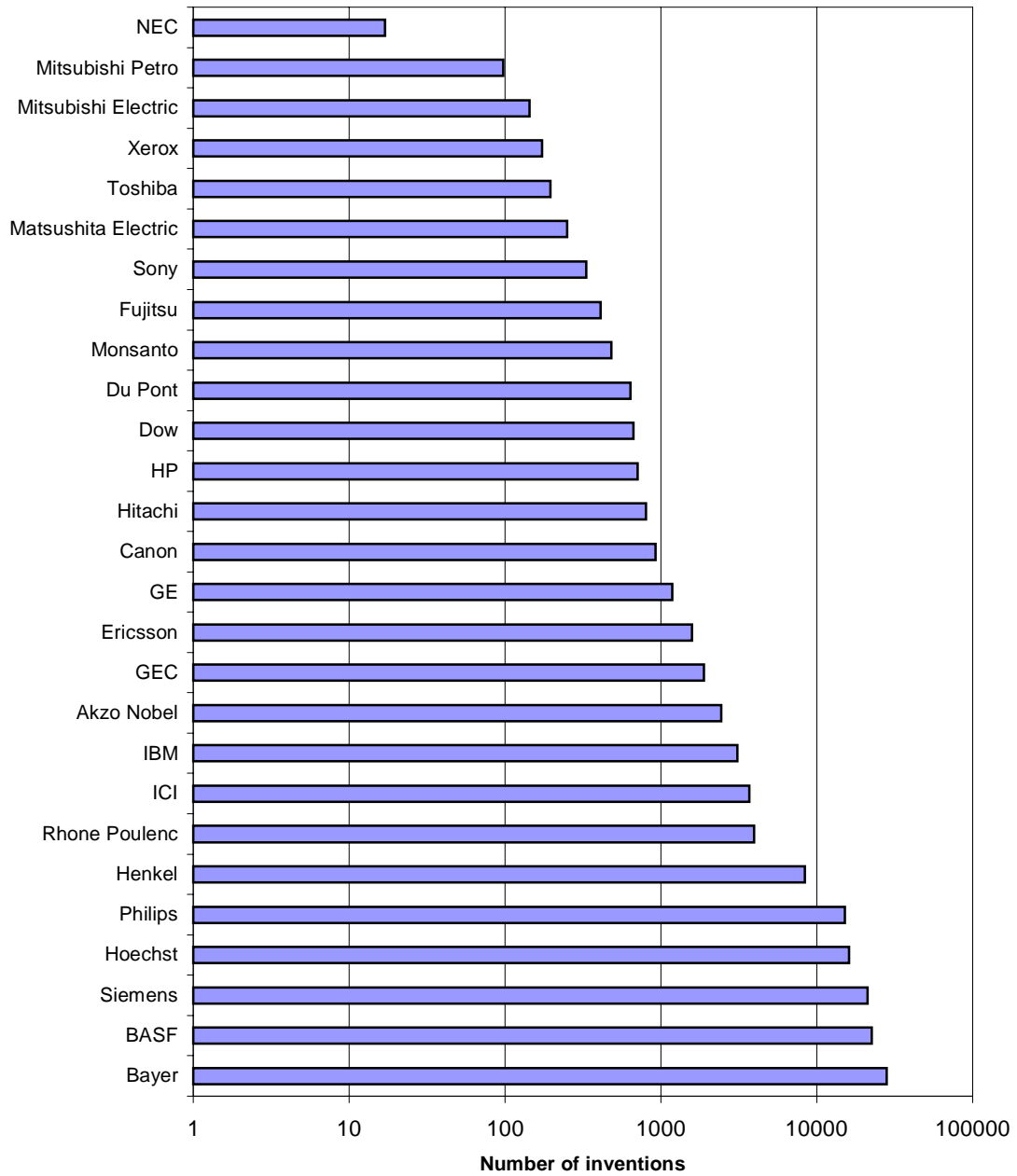


Figure 3. Number of patents ('full-counted') in the 109 European regions, per firm, 1983-1995

4. Model setup and estimation

We first set up a list of all patents applied for by firms in the sample in the period 1983 - 1995. For this list, all citations between patents on the list are identified. Then the citations with citation lags longer than six years are removed. The resulting list of citations is transformed into a 27x27 firm-by-firm citation matrix, with rows indicating spillovers received (citing patents) and columns indicated spillovers generated (cited patents). This matrix will be referred to as the ‘data matrix’. Our main dependent variable consists of a stacked version of this matrix. Because we will leave out the intra-firm citations (the diagonal of the matrix), we have $27*27-27 = 702$ observations.

The model that will be estimated can be formulated in the following general way:

$$C_{ij} = f(gd_{ij}, td_{ij}, P_i, P_j, INTRAC, INTRAE, DR_i, DT_j, c). \quad (1)$$

In this equation, C_{ij} is the number of citations by firm i to patents applied for by firm j , gd is geographical distance, td is technological ‘distance’, P is the number of patent applications (during the period for which citations are considered), $INTRAC$ and $INTRAE$ are two dummies for intra-sectoral citations in chemicals and electricals (computers and electronics), respectively, DR and DT are firm dummies, and c is a constant. The variables are constructed/defined in the following way:

C_{ij} = number of citations constructed using the citation window during the period 1983-1989 as described above.

gd_{ij} = geographical distance based on distances between regions. The starting point is the set of NUTS regions introduced above. Distances between these regions are measured in a stylized way, i.e., by counting the (minimum) number of borders one has to cross to reach region p from q . For example, if p and q are border regions, the distance will be one, if there is one region between p and q , the distance will be two. In this calculation, sea areas are generally counted as one region. This method is the same as applied in Maurseth and Verspagen (1999), and is often used in the geographical literature (Hagget, Cliff *et al.*, 1977). The distance between regions is used together with patent shares of firms per region to calculate distances between firms (the variable gd). This is done as follows:

$$gd_{ij} = \sum_{k=1}^{109} \sum_{l=1}^{109} d_{kl} s_{ik} s_{jl}, \text{ where the indices } k \text{ and } l \text{ denote regions, } d \text{ is the regional distance as}$$

explained above, and s_{ik} is the share of patents of firm i located in region k . Note that when a firm’s patents are located in more than one region, the diagonal of the matrix with elements gd will not be zero, i.e., firms have a positive distance to themselves. Also, this matrix is symmetric, i.e., $gd_{ij} = gd_{ji}$.

td_{ij} = technological ‘distance’, calculated in two different ways. Both measures are based on the distribution of firms’ patents over 4-digit IPC classes. The general idea behind both measures is that when firm patent ‘portfolios’ (in terms of the IPC classes) are similar, the technological distance between them is low (Jaffe, 1986). Both measures start by defining σ_{iz} as the share of patents of firm i

in IPC class z . The first measure of technological distance is defined as $td_{ij}^1 = \sum_{z=1}^{545} (\sigma_{iz} - \sigma_{jz})^2$. This

measure applies equal weights to all IPC classes. The second measure of technological distance

assumes that large differences in patent portfolios are more important when they apply to ‘important’ or large IPC classes. This is taken into account by weighting the IPC classes by their rank in the patent portfolio of firm i , as follows: $td_{ij}^2 = \sum_{z=1}^{545} R_{iz}^\beta (\sigma_{iz} - \sigma_{jz})^2$. In this equation, R_{iz} is the rank of IPC class z in the portfolio of firm i . The exponent β is used to scale the steepness of the weighting scheme, and is set to 0.5 in all calculations in this paper.

Instead of entering both P_i and P_j in the equation separately, we choose to enter them jointly in the form of $\ln(P_i P_j)$, denoted by lpp . This is similar to the procedure used in Maurseth and Verspagen (1999).

The definitions or measurement of other variables are straightforward.

Our expectations on the signs of the coefficients to be estimated are as follows. For geographical and technological distance, a negative sign is expected. In both dimensions, the theory says that larger distance (higher proximity) leads to less (more) spillovers. Note that we include the technological distance variable mainly to control for the effect that similar (technological) activities tend to cluster in space. Naturally, such clustering may be related to the existence of spillovers. Hence it would be desirable to estimate a more elaborate model. However, in order to control for the possibility that citations are clustered in space because inventions are clustered in space due to some exogenous factor (like the spatial concentration of public research), we include the technological relatedness variable.

Because patenting is a pre-requisite for patent citations (and hence knowledge spillovers) to occur, a positive sign is expected on the variable lpp . Moreover, Maurseth and Verspagen (1999b) expect an elasticity of around one for this variable. However, as will be argued in more detail below, that expectation may in fact be related to the regional aggregation level they apply, and, hence, may not apply to the present case of micro-level firm data. The intra-sectoral dummies to a certain extent measure the same phenomenon as the technology distance variables. Within sectors, one would expect the technological distance between firms to be smaller than between sectors. We include the intra-sectoral dummies mainly because the way in which we can measure technological distance is relatively crude. In addition to a negative sign on technological distance, we would therefore expect a positive sign on the intra-sectoral dummies. We do not have specific expectations on the sign of the two types of firm dummies, and will hence not document the estimation results for these variables. These dummies are included to control for unexplained (but fixed) effects related to firm strategies.

Because the data are count data and not normally distributed, OLS is not an adequate technique. We will therefore apply four different sets of estimations, each providing one alternative to the common OLS. The first set of estimations consists of a probit model. For this, we have to re-code the dependent variable as follows:

$$\tilde{C}_{ij} = \begin{cases} 0 & \text{if } C_{ij} = 0 \\ 1 & \text{otherwise} \end{cases}$$

By its nature, the probit model thus attempts to predict the probability of any citations occurring between two firms, rather than predicting the number of citations. Other models that will be estimated below will attempt to do the latter. The independent variables are entered in a linear way in the probit specification. The results for the probit specification are given in Table 1.

The first two equations do not include any dummy variables. The specifications of these two equations are similar, except for the fact that they use different measures for technological distance (one weighted, the other not). Note that in equation (1), all of the independent variables are ‘symmetric’ between firm pairs, while the dependent variable is not. This is obviously a problem, that can only be solved by including a-symmetric variables as explanatory variables. The only a-symmetric variables we have are dummy variables, or the weighted technological distance measure (td^2), which enter the equations (2)-(10) to various degrees.

Equation (1) and (2) in Table 1 both indicate that geographical and technological proximity have a positive effect on the probability that spillovers will take place. Both variables are significant at the 1% level. The number of patents (lpp) is not significant, but has the expected sign (positive). The value of the coefficient on lpp is significantly lower than what was found in a similar probit model in Maurseth and Verspagen (1999b). We explain this difference by the fact that the number of citations we include is lower than in Maurseth and Verspagen (due to firm selection and cut-off of citations at a fixed lag), and by the fact that we use ‘full-counting’, as explained above. Consequently, we do not attach great value to the exact value of the coefficient on lpp , as long as it is positive.

The next two equations (3 and 4) include the intra-electronics dummy. Note that within the chemicals sector, all firms cite each other, so the partition of the full data-matrix related to intra-chemicals citations is filled with ones only. Hence we cannot include the intra-chemicals dummy in the estimations. The intra-electronics dummy has the expected positive sign, and is significant at the 1% level. Moreover, including this variable increases (in fact, almost doubles) the absolute value of the coefficient on geographical distance (compare to equations 1 and 2). Also, the lpp variable is now significant with the expected positive sign. Including the intra-electronics dummy also raises the goodness of fit as indicated by the pseudo- R^2 statistic.

Next, we include firm dummies, first for the spillover ‘receiving’ (i.e., citing) firm (equations 5 and 6). This again raises the goodness of fit significantly. It leaves the signs and significance of the other variables unaffected. Including dummies for the spillover ‘generating’ (i.e., cited) firm, but leaving out dummies for the receiving firm, is done in equations (7) and (8). This largely has the same effect as in (5) and (6), although the pseudo- R^2 is now somewhat higher, and also the exact values of the other coefficients differ somewhat relative to equations (5) and (6).

Thus, so far, the probit estimation results provide strong support for the positive effect of proximity on spillovers. In all equations estimated so far, geographical distance has a significantly negative effect on the probability of a spillover occurring. This result breaks down in the last two columns of Table 1. Here we include both dummies for the receiving firm and the generating firm. This means that we include 54 firm dummies in total. In fact, this large number of dummies may imply that some of the effects of geographical distance or technological distance are captured by the dummies, rather than by the distance variables. Our measure of distance (technological or geographical) is admittedly crude, and there is significant correlation between these measures and the firm dummies. In other words, some firms tend to be remote (in a geographical or technological way) from all other firms. Hence, we interpret the results in equations (9) and (10) in Table 1 with some care. In any case, equations (9) and (10) confirm the negative effect of technological distance, but not of geographical distance. As implied

Table 1. Results from the probit regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Constant	2.033***	1.997***	1.074**	1.046**						
<i>gd</i>	-0.178***	-0.180***	-0.302***	-0.303***	-0.251***	-0.257***	-0.304***	-0.304***	0.228	0.156
<i>td</i> ¹	-7.749***		-7.925***		-11.635***		-12.234***		-269.057***	
<i>td</i> ²		-0.318***		-0.333***		-0.506***		-0.495***		-6.497***
<i>lpp</i>	0.024	0.022	0.101***	0.099***	0.131***	0.129***	0.134***	0.131***	0.305	-0.644**
<i>INTRA</i> <i>E</i>			1.915***	1.926***	2.471***	2.498***	2.375***	2.410***	-3.154	5.256***
<i>DR</i> present	No	No	No	No	Yes	Yes	No	No	Yes	Yes
<i>DT</i> present	No	No	No	No	No	No	Yes	Yes	Yes	Yes
<i>n</i>	702	702	702	702	702	702	702	702	702	702
Pseudo <i>R</i> ²	0.14	0.13	0.30	0.29	0.41	0.41	0.46	0.44	0.77	0.74
<i>logl</i>	-308.333	-313.189	-250.69	-253.743	-210.199	-213.058	-195.484	-199.345	-83.499	-94.646

One, two and three stars denote significance in a 2-tailed *z*-test at the 10, 5 and 1% level, respectively.

Table 2. Results from the Poisson regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Constant	4.809***	4.798***	1.287***	1.237***						
<i>gd</i>	-0.147***	-0.150***	-0.175***	-0.178***	-0.107***	-0.109***	-0.130***	-0.131***	-0.058***	-0.058***
<i>td</i> ¹	-12.251***		-6.337***		-11.260***		-8.417***		-16.398***	
<i>td</i> ²		-0.514***		-0.263***		-0.481***		-0.351***		-0.706***
<i>lpp</i>	-0.007***	-0.009***	0.098***	0.098***	0.099***	0.099***	0.098***	0.097***	-0.204***	-0.227***
<i>INTRAC</i>			1.304***	1.351***	0.836***	0.886***	1.264***	1.330***	0.294***	0.795***
<i>INTRA</i> <i>E</i>			2.502***	2.541***	2.303***	2.365***	2.242***	2.290***	2.094***	1.740***
<i>DR</i> present	No	No	No	No	Yes	Yes	No	No	Yes	No
<i>DT</i> present	No	No	No	No	No	No	Yes	Yes	Yes	Yes
<i>n</i>	702	702	702	702	702	702	702	702	702	702
Pseudo <i>R</i> ²	0.22	0.20	0.53	0.53	0.66	0.67	0.68	0.67	0.85	0.85
<i>logl</i>	-15237.871	-15563.812	-9039.273	-9100.770	-6294.109	-6354.438	-6258.590	-6314.699	-2896.457	-2942.586

One, two and three stars denote significance in a 2-tailed *z*-test at the 10, 5 and 1% level, respectively.

by our discussion of this specification, we do not, however, take this as a final rejection of the proximity effect. It does point out, however, that further estimations may be useful.

We therefore undertake estimations using other methods. The next set of estimations we perform uses a Poisson regression model. This model (Greene (1995)) estimates the average number of citations between two firms as a loglinear function of the independent variables. To be more precise, the (standard) assumption is the following: $\ln \lambda = f^{lin}$, where f^{lin} is a linear variant of f in equation (1) above, and λ is the arrival rate. The Poisson model has been applied to the case of an innovation production function (specifying the relationship between patents as a measure of innovative output, and R&D as a measure of input) by various authors, including Hausman, Hall *et al.* (1984), Crepon and Duguet (1997), Crepon and Duguet (1997), and Cincera (1997). The results from the Poisson regressions are documented in Table 2.

The specification (inclusion of variables) of equations (1) - (10) in Table 2 is similar to that in Table 1. In the simplest form (equations 1 and 2), the geographical and technological distance variables have the expected negative sign and are significant. However, the lpp variable is negative and significantly so. Despite the fact that the absolute value of the coefficient is small, this is an unexpected result. The next two equations, which include intra-sectoral dummies (this time both for chemicals and electronics) provides positive (and significant) coefficients for lpp . Moreover, the other variables also have the expected signs, and are significant, including the intra-sectoral dummies. The fact that the intra-electronics dummy is larger than the intra-chemicals dummy may result from the fact that we have more electronics firms in our sample. The pseudo- R^2 values for these two equations are fairly high.

Including firm dummies in equations (5) - (8) does not change the results drastically. The coefficients for geographical distance and technological distance always remain negative and significant. Also, the value for lpp remains stable at a positive level close to 0.1 and significant. Finally, the intra-sectoral dummies remain positive. In line with Table 1, the impact of geographical distance on citations drops somewhat when both types of firm dummies are included at the same time (equations 9 and 10). However, in the Poisson specification, the geographical variable remains and significant and negative. For all specifications with firm dummies, the pseudo- R^2 values are high. Concluding, one may say that the Poisson specification is relatively insensitive to the inclusion of firm specific fixed effects, and confirms the proximity effect.

As is well-known, the Poisson model assumes that the mean of the dependent variable is equal to its variance. In many cases, this is an undesirable assumption that is refuted by the data, i.e., the data show so-called overdispersion. In this case, the negative binomial model is more adequate. This model is a close cousin of the Poisson model, and can be derived by assuming cross-firm heterogeneity in the arrival rate. The negative binomial model involves estimating an additional parameter, denoted by α , that enters into the implicitly estimated variance of the dependent variable. The negative binomial model can be tested against the simpler Poisson form using a significance test of this parameter α . The results are documented in Table 3.

The first notable result is that the negative binomial specification is preferred to the Poisson specification in all columns. Also, the pseudo- R^2 values in Table 3 are quite a lot lower than in Table

2. The reason for this lies in the relative importance of cross-firm heterogeneity, which is captured by the α -parameter. The pseudo- R^2 statistic measures increases in the log likelihood relative to a model including the α and a constant only.

Otherwise, the negative binomial results in Table 3 are similar to the Poisson results in Table 2. Hence, including cross-firm heterogeneity seems to be important, but does not change the main conclusions. There is still support for the proximity effect on spillovers in all 10 specifications in Table 3.

Finally, we implement a Tobit model for censored dependent variables. In our interpretation, the variable C_{ij} is censored at zero. Again, a linear specification for f is used in this case. The results are documented in Table 4.

The pseudo- R^2 values in Table 4 are relatively low, as compared to Tables 2 and 3. Because of the different model specification, the value of the coefficients is naturally different, but overall the signs and significance levels are similar as the ones obtained before. Thus, geographical distance and technological distance always have a positive sign. The lpp variable is always positive, except when no dummies at all are included (equations 1 and 2). Equations (9) and (10) show results that are somewhat out of line with those in equations (1) - (8). When including both types of firm dummies, the geographical distance variable loses significance, and the intra-chemicals dummy becomes negative (and significant). Given the rather strong evidence in the other columns (and tables), we conclude that these anomalous results are probably due to the fact that too many dummies are included in equations (9) and (10), as argued already above.

Table 3. Results from the negative binomial regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
<i>Constant</i>	5.026***	4.980***	1.433***	1.405***						
<i>gd</i>	-0.183***	-0.189***	-0.310***	-0.312***	-0.188***	-0.190***	-0.212***	-0.214***	-0.072**	-0.070**
<i>td¹</i>	-9.483***		-6.671***		-9.502***		-8.162***		-23.380***	
<i>td²</i>		-0.386***		-0.278***		-0.404***		-0.333***		-1.011***
<i>lpp</i>	-0.029	-0.028	0.118***	0.116***	0.117***	0.116***	0.130***	0.128***	0.018	0.018
<i>INTRAC</i>			1.129***	1.178***	1.009***	1.068***	1.356***	1.431***	0.662**	0.840***
<i>INTRAЕ</i>			2.682***	2.713***	2.769***	2.814***	2.666***	2.715***	1.389***	1.423***
α^+	2.487***	2.531***	1.274***	1.288***	0.885***	0.899***	0.819***	0.832***	0.269***	0.285***
<i>DR present</i>	No	No	No	No	Yes	Yes	No	No	Yes	Yes
<i>DT present</i>	No	No	No	No	No	No	Yes	Yes	Yes	Yes
<i>n</i>	702	702	702	702	702	702	702	702	702	702
Pseudo R^2	0.02	0.02	0.10	0.10	0.14	0.14	0.15	0.15	0.25	0.24
<i>logl</i>	-2661.042	-2667.453	-2441.638	-2444.927	-2328.405	-2332.635	-2301.060	-2305.723	-2042.807	-2055.713

One, two and three stars denote significance in a 2-tailed z -test at the 10, 5 and 1% level, respectively.

+ Significance based on LR test vs Poisson model, high significance points to acceptance negative binomial model.

Table 4. Results from the Tobit regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
<i>Constant</i>	76.175***	75.670***	2.632	-3.670						
<i>gd</i>	-6.524***	-6.640***	-8.510***	-8.579***	-5.392***	-5.431***	-6.051***	-6.114***	-0.801	-0.839
<i>td¹</i>	-323.064***		-201.152***		-349.350***		-289.447***		-936.661***	
<i>td²</i>		-13.559***		-8.411***		-15.407***		-11.900***		-41.699***
<i>lpp</i>	-0.185	-0.244	3.263***	3.229***	3.153***	3.132***	3.273***	3.183***	1.259	3.253***
<i>INTRAC</i>			17.137***	18.669**	5.312	6.393	11.649**	14.323***	-34.662***	-15.209***
<i>INTRAЕ</i>			64.169***	65.201***	59.192***	60.451***	61.901***	63.460***	26.910***	12.565**
<i>DR present</i>	No	No	No	No	Yes	Yes	No	No	Yes	Yes
<i>DT present</i>	No	No	No	No	No	No	Yes	Yes	Yes	Yes
<i>n</i>	702	702	702	702	702	702	702	702	702	702
Pseudo R^2	0.02	0.02	0.07	0.07	0.09	0.09	0.10	0.10	0.17	0.17
<i>logl</i>	-3033.245	-3039.295	-2898.151	-2900.245	-2819.339	-2820.918	-2805.290	-2808.713	-2579.591	-2580.830

One, two and three stars denote significance in a 2-tailed z -test at the 10, 5 and 1% level, respectively.

5. Conclusions and summary

The main purpose of this paper is to test for the proximity effect on technological spillovers. This effect states that spillovers occur relatively frequently at small geographical distances. We use patent citations between 27 large multinational groups as an indicator of technology spillovers. The patent data refer to European patents. We only consider patent citations between patents that originate from European regions, covering almost the total space of the present EU members. In order to test for the proximity effect, we measure geographical distance at the level of firms, using data on the location of inventive activities of the firms.

Because of the specific nature of our dependent variable (count data, non-normally distributed), we have to rely on non-OLS techniques. Our choice was to implement four different estimation methods: a probit model (re-coding our dependent variable in to zero or positive citations), a Poisson regression model, a negative binomial regression model, and a Tobit model that assumes the dependent variable is left-censored at zero. Besides geographical distance, we generally include ‘technological distance’ between firms, the number of patents of citing and cited firms, two intra-sectoral dummies, and two different sets of firm dummies to capture fixed effects related to firms’ strategies.

The results generally confirm the proximity effect on spillovers. We usually find significantly negative coefficients on the geographical distance variable. The results for the other variables are also in line with our *a priori* expectations. The only exception to this are the equations using either a probit or Tobit model where both types of firm dummy variables are included at the same time. In the Poisson and negative binomial model, the inclusion of the two types of dummies also weakens the proximity effect, but in this case, the effect is still significant. We attribute the probit and Tobit results for this specification to colinearity as a result of too many dummies (54) included at the same time. Hence, we conclude that, overall, our results support the proximity effect on technological spillovers.

This implies that invention is a process that is strongly rooted in space. Building on the existing literature, we attribute this to the tacit nature of technological knowledge, which implies that personal contact is important in transferring spillovers. This phenomenon has important implications for countries or regions that attempt to implement a technological catching-up strategy based on imitation knowledge developed elsewhere. Such a strategy may in fact be stimulated by trying to attract research activities of (foreign) multinational firms.

References

- Basberg, B. L. (1987). “Patents and the Measurement of Technological Change: A Survey of the Literature.” *Research Policy* 16: 131-141.
- Caniëls, M. C. J. (1999). *Regional Growth Differentials*. Maastricht, Universitaire Pers Maastricht.
- Cincera, M. (1997). “Patents, R&D and technological spillovers at the firm level: some evidence from econometric count models for panel data.” *Journal of Applied Econometrics* 12: 265-280.
- Cohen, W. M. and D. A. Levinthal (1989). “Innovation and Learning: The Two Faces of R&D.” *Economic Journal* 99: 569-596.
- Crepon, B. and E. Duguet (1997). “Estimating the innovation function from patent numbers: GMM on count panel data.” *Journal of Applied Econometrics* 12: 243-263.

- Crepon, B. and E. Duguet (1997). "Research and development, competition and innovation pseudo-maximum likelihood and simulated maximum likelihood methods applied to count data models with heterogeneity." *Journal of Econometrics* 79: 355-378.
- David, P. A. (1993). Intellectual Property Institutions and the Panda's Thumb: Patents, Copyrights, and Trade Secrets in Economic Theory and History. *Global Dimensions of Intellectual Property Rights in Science and Technology*. M. B. Wallerstein, M. E. Mogee and R. A. Schoen. Washington, D.C., National Academy Press: 19-64.
- Greene, W. (1995). *Econometric Analysis, 3rd ed.* New Jersey, Prentice Hall.
- Griliches, Z. (1979). "Issues in Assessing the Contribution of Research and Development to Productivity Growth." *The Bell Journal of Economics* 10: 92-116.
- Griliches, Z. (1990). "Patent Statistics as Economic Indicators: A Survey." *Journal of Economic Literature* XXVIII: 1661-1707.
- Grossman, G. M. and E. Helpman (1991). *Innovation and Growth in the Global Economy*. Cambridge MA, MIT Press.
- Hagget, P., A. D. Cliff and A. Frey (1977). *Locational Models*. London, Edward Arnold.
- Hausman, J., B. H. Hall and Z. Griliches (1984). "Econometric Models for Count Data with an Application to the Patents-R&D Relationship." *Econometrica* 52: 909-938.
- Jaffe, A. B. (1986). "Technological opportunity and spillovers of R&D: evidence from firms' patents, profits and market value." *American Economic Review* 76: 984-1001.
- Jaffe, A. B., M. Trajtenberg and R. Henderson (1993). "Geographic localization of knowledge spillovers as evidenced by patent citations." *The Quarterly Journal of Economics* CIIIX: 577-598.
- Los, B. (1999). *The Impact of Research and Development on Economic Growth and Structural Change*. Enschede, Ph.D. thesis University of Twente.
- Maurseth, P.-B. and B. Verspagen (1999a). Europe: One or Several Systems of Innovation? An Analysis based on Patent Citations. *The Economic Challenge to Europe. Adapting to Innovation Based Growth*. J. Fagerberg, P. Guerrieri and B. Verspagen. Aldershot, Edward Elgar.
- Maurseth, P.-B. and B. Verspagen (1999b). Knowledge Spillovers in Europe. A Patent Citation Analysis, Paper presented at the CRENOS conference on "Technological Externalities and Spatial Localisation", September 1999.
- Oerlemans, L. A. G., M. T. H. Meeus and F. W. Boekema (2000). "On the spatial embeddedness of innovation networks: an exploration of the proximity effect." *Journal of Economic and Social Geography*, forthcoming.
- Patel, P. and K. Pavitt (1991). "Large Firms in the production of the world's technology: an important case of non-globalisation." *Journal of international business studies* 22: 1-21.
- Patel, P. and K. Pavitt (1994). "Uneven (and divergent) technological accumulation among advanced countries: evidence and a framework of explanation." *Industrial and Corporate Change* 3: 759-790.
- Saxenian, A. (1994). *Regional Advantage. Culture and Competition in Silicon Valley and Route 128*. Cambridge MA and London, Harvard University Press.
- Sjöholm, F. (1996). "International transfer of knowledge: the role of international trade and geographic proximity." *Weltwirtschaftliches Archiv* 132: 97-115.
- Verspagen, B. (1997). "Estimating International Technology Spillovers Using Technology Flow Matrices." *Weltwirtschaftliches Archiv* 133: 226-248.