7. Europe: one or several systems of innovation? An analysis based on patent citations

Per Botolf Maurseth and Bart Verspagen

1 INTRODUCTION

European integration has traditionally been aimed at reduction of barriers to intra-European trade and factor mobility. This has been achieved by the abolition of tariffs and import duties, by liberalization of capital movements and reduced barriers to foreign direct investments, by legislation facilitating mobility of people across the European Union, and by the abolition of various so-called non-tariff barriers to trade under the 1992 programme.

Has a similar degree of integration been reached in the field of technology and innovation? This is an important question since the extent to which a nation or region can assure access to external technological knowledge may be decisive for its economic performance.

The literature on so-called national systems of innovation (for instance, Lundvall, 1992) focuses on the ways in which knowledge flow takes place. Knowledge flows and the factors which have an impact on them are not easy to quantify, however. In the national systems of innovation approach, various factors related to institutions, culture and history are identified as important. From another perspective, Coe and Helpman (1995) have pointed to trade as an important channel for diffusion of technology embodied in physical goods. Diffusion of disembodied technology, in the form of general or specific knowledge, is also believed to be of importance. Although knowledge flows are to some extent related to trade flows, the general tendency to liberalize trade flows in the European Union does not necessarily imply a proportionate increase of knowledge flows.

Much of the recent literature argues that technology spillovers are to an important extent local (Morgan, 1997; Jaffe, Trajtenberg and Henderson, 1993). The reason for this is that, despite modern communications techniques, due to the tacitness of knowledge, frequent face-to-face contact, or mobility of knowledge workers are still important channels of knowledge spillovers. In
this chapter, we use patent citations as an indicator of disembodied technology flows. Such citations may be interpreted as external knowledge flows or technological spillovers. In the analysis presented here, we will focus on the impact of geography, national borders, economic development and technological specialization on the flow of technological spillovers between a set of European regions.

The rest of this chapter is organized as follows. Section 2 discusses some implications of spillovers in general and localized spillovers in particular for economic development at the country and regional levels. Section 3 describes innovative capability in European regions with use of data on patenting. Section 4 provides a descriptive analysis of regional technological interaction as evidenced by patent citations. In section 5 econometric evidence on the determinants of technological spillovers between European regions is presented. The concluding section 6 summarizes the empirical findings and points out some directions for future research. Data construction and data sources are discussed in Appendix 7.1.

2 ECONOMIC GROWTH, REGIONAL DEVELOPMENT AND TECHNOLOGICAL SPILLOVERS

An important characteristic of technological knowledge is that it can be used without being exhausted. Technology is also cumulative in nature, because it is based on previously gained insights. Furthermore, technological knowledge is seldom (completely) limited to the person or firm that developed it and has the property rights to it. In other words, technological spillovers take place. In the recent formal growth models in the neoclassical tradition, increasing returns through spillovers make endogenous growth possible. Without such spillovers, economic growth either ceases in the long run (see Grossman and Helpman, 1991, Chapter 3), or is 'explained' as a completely exogenous process (as in the old neoclassical model of the 1950s).

When spillovers are industry-specific, specialization in certain industries may result in higher growth than specialization in other industries, and the specialization pattern of a country or region is then likely to have an impact on economic growth. If spillovers are geographically concentrated, knowledge stocks may accumulate in proportion to local industrial activity. Thus increasing returns resulting from spillovers may be bounded within geographical limits. Localized spillovers thereby facilitate clustering of economic activity.\(^1\) To reap the benefits of local spillovers, production is established near to pre-existing production. External effects from neighbouring establishments increase profitability of further establishments.
Europe: one or several systems of innovation?

Alfred Marshall observed early on that knowledge spillovers may play a crucial role for clustering of economic activity. In addition to obvious explanations such as endowments of natural resources, Marshall referred to technological spillovers as one of three possible explanations for clustering of economic activity (Marshall, 1948, p. 271):

When an industry has thus chosen a locality for itself, it is likely to stay there long: so great are the advantages which people following the same skilled trade get from near neighbourhood to one another ... If one man starts a new idea, it is taken up by others and combined with suggestions of their own; and thus it becomes the source of further ideas.

The importance of geography for diffusion of knowledge was also recognized by Raymond Vernon as a basis for his product cycle theory. Vernon showed how localized knowledge and technological opportunities might envisage introduction and production of new products in advanced markets.

Kaldor (1978, p. 143), reflecting on uneven regional development, analysed the role of localized dynamic increasing returns as a result of, among other factors, ‘the opportunities for easy communication of ideas and know-how’. Kaldor, inspired by Allyn Young’s theorizing on technological spillovers as a source of aggregate increasing returns (Young, 1928), hypothesized that regional development was subject to a principle of ‘circular and cumulative causation’ in which regional economic progress (or stagnation) is the seed of further progress (or stagnation). Thus, uneven regional development may be an inherent outcome of decentralized economic processes in the absence of counteracting economic policy. Kaldor pointed out that such processes of cumulative causation made a case for regional policies.

Previous empirical research has established that geography may indeed be important for technology spillovers. Analysing patent citations, Jaffe, Trajtenberg and Henderson (1993) found intra-national citations (national patents citing national patents) and intra-state citations (citations to patents originating in the same state) to occur more often than expected from the distribution of patenting activity, using US patent statistics. Similar results were obtained in Jaffe and Trajtenberg (1996) where it also was found that the geographical concentration of spillovers decreased over time. In the case of Sweden, Sjöholm (1996 and 1997) found citations to patents in neighbouring countries to occur more often than to patents originating in more distant countries. Whether these findings extend to Europe as a whole is one of the major issues to be addressed in this chapter.

There are also factors that can be identified as stimulating the flow of knowledge. The so-called technology gap theory on economic growth and international trade deals with the (international) diffusion of technological
knowledge (Fagerberg, 1994). This theory focuses on how countries ranking low on the productivity ladder may catch up with leading countries. Diffusion of technology facilitates the potential for catch-up, but technological progress on the frontier increases the height of the ladder to climb.\(^\text{5}\) The ability to adapt new technologies depends on institutional infrastructure, education, geography and resources devoted to R&D. The technology gap theory has increased the understanding of critical factors of catch-up with the technological leading countries (for instance, finance, the educational system and politics, see, *inter alia*, Abramovitz, 1985; Fagerberg, 1988; Verspagen, 1991). Fagerberg (1994) concludes a survey on the literature on the catch-up debate with the following: 'Indeed, what the whole literature, from Gerschenkron onwards, suggests is that catching up is very difficult, and that only countries with appropriate economic and institutional characteristics will succeed.'

It thus appears that the absorptive capability of a country or region is crucial for clustering. The cumulative nature of technology and the localness of spillovers bring with them a tendency for clustering, and the extent to which this tendency will be counteracted by wider technology diffusion depends on absorptive capacity. If there are large differences in terms of absorptive capacity, a considerable degree of clustering may arise (depending on whether the peripheral regions have high or low absorptive capacity), whereas if all regions have high absorption capability, spillovers flow easily, and the spread of economic activity will be more even.

An important question, in an increasingly integrated Europe, is the extent to which national systems of innovation are still relevant. Increased integration indicates less importance of national borders. In Europe, economic integration is combined with supranational institution building to support regional development (structural funds), exchange of students, cooperation between universities and R&D laboratories and infrastructure projects. This process raises the question whether a European system of innovation will come to supplement the national systems.

On the other hand, studies of national systems of innovation highlight important path-dependent aspects of such systems. One example is technological spillovers that are somewhat specific in scope (sectorally, geographically, etc.). Reduced importance of national borders and policies may imply that systems of innovation become less national but still geographically concentrated. Thus analyses of the innovative capacity in Europe should incorporate both distinct European aspects (for instance in terms of a European system of innovation) and variety at national and regional levels (national and regional systems).

The system of innovation, and whether it can be characterized as European, national or regional, thus provides a crucial link between localized spillovers (which lead to clustering) and diffusion of technological knowledge (leading
Europe: one or several systems of innovation? 153

to convergence. Our analysis in the next sections will be aimed at answering the question of how knowledge flows in the European innovation system. Can we still observe, despite increased integration since the 1950s, factors that hinder the flow of knowledge through the system? Do we see one truly European system of innovation, in which knowledge spillovers flow between all relevant units (for instance, regions), or do we have instead many isolated innovation systems that only interact marginally with each other?

3 TECHNOLOGICAL COMPETENCIES IN EUROPEAN REGIONS

As is well known from evidence at the country level, there are large differences between European countries in terms of technological competencies (see Chapter 1). In terms of R&D intensity (R&D expenditures as a percentage of GDP), large differences between European countries exist. From the analysis of differences in GDP per capita in the European Union (see Chapter 6 by Cappelen, Fagerberg and Verspagen in this volume), we know that regional differences in GDP per capita are much larger than across countries. Because there is a close correlation between technological competencies and GDP per capita (see Fagerberg, Verspagen and Caniëls, 1997), one might expect that regional differences in terms of technology in the European Union are large.

It is the aim of this section to investigate the extent of these differences. Patent statistics will be used to this end. Such statistics are often used as an indicator of technological strength of a country or region. The fact that patent statistics are output indicators rather than input indicators has some advantages as well as disadvantages. The main advantage is that one is able to circumvent the issue of R&D productivity ('the number of innovations per unit of R&D'), and that patent statistics are available for a wider set of regions and a longer time period than R&D statistics. The main disadvantages lie in the problem that simple patent counts do not take into account differences in the quality of innovations, that many patents do not lead to innovations, and that the propensities to patent may differ between sectors. Despite these differences, patent statistics are widely used to analyse regional differences in innovation in the European Union (for instance, Caniëls, 1996; Paci and Usai, 1997; Verspagen, 1997a and 1997b).

The expectation of a correlation between GDP per capita and patenting between European regions is indeed confirmed by the data. The rank correlation between GDP per capita in 1994 and the share in patent applications at the European Patent Office (EPO) over the period 1979–96 is 0.67.
Figure 7.1 gives an overview of patenting activity in European regions. The map gives four groups (quartiles) of regions, based on the number of patent applications at EPO over 1979–96. All applications are assigned to the region of the home address of the inventor, so we rule out any bias resulting from the fact that patents are often applied for from a location other than the one where the invention was made.\textsuperscript{7} The ranking of the regions by number of patent applications is indicated in the figure.

Germany comes out with the highest activity. All regions in the former West Germany rank in the highest group. Even the eastern part, however, has high values, with all regions except one ranking in either the first or second group. This may partly have to do with the fact that the main office of EPO is located in Munich, but it is unlikely that this has a strong impact. Patents can be filed in any language, and all countries have patent lawyers fully qualified to handle

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure7.1.png}
\caption{Patent applications at EPO 1979–96 (share in total)}
\end{figure}
EPO applications. We thus interpret the German result as confirming the technological leadership of this country in the European context.

The other members of the high patenting activity group are spread out over six countries: the United Kingdom, France, Italy, Austria and Sweden. It is noteworthy that there is a clear amount of clustering in two areas: north Italy combined with southeast France, and England. Also in the Netherlands, patenting activity clusters in two adjacent regions. In Belgium, France, Austria and Sweden, the regions with capital cities rank high.

Note that in general, patenting activity in the south of Europe is lower than in the north, as could be expected on the basis of GDP per capita data. Portugal, Spain and Greece are the only countries without a region ranking in the highest activity quartile, and in Italy the high activity regions are located in the north. In fact, in the set of regions consisting of Portugal, Spain, Greece and south Italy, there are only two regions which rank in the 'intermediate high' quartile (regions around Rome and Barcelona); all other regions rank lower. In Portugal, all regions rank in the lowest quartile.

Figure 7.2 adjusts the raw patent applications data for the size of the region by dividing the number of patents used in Figure 7.1 by the population of the region in 1990. This mainly has the effect of reducing the dominance of the larger countries such as Italy, the UK and Germany, in favour of smaller countries such as Austria and the Netherlands. Austria now ranks almost fully in the highest quartile, with only two regions ranking in the second group. In the Netherlands, four regions rank in the top group, as do two in Sweden. The large clusters of high activity in north Italy and England are reduced in size, although most of the regions in these clusters still rank in the 'intermediate high' group.

The division between north and south still remains clearly visible. Portugal still ranks completely in the lowest group, as does Greece, and the largest part of Spain (three regions in this country rank in 'intermediate low', the rest in 'low'). South Italy also ranks very low. In fact, no region north of the Pyrenees ranks in the 'low' group. 8

The conclusion on the geographical spread of inventive or innovative activity over Europe is thus that there is a fair amount of concentration. This concentration occurs in various dimensions. It is perhaps most visible in the north–south context, where we find the familiar pattern of high innovation activity in the north, and low activity in the south. However, also in the within-country dimension, concentration occurs. Each of the countries in our maps clearly shows some geographical concentration of patenting.

The degree to which patenting is geographically concentrated differs, however, between industries. In order to investigate this dimension, we assigned each of the patent applications to one or more of 22 manufacturing
industries, according to the MERIT concordance table between IPC and ISIC (Verspagen, van Moergastel and Slabbers, 1994). We thus have, for each region, not only the total number of patent applications, but also the spread of these over 22 sectors.

It turns out that for each sector, particularly the so-called high-tech ones, the (rank) correlation between the total amount of patents in the regions and the sector-wise number of patents by region is quite high. It would thus be redundant to repeat the maps shown for total patenting at the sectoral level. Instead, we present, in Table 7.1, a more synthetic measure of concentration, in the form of the Herfindahl index. This index is defined as the sum (over regions) of squares of regional shares of patenting. A high number indicates high concentration.
Table 7.1 presents the results of these calculations for the 22 sectors. The two sectors with the highest geographical concentration are both high-tech sectors, that is, pharmaceuticals and computers. Other high-tech sectors also show high concentration (for instance, aerospace, electronics). The other sectors for which patenting is relatively concentrated are all scale-intensive sectors: ferrous basic metals, chemicals, electrical machinery, refined oil, ships and boats. The high geographical concentration of patenting in these sectors may thus well be the result of a high concentration of economic activity over space, rather than the result of some inherent tendency for technological activity to be clustered.

Table 7.1 Concentration of patenting over European regions for 22 manufacturing sectors; total patenting and high-tech patenting

<table>
<thead>
<tr>
<th>Sector</th>
<th>HF-index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pharmaceuticals</td>
<td>0.075</td>
</tr>
<tr>
<td>Computers and office machines</td>
<td>0.073</td>
</tr>
<tr>
<td>Ferrous basic metals</td>
<td>0.071</td>
</tr>
<tr>
<td>Electrical machinery</td>
<td>0.068</td>
</tr>
<tr>
<td>Aerospace</td>
<td>0.067</td>
</tr>
<tr>
<td>Chemicals</td>
<td>0.064</td>
</tr>
<tr>
<td>Refined oil</td>
<td>0.059</td>
</tr>
<tr>
<td>Ships and boats</td>
<td>0.056</td>
</tr>
<tr>
<td>Electronics</td>
<td>0.055</td>
</tr>
<tr>
<td>Motor vehicles</td>
<td>0.051</td>
</tr>
<tr>
<td>High-tech aggregate</td>
<td>0.049</td>
</tr>
<tr>
<td>Other transport</td>
<td>0.048</td>
</tr>
<tr>
<td>Paper and printing</td>
<td>0.047</td>
</tr>
<tr>
<td>Textiles, apparel, leather</td>
<td>0.046</td>
</tr>
<tr>
<td>Instruments</td>
<td>0.046</td>
</tr>
<tr>
<td>Non-ferrous basic metals</td>
<td>0.045</td>
</tr>
<tr>
<td>All sectors aggregate</td>
<td>0.043</td>
</tr>
<tr>
<td>Machinery</td>
<td>0.042</td>
</tr>
<tr>
<td>Metal products</td>
<td>0.042</td>
</tr>
<tr>
<td>Wood and wooden products</td>
<td>0.041</td>
</tr>
<tr>
<td>Food, drinks and tobacco</td>
<td>0.040</td>
</tr>
<tr>
<td>Glass, stone and clay</td>
<td>0.039</td>
</tr>
<tr>
<td>Rubber and plastic products</td>
<td>0.035</td>
</tr>
<tr>
<td>Other manufacturing</td>
<td>0.034</td>
</tr>
</tbody>
</table>

Source: Calculations based on EPO data.
4 TECHNOLOGY SPILLOVERS BETWEEN EUROPEAN REGIONS

From the point of view of a European innovation system, what matters is not only the distribution of activities over regions, but also the way in which regions 'interact' with respect to technology. As pointed out in section 2, technology spillovers may have important effects on economic development in and across regions and countries.

Griliches (1979) distinguishes between two types of spillovers, that is, rent spillovers and pure knowledge spillovers. Rent spillovers are pecuniary spillovers which result when the innovating firms are unable to raise prices proportionally to the quality improvements of their products. For the firms that use these products as inputs, this results in a better quality-price ratio, which is interpreted as a spillover.11

The concept of pure knowledge spillovers, on the other hand, is related to 'the impact of the discovered ideas or compounds on the productivity of the research endeavours of others' (Griliches, 1992, pp. 30–1). This corresponds to the impact of 'general knowledge' on the productivity of R&D in Romer's (1990) model. In this context, one may think of 'imitative' spillovers (that is, one firm copying an innovation by another firm), or 'idea-creating' spillovers (when an innovation leads to an idea for another innovation).

Because the impact of rent spillovers is largely related to traded inputs, one could expect that the importance of such spillovers increases when trade barriers in Europe are reduced. This does not necessarily hold for pure knowledge spillovers. Unfortunately, pure knowledge spillovers are a difficult concept to operationalize, given the available indicators. Because no data are available on the R&D-financing links between regions or countries, we cannot follow the more traditional methodology to use these data as an indicator of interaction.12

Instead we use data on patent citations as an indicator of interaction. This follows earlier contributions (using data at the national level) in Verspagen (1997a, 1997b), building on a method proposed by Jaffe, Trajtenberg and Henderson (1993).

Each patent application must refer to previous patent applications. The purpose of patent references is to preclude double patenting of innovations and eventually to limit patent protection. Also, patent references indicate relevant established knowledge for new innovations. One may thus straightforwardly interpret such patent references as indicators of spillovers of knowledge from the cited patent to the citing patent. For the purpose of this chapter, citations in European patents are used as our measure of knowledge spillovers. However, it has to be kept in mind that the large majority of the patent citations is added
by the EPO patent examiners, which implies that the inventors may not have been aware of the cited patent. Still, the citation link may be seen as an indicator of technological relevance. This certainly indicates potential spillovers, although this potential may not have been realized in all cases. However, patents are public knowledge, so professional R&D laboratories can to a certain extent be assumed to be able to extract useful knowledge from existing patents.

It should be emphasized that the concept of knowledge spillovers is much broader than what is captured by our indicator. In terms of the distinction by Griliches introduced above, we look at a specific form of pure knowledge spillovers. Rent spillovers are left out of the analysis completely. Even within the category of pure knowledge spillovers, however, patent citations are only a part of the complete story. In order for patent citations to take place, both the spillover-receiving and spillover-generating region must be actively engaged in R&D, leading to patent applications.

Thus our analysis will refer to a specific form of knowledge spillovers. This may seem to be a narrow perspective, but has the advantage that we can make use of a very detailed and precise database, in contrast to the more general indicators of spillovers that have been used in other parts of the literature.

The citation data were used to set up a list of pairs of cited and citing patent applications. The regional citation list consists of pairs of 112 European regions, plus seven country aggregates (so the total number of observations is 119 x 119). Four of these countries are European ones for which no regional breakdown of the data is available. The remaining three are the USA, Japan and ‘other countries’ (other than the USA, Japan or European countries).

The region-by-region list has a relatively high concentration of intra-regional citations. Approximately 35 per cent of all citations are for observations when the cited and the citing region are one and the same (note that only 0.84 per cent of all observations in a 119 x 119 list are for the cases when the citing and the cited region are identical). The data do not enable us to make a distinction between citations within a firm or establishment, and citations between two different firms located in the same region. Obviously, this distinction is of great relevance for a discussion on regional innovation systems. The distinction between Silicon Valley and Route 128 is illustrative in this respect (see Saxenian, 1994). In Route 128, one finds large, vertically integrated firms which can be expected to have a large number of intra-firm citations. In Silicon Valley one finds many more small firms, operating in an open and interactive system, with, expectedly, more interfirm (but intra-regional) citations. Saxenian (1994) argues that the two innovation ‘systems’ can be expected to show quite different levels of performance (the Silicon Valley system is argued to be more efficient).

With the data we have, however, we are unable to separate the Route 128s from the Silicon Valleys. This is the reason why we will generally leave out the
intra-regional citations from the analysis. This does not imply that intra-regional and intra-firm citations are not an interesting phenomenon. It merely indicates the limitations of the data.

The inter-regional citations show a highly skewed distribution. This is shown in Figure 7.3. Slightly more than half of all regional pairs (7609 pairs of regions) never cites each other's patents. The frequency of citations gradually declines for more intensive citation links. There are only 71 pairs of regions for which the number of citations is 200 or more. Note that the total number of inter-regional citations is approximately 110,000. The average number of citations for all pairs of regions is 8.6 and the average for those that cite each other is about 17.

From this we conclude that strong technological spillovers between European regions, as far as they are related to patent citations, are only found between a relatively small number of regions. These are also the regions which are relatively active in terms of patenting (see Figures 7.1 and 7.2).

Figure 7.3  Frequency of inter-regional citations (vertical axis) vs number of citations (horizontal axis)
5 SPILLOVERS BETWEEN EUROPEAN REGIONS – ECONOMETRIC RESULTS

The aim of this section is to systematically investigate the pattern of patent citations in Europe based on the data just described. Of particular interest is the extent to which geography affects the technological interaction between European regions. The previous sections, however, have indicated several other factors of potential importance for knowledge flows. Thus the effects of technological specialization, productivity gaps (between spillover-receiving and spillover-producing regions), innovative activity, and of national systems of innovation are also taken into account. The analysis is confined to the European regions only (minus Guernsey and Isle of Man), as comparable data on all variables are not available on overseas regions. The new data set consists of 112 × 112 observations (including the four European countries for which regional breakdowns are not available).

To explore whether spillovers are influenced by geography and country borders, we have to control for other factors which are also likely to matter for spillovers. For example, if patent seekers from Ludwigshaven in Germany (the headquarters of the chemistry giant BASF) cite other German patents, this may be due to several distinct effects which do not all necessarily reflect the impact of agglomeration.

First, as pointed out in section 3 above, German regions are highly innovative as compared to the European average. Thus German patents are more likely to cite each other just because of the above-average German patent activity. This effect should not be taken as evidence of clustering effects of spillovers. Technologically active regions are a priori supposed to cite each other more often than technologically inactive regions. In fact, the data reveal a high correlation between citations and the numbers of patents in the related regions.14 This is largely due to the fact that patenting in both the cited and the citing region is a necessary condition for any reference between them at all. For this reason, we construct a dependent variable where the numbers of citations between two regions are expressed as a fraction of the sum of patents in the citing and the cited regions. In other words, what we are trying to explain is not the absolute amount of spillovers as indicated by patent citations, but rather the intensity of this flow compared to total patenting activity in the regions.

Second, patent applicants affiliated to, say, BASF are a priori more likely to cite patents related to chemistry than patents related to other sectors. If Germany is specialized in chemistry, patent applicants from Ludwigshaven are more likely to cite other German patents because of the pre-existing pattern of economic specialization. This type of spillover may be taken as evidence of...
industry-specific spillovers. As discussed in section 2, industry-specific spillovers may have important implications for the effects of economic specialization.

To take industrial specialization into account, we construct a variable called the 'compatibility index', which makes use of the observed pattern of citations between sectors and the regions' sectoral specialization in patenting. If two regions are specialized in sectors that are often observed to cite each other, this combination of regions receives a high score on the compatibility index. The technicalities concerning this index are discussed in Appendix 7.1. The compatibility index, denoted by \( s_{ij} \), ranges between minus one and one, and the impact of the index on the spillovers between two regions is expected to be positive.\(^{15}\)

Distance data were constructed by counting the number of regional borders one has to cross to reach one region from another. This yields a list of distances between all pairs of European regions. The distance variable is denoted by \( d_{ij} \). Technicalities regarding the distance data are discussed in Appendix 7.1.

The literature on national systems of innovation seeks to explore how differences in national history, institutions, policy and traditions may affect countries' innovative capability and competencies. This chapter does not aim to explore all aspects of such national systems. To take into account possible effects of national systems of innovation, however, we include a dummy variable for intra-country citations, as well as dummy variables for each cited and citing country in the sample.

Technology gap models point to a potential for poor countries to catch up with economic and technological leaders. However, as noted in section 2, spillovers do not only depend on technology gaps, but also on absorption capability in the lagging country or region, and technological congruence. Taking into account these two variables, Verspagen (1991) argued that while a large gap indicates a large potential for spillovers, it may also imply a low capacity to assimilate spillovers. Thus (very) low-income regions may become stuck in a kind of underdevelopment trap, unable to make use of spillovers from advanced countries. Medium-income countries may be better placed to take advantage of knowledge created in other countries. In other words, the amount of realized spillovers may well be a non-monotonic function of the size of the gap.

It is hard to judge on a priori grounds what constitutes a large gap (that is, one that hinders spillovers more than it creates potential for it) or a small gap (that is, one that stimulates spillovers). Fagerberg and Verspagen (1996), Paci (1997) and Chapter 6 by Cappelen, Fagerberg and Verspagen in this volume indicate that the productivity gaps between some European regions are substantial, and that the disparity is substantially larger at the regional level.
than at the national level. Thus one would want to allow for a positive as well as a negative impact of the productivity gap on technology spillovers. We therefore allow for a non-linear relationship between spillovers and the productivity gap, by including a GAP-variable, as well as its squared value. The productivity gap variable GAP is defined as the log of the ratio of GDP per capita in the spillover-receiving and the spillover-generating region.

Finally, we take into account the possibility that the distribution of patents between the receiving and the generating region may have an impact on the intensity of patent citations. In order to quantify this, we include (the log of) the two regions’ share of their total patenting, \( P_i/(P_i + P_j) \) and \( P_j/(P_i + P_j) \), as explanatory variables. If these two variables receive an equal coefficient in the regressions, this indicates that an equal distribution of total patenting (the spillover-receiving region patents as much as the spillover-generating region) is most conducive for growth. Should any of the two variables receive a higher coefficient than the other, this indicates that a distribution in favour of that particular region is most conducive to spillovers.

We arrive at the following regression model for our indicator of spillovers between regions, \( SP_{ij} \):

\[
SP_{ij} = \ln \left( \frac{C_{ij}}{P_i + P_j} \right) = \alpha_0 + \alpha_3 \ln \left( \frac{P_i}{P_i + P_j} \right) + \alpha_2 \ln \left( \frac{P_j}{P_i + P_j} \right) + \alpha_3 \ln d_{ij}
+ \alpha_4 \text{COUNT} + \alpha_5 \ln GAP_j + \alpha_6 \left( \ln GAP_j \right)^2 + \alpha_7 s_{ij}
+ \sum_{n=1}^{14} \alpha_{9n} \text{CitedCOUNTRY} + \sum_{n=1}^{14} \alpha_{9n} \text{CitingCOUNTRY} + \varepsilon
\]

\text{COUNT} is a dummy variable for intra-country spillovers. \text{CitingCOUNTRY} and \text{CitedCOUNTRY} are dummies for the citing and cited individual countries (14 of each). A significant \text{CitingCOUNTRY} result indicates that the country’s capacity to absorb spillovers differs significantly from the average. A significant \text{CitedCOUNTRY}-variable indicates country-specific effects in terms of producing spillovers. \( \varepsilon \) is the error term in the regression. The model is estimated by heteroscedasticity consistent least squares.

The results are reported in Table 7.2.\(^6\) \( p \)-values are given in parentheses. The general impression is that the model fits the data well. The relatively high \( R^2 \) indicates an overall good fit, and most of the coefficients for the structural variables are significant at better than 1 per cent probability level.
Table 7.2 Regression results on spillovers, least squares regression, excluding observations with zero citations. Heteroscedasticity-consistent P-values

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>$R^2 = 0.47$</th>
<th>Number of observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\ln(C_i/(P_i + P_j))$</td>
<td>$0.5174105$</td>
<td>$6341$</td>
</tr>
<tr>
<td>$\ln(P_i/(P_i + P_j))$</td>
<td>$0.4984828$</td>
<td>$0.000$</td>
</tr>
<tr>
<td>$\ln(d_{ij})$</td>
<td>$-0.3693099$</td>
<td>$0.000$</td>
</tr>
<tr>
<td>$\text{COUNT}$</td>
<td>$0.441461$</td>
<td>$0.000$</td>
</tr>
<tr>
<td>$x_{ij}$</td>
<td>$0.758228$</td>
<td>$0.000$</td>
</tr>
<tr>
<td>$\ln(GAP_{ij})$</td>
<td>$0.0657392$</td>
<td>$0.084$</td>
</tr>
<tr>
<td>$(\ln(GAP_{ij})^2$</td>
<td>$-0.2888199$</td>
<td>$0.000$</td>
</tr>
</tbody>
</table>

Note: 28 country-specific dummy variables included in regression, but not reported.

The results indicate that there are important barriers to technology spillovers in Europe. This is seen by several results. First, spillovers between a pair of regions decrease significantly with the distance between them. Even if the dependent variable is weighted, making the interpretation harder, the magnitude of the coefficient of the distance variable is large. A 1 per cent increase in distance decreases the expected spillovers (in terms of the constructed weighted dependent variable) between two regions by 0.37 per cent. As an example of the impact of distance, consider the spillovers between Paris and its neighbouring region Picardie. Compared to Picardie, a region that is at distance 2 from Paris (for instance, Wallonie in Belgium) receives 22 per cent less spillovers (ceteris paribus). For a region at distance 8 from Paris (such as Sicily), the percentage is 53.17

Second, the intra-country dummy variable (COUNT) is positive and significant. The magnitude of this variable indicates that country borders significantly hinder knowledge spillovers. This finding gives some support for the importance of national systems of innovation, although we have not investigated whether this result is due to language, institutions or other factors.18

Third, the impact of the GAP-variable should be noted. The sign of the estimated coefficients indicates a hill-shaped parabola, that is, knowledge spillovers decrease with the size of the gap (on both sides of the vertical axis). The top of the parabola (that is, the value of the GAP-variable that, ceteris paribus, maximizes the amount of spillovers) occurs for a value of $\ln(GAP$
slightly less than zero. This indicates that a small productivity gap (that is, the spillover-receiving region lags somewhat behind relative to the spillover-generating region) is most conducive for spillovers. Very poor regions and very rich regions do not receive many spillovers from other regions. For a very rich region, this may indicate that its high technological competency relative to the other region reduces the potential for learning. For very poor regions, the result indicates that poor regions lack absorptive capacity to benefit from technology developed elsewhere. This result thus gives support for the existence of low-growth underdevelopment traps, as found, for instance, in Verspagen (1991) for a large sample of countries at different levels of development. It also indicates that spillovers do not flow so easily between core and periphery, but rather tend to stay within a group of already relatively well-developed regions.

The importance of technological compatibility for spillovers is also supported. Regions specialized in sectors that are observed to cite each other often, do in fact cite each other more often than average regions do.

Finally, patenting activity in both the cited and the citing region impacts positively, (almost) symmetrically and very significantly on spillovers between each pair of regions. This indicates that regions that patent in approximately equal amounts share most spillovers.

Because of the logarithmic relation used, all observations in which either the number of patents or the number of citations is zero were not included in the regressions. About half of all pairs of regions do not cite each other. Thus, excluding these leaves a biased and not representative sample. The easiest way to include the zero observations is to add a small value to the (raw) citations variable. This makes it possible to take the log of these numbers, although they receive small weights due to the fact that logarithmic functions approach minus infinity for numbers approaching zero. The results reported in Table 7.3 are obtained by adding 0.0001 to the observations of inter-regional patent citations.

The signs of the coefficients are the same as in Table 7.2. However, the magnitudes of the coefficients (except for the compatibility index) are larger than in the case when the zero observations were excluded. Also the fit is somewhat reduced and the linear GAP-variable loses significance. These changes relative to Table 7.2 may partly reflect the fact that the logarithmic expression gives a large and negative effect on observations that are very small (smaller than zero).

Since least-square regressions are based on assumptions of normally distributed residuals, and because our addition of 0.0001 is rather arbitrary, the results in Table 7.3 may be based on a mis-specified model. In order to test for a different specification, we experimented with a probit regression model. This model estimates the (marginal) effects on the probability (not the propensity) that two regions cite each other by a (marginal) increase in the explanatory
### Table 7.3  Estimation result when observations of no citation were included - least squares regression. Heteroscedasticity-consistent P-values

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>Number of observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \ln([C_{ij} + 0.0001]/[P_i + P_j]) )</td>
<td>( R^2 = 0.41 )</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>( P )-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \ln(P_i/[P_i + P_j]) )</td>
<td>0.8919232</td>
<td>0.000</td>
</tr>
<tr>
<td>( \ln(P_j/[P_i + P_j]) )</td>
<td>0.9304383</td>
<td>0.000</td>
</tr>
<tr>
<td>( \ln(P_j) )</td>
<td>-1.839696</td>
<td>0.000</td>
</tr>
<tr>
<td>( \ln(GAP_{ij}) )</td>
<td>0.2210878</td>
<td>0.084</td>
</tr>
<tr>
<td>( s_{ij} )</td>
<td>1.022066</td>
<td>0.000</td>
</tr>
<tr>
<td>( \ln(GAP_{ij}^2) )</td>
<td>0.1840636</td>
<td>0.143</td>
</tr>
<tr>
<td>( (\ln(GAP_{ij}^2)^2 )</td>
<td>-0.3519486</td>
<td>0.033</td>
</tr>
</tbody>
</table>

Note: 28 country-dummies included in regression, but not reported.

### Table 7.4  Estimation result when observations of no citation were included - probit estimation. Heteroscedasticity-consistent P-values

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>Number of observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>( Y_{ij} = C_{ij} )</td>
<td>( R^2 = 0.61 )</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>( P )-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \ln(P_i) )</td>
<td>0.7792036</td>
<td>0.000</td>
</tr>
<tr>
<td>( \ln(P_j) )</td>
<td>0.8011874</td>
<td>0.000</td>
</tr>
<tr>
<td>( \ln(P_j) )</td>
<td>-0.4117427</td>
<td>0.000</td>
</tr>
<tr>
<td>( \ln(GAP_{ij}) )</td>
<td>0.5306639</td>
<td>0.000</td>
</tr>
<tr>
<td>( s_{ij} )</td>
<td>0.6786857</td>
<td>0.000</td>
</tr>
<tr>
<td>( \ln(GAP_{ij}) )</td>
<td>0.116141</td>
<td>0.129</td>
</tr>
<tr>
<td>( (\ln(GAP_{ij})^2 )</td>
<td>0.0043662</td>
<td>0.966</td>
</tr>
<tr>
<td>Log-likelihood*</td>
<td>-3340.078</td>
<td></td>
</tr>
</tbody>
</table>

Notes:
28 country-dummies included in regression, but not reported.

Pseudo-\( R^2 = 1 - L_i/L_0 \), where \( L_0 \) is the value of the log-likelihood function with the constant only; \( L_i \) is the value with all the variables included.

*Log-likelihood after 6 iterations.
Europe: one or several systems of innovation?

variables in the model. We thus estimate the probability that (any) spillovers occur, rather than the magnitude of the spillovers. Thus this model has a quite different interpretation to the previous one. Given this different interpretation, we use the number of citations as dependent variable, rather than the relative citation variable used so far. As a consequence, we also substitute \( \ln P_i \) and \( \ln P_j \) for \( \ln(P_i/(P_i + P_j)) \) and \( \ln(P_j/(P_i + P_j)) \) as explanatory variables. The results are given in Table 7.4.

The signs of the estimated coefficients in Table 7.4 are similar to the ones in Tables 7.2 and 7.3, and mostly significant. The only two variables that are not significant are the GAP-variables. Thus, while the amount of spillovers appears to be a non-linear function of the GAP, this does not seem to hold for the probability that any spillovers occur. This may also be the reason for the findings that the significance of the linear GAP-variable disappeared when observations of zero citations were included in linear regression. The obtained pseudo \( R^2 \) is quite high, indicating an overall good fit.

6 CONCLUSIONS

This chapter investigates whether knowledge spillover flows in Europe take place within one large European system of innovation, or within several localized systems of innovation, with little flows between them. The descriptive analysis of innovation activities as measured by patenting statistics revealed that there is indeed a large degree of concentration in terms of patenting. Thus there are clearly some regions or clusters of regions that can be characterized as 'high-tech', and others as 'low-tech'.

We used patent spillovers as an indicator of knowledge spillovers. This means that we focused on a particular part of spillovers only, namely that part which is most directly related to the innovation process itself (so-called pure knowledge spillovers). In other words, our conclusions relate to the impact of spillovers on the efficiency of the invention process rather than to the broad economic impact of technology spillovers. Our analysis reveals that there are four main factors that limit technology flows across Europe.

First, spillovers are more extensive between regions with similar or complementary specialization patterns. Partly this is due to the fact that knowledge flows more easily within sectors than between them. Inter-sectoral spillovers occur mostly between sectors that are technologically linked (for instance electronics and computers).

Second, distance matters a lot for inter-regional citations. There is a clear negative and strong impact of distance on patent citations. There are several aspects that need more investigation, however. The data used for this chapter have no time dimension. Jaffe and Trajtenberg (1996) found evidence that the
impact of distance decreases over time. This study does not allow any
conclusions on this question.

Third, the data reveal that knowledge flows more freely within than across
national borders. Intra-country spillovers are more extensive than inter-country
spillovers. Thus the concept of national systems of innovation seems to be
relevant for technological competencies among European regions. However,
the impact of national systems of innovation may be reduced by developments
in communication technology and by European integration. Thus future
research should investigate whether the impact of country borders has
decreased over time.

Fourth, productivity gaps play an important role in the spillover process
through their impact on absorptive capacity. Spillovers are most effective when
the receiving region lags somewhat (but not too much) behind the spillover-
generating region. Also, we find that regions that patent in approximately equal
amounts (that is, regions that are on approximately equal technology levels)
share most spillovers. This result is to be taken as a confirmation of the result
from technology gap theories that technology diffusion is in no sense
automatic, but demands a certain level of economic development, in addition
to innovative efforts and favourable institutional settings. In particular, this
shows that in the European context, spillovers are mostly taking place between
a limited set of already fairly highly developed regions.

This leads us to the conclusion that the European system of innovation, as
far as the role of knowledge spillovers is concerned, is to be characterized as
one with polarization between several centres, rather than a single system
without major barriers for knowledge flows. Within these individual centres
knowledge flows relatively freely, helped by relatively small productivity gaps,
small geographical distances, absence of national borders and similar or com-
plementary specialization patterns. Across these centres, and between these
and more peripheral regions, there are much fewer technology spillovers.
APPENDIX 7.1

Patents

Data on patents are from the European Patent Office (EPO, 1996). The address of the innovator(s) is used to assign regions (NUTS level 2) to patents. In the data provided by EPO each patent is assigned to one main technology class. Based on the concordance table in Verspagen, van Moergastel and Slabbers (1994) these technology classes are assigned to 22 economic sectors. Thus, the EPO patent data are used to construct data on patenting in 112 European regions, including Ireland, Norway, Denmark and Finland for which regional breakdowns are not available. In the data used in this chapter one mid-Swedish region is lacking. EPO also provides data on patent citations. To preclude or limit patent protection, the patent offices search previous patent applications in the same and related technology classes, and refer to relevant existing patents. These references constitute links between related patents and are thus used as indicators of technological spillovers.

Geography

There is no coherent data available on distance between European regions. For the purpose of this chapter, such data were constructed on the basis of maps of European regions from Eurostat (NUTS level 2) (Eurostat, 1995). The distance between two regions was set as the smallest number of regions one has to cross to reach one region from another one. Thus intra-regional distances were set equal to zero, the distance between two adjacent regions was set equal to one and so on. In the case of sea separating two regions a 'dummy region' was constructed. Thus the distance from, for instance, French regions next to the English Channel and the corresponding English regions was set equal to two. The procedure yielded a list of distances between all pairs of the 113 European regions (one Swedish region is lacking in the other data used).

Economic Development

The data on economic development are regional (NUTS level 2) GDP in PPP per habitant in ECU from 1992 and are taken from Eurostat (1997). GDP data from Norway are from Statistics Norway (1997).

Regional Compatibility

The index for regional sectoral compatibility between two regions (regions $i$ and $j$), $s_{ij}$, was calculated in the following way. The starting point is a sector-by-sector matrix $Z$ which describes the sectoral citation relations. In this matrix,
the element $Z_{pq}$ denotes the number of patents originating from sector $p$ cited by sector $q$. We construct a new matrix $z$ by dividing the elements of $Z$ by the column sums, that is, $z_{pq} = Z_{pq} / \Sigma_p Z_{pq}$. The matrix $z$ describes the distribution of a sector’s received spillovers over spillover-generating sectors. For each region $i$, we now calculate the share of sector $p$ in total patenting as $\sigma_{ip} = P_{ip} / \Sigma_p P_{ip}$, where $P$ is the number of patents. The next step is to calculate, for each region, 22 (that is, the number of sectors) correlation coefficients $\rho_{ip}$ between $z_{pq}$ and $\sigma_{ip}$. Now calculate the share of a region in patenting of sector $p$ as $\chi_{ip} = P_{ip} / \Sigma_i P_{ip}$. The regional sectoral compatibility between regions $i$ and $j$ is now calculated as the correlation coefficient between the 22 observations on $\rho_{ip}$ and $\chi_{ip}$. This correlation coefficient measures to what extent the sectoral patenting structure of region $j$ is likely to be cited by region $i$, given the sectoral structure of $i$ and the sectoral citation linkages. The range of the compatibility index is between minus one and one, where minus one denotes that there is no probability of region $i$ citing region $j$, and one means that the sectoral distribution in patenting in the two regions is perfectly conducive to citation. The matrix of regional compatibility indexes was stacked into a list for all pairs of regions. Note that this measure of regional sectoral compatibility is not symmetric so that, generally, $s_{ij} \neq s_{ji}$. 
NOTES

- We thank Marjolein Caniëls for help in constructing the distance data used in this chapter. We also thank Jan Fagerberg, Bart Los, Eddy Szirmai, Erik Bion, Karl Ove Moene, Arne Melchior and participants at the TSER workshop in Gothenburg, 4–6 June 1998, and at a seminar at the University of Linz (Economics Department), June 1998, for useful comments.


3. The other two were local markets for specialized skill (labour market pooling) and for specialized intermediates.

4. Vernon (1966, p. 192) states that: 'There is good reason to believe, however, that the entrepreneur's consciousness of and responsiveness to opportunity are a function of ease of communication; and further, that ease of communication is a function of geographical proximity.'

5. Krugman (1979) constructs a model of technology gaps in which laggard countries continuously take over old-fashioned products developed in the most advanced countries, which gives rise to a product cycle theory in the Vernon fashion. Krugman (1986) extends this technology gap theory and demonstrates that catch-up may harm the most advanced countries, while technological progress on the frontier increases income in both advanced and developing countries.

6. It is beyond the scope of this chapter to examine the (dis)advantages of patents or R&D indicators in detail (see the survey by Griliches, 1990).

7. In fact, it might be the case that the inventor lives in one region, but works in a different region. (This may be particularly so when inventors are well-paid employees who can afford to live in nice locations at a relatively far distance from their workplace.) However, given that the regional grouping we use consists of rather large geographical areas (often NUTS-1 level), this problem is unlikely to be severe. To assess its impact, we also calculated the numbers of patents based on applicants rather than inventors, and the correlation between the two measures was high.

8. Note that Greece appears more north on the map than it actually is. This is done for typographical reasons.

9. High-tech sectors are usually defined on the basis of R&D intensity (see OECD/EUROSTAT, 1995). The sectors pharmaceuticals (3522), computers and office machines (3825), electronics (3832), aerospace equipment (3845) and instruments (385) are usually considered as high tech (ISIC rev 2 numbers between brackets). We adopt this definition in Table 7.1.

10. Formally, the Herfindahl index is defined as $\sum_i x_i^2$, where $i$ indicates regions and $x_i$ is defined as $X_i \Sigma x_i^2$ ($X_i$ denotes the number of patents).

11. Studies estimating the impact of so-called indirect R&D embodied in traded inputs on productivity (e.g. Coe and Helpman, 1995; for a survey see Mohnen, 1992) are generally within this interpretation of spillovers.

12. Data for R&D-financing links are available at the level of institutional sectors within countries. Thus, for example, one has information on which part of business R&D is financed by government in a particular country. At the regional level, however, these data are not available. Moreover, even the data on the institutional shares in regional R&D are so incomplete as to prevent us from using them in a sample as wide as the one we have in Figures 7.1 and 7.2.

13. In principle, this list can be blown up to take into account economic sectors, that is, a region-by-region by sector-by-sector list. Obviously, this would increase the number of observations dramatically (more than seven million), so only selective use will be made of the sector dimensions, and most of the work will concentrate on the regional dimension. Verspagen (1997b, 1997b) analyses sector-by-sector citations.

14. The correlations between (the log of) the number of citations between two regions and (the log of) these regions' individual number of patents are very close to 0.50.

15. We also experimented with an alternative index that reflected overlap in industrial technological specialization between regions (the exact index was defined as the sum of the absolute
value of differences between sectoral shares in patenting). As the compatibility index captures technological linkages also between different industries, this index was preferred. The regression results were not sensitive to the index used.

16. All reported estimations exclude intra-regional spillovers. These are excluded because we have no way to distinguish between intra-regional, extra-firm spillovers, and intra-regional, intra-firm spillovers, which makes the interpretation of such regressions difficult. We have experimented, however, with regressions including intra-regional spillovers, and this did not change the results markedly. These results are available from the authors on request.

17. Suppressing the other variables, spillovers depend on distance according to the following expression: Spillovers = Adj^0.37.

18. This result indicates that country borders significantly affect the intercept of the regression line. We tested whether this applied also to the slope coefficients of the other variables, but it was not possible to demonstrate any clear effects. Detailed results are available from the authors on request.

19. We tested whether the value of the lnGAP variable for which the top of the parabola occurs is significantly different from zero by using a non-linear Wald Test. The null hypothesis (that is, that the maximum value occurs at lnGAP = 0) is rejected at the 10 per cent level, but not at the 5 per cent level (P = 0.1, F = 2.71).

20. See also Verspagen and De Loo (1998) for a discussion of the time dimension in European patent citation statistics.

REFERENCES


Europe: one or several systems of innovation?


The economic challenge for Europe


