The increasing importance of geographical proximity in knowledge production: an analysis of US patent citations, 1975 - 1997

Jung Won Sonn, Michael Storper

To cite this version:

Jung Won Sonn, Michael Storper. The increasing importance of geographical proximity in knowledge production: an analysis of US patent citations, 1975 - 1997. Environment and Planning A, SAGE Publications, 2008, 40, pp.1020-1039. hal-01022386

HAL Id: hal-01022386
https://hal-sciencespo.archives-ouvertes.fr/hal-01022386
Submitted on 10 Jul 2014

Environment and Planning A 2008, volume 40, pages 1020 – 1039

DOI:10.1068/a3930

The increasing importance of geographical proximity in knowledge production: an analysis of US patent citations, 1975 – 1997

Jung Won Sonn
Bartlett School of Planning, University College London, Wates House, 22 Gordon Street,
London WC1H 0QB, England; e-mail: j.son@ucl.ac.uk
Michael Storper
Department of Geography and Environment, London School of Economics, Houghton Street,
London WC2A 2AE, England; and Institute of Political Studies, 98 rue de l'Université,
75007 Paris, France; e-mail: m.storper@lse.ac.uk, michael.storper@sciences-po.fr
Received 25 January 2006; in revised form 15 August 2006;

Abstract. Much literature suggests that knowledge-production activities are still heavily dependent upon geographically proximate sources of information, in spite of rapid development in telecommunications technology. Some analysts believe that the importance of proximity in knowledge production will eventually disappear with the continued development of telecommunications. The authors analyse patent citations and find that, after controlling for the existing distribution of knowledge-production activities, the proportion of local citations has increased over time. This finding reinforces the notion that in contemporary knowledge production and innovation the role for geographical proximity is increasing.

1 Geographical proximity and innovation: 'not yet dead' or 'increasingly important'?

Advances in telecommunications and information technologies allow information, at least in its codified forms, to flow without much friction from one place to another. At the same time, the transition to a 'knowledge-based' economy has increased the incentives to exploit knowledge produced in other parts of the world (Archibugi and Lundvall, 2001; Lundvall and Johnson, 1994). These two changes, some believe, are causing knowledge-production activities increasingly to be dispersed across space.

Researchers in geography and urban economics, however, have accumulated substantial evidence that geographical proximity between the people and the organizations that produce knowledge remains central to their ability to stay innovative. Case studies of innovative sectors such as biotechnology, semiconductor, aerospace, and media reveal that clustering is a common phenomenon (Aharonson et al, 2004; Bathelt and Graft, 2006; Saxenian, 1994; Scott, 1993). The Internet industry, which is believed to be creating a geography-free world, is itself clustered (Zook, 2000; 2002). These studies show that certain kinds of knowledge flow are subject to constraints related to the cost or feasibility of covering distance (Acs, 2002; Audretsch and Feldman, 1996; Glaeser, 1999; Patel and Pavitt, 1991; Porter, 2000; Storper, 1997; Zucker et al, 1998).

Nonetheless, most of the existing studies track only the locations and flows of knowledge at a fixed point in time. Thus, they do not indicate whether the role of proximity is increasing or declining. This is an important limitation because, without understanding temporal change, we cannot tell whether the current clusterings of knowledge-production activities are mere residuals of the past that will soon disappear or whether they are likely to persist. The few studies that do use time-series data have methodological problems that limit their ability to shed light on this issue. To address this gap, in this paper we investigate historical changes in inventors' tendency to cite locally, and use methodology specifically adapted to this task. We also illuminate the differences among inventor types in the propensity to cite local patents.

---

2 The dearth of empirical research on the changing role of geographical proximity in innovation

Jaffe (1989) pioneered work on the relationship between knowledge production and geographical proximity, by showing that for each U.S state, the number of patents by businesses is positively correlated with the R&D expenditure of universities located in the same state. On the basis of this finding, he infers that there may be knowledge spillovers from universities to businesses. Similar results, but with different data, were obtained by Acs et al (1992). Kelly and Hageman (1999) tested whether knowledge-production activities are still clustered even after controlling for the geographical distribution of production. Their assumption was that in the absence of knowledge spillovers the distribution of knowledge-production activities (as measured by the number of patents) should be analogous to the distribution of production (as measured by the number of workers). By employing a quality-ladder model, with patent counts at the state level, they found that patenting is geographically clustered even after controlling for the location of employment.

The rare studies investigating temporal changes in knowledge flows include those of Jaffe et al (1993), Jaffe and Trajtenberg (1996), and Johnson et al (2006). Jaffe and Trajtenberg (1996) found that in the early years after a patent is approved, citations are made disproportionately by inventors in the same country (a localization effect), but the proportion of foreign citations increases with time, suggesting a first-mover benefit to proximity. However, they did not investigate whether the localization effect today is stronger than it was in the past. Jaffe et al (1993) ran separate t-tests for the 1975 and the 1980 samples of patent citations. From this they discovered that the localization effect was stronger for the 1980 patent citations than it was for the 1975 citations. However, they could not conclude that the localization effect was increasing because they did not completely control for the citation lag. Their 1975 sample had citations with time lags of up to thirteen years and their 1980 sample data contained time lags of up to nine years. If they had removed citations with time lags of more than nine years from their 1975 sample, they would have been able to compare the citation effects of the two samples.

In apparent contrast to these findings, Johnson et al (2006) conclude that the distances between patent coauthors and between citing and cited patents have increased over time. Their indicator is simple physical distance between two points. With today's transport and communications technologies, however, time and costs of travel are not linearly related to distance. Rather, they resemble a step function in which marginal costs rise very steeply once an overnight stay is required. There are also nonlinear social, political, and psychological costs to distance. Distance should therefore be defined as a discontinuous variable with distinctive threshold values. Furthermore, because Johnson et al (2006) do not control for the existing distribution of inventive activities, some of the decreasing importance of distance they find could be due to the long-term historical shift of inventive activities from the Northeastern region to the South and the West of the USA. Since the geographical center of patenting for the nation as a whole shifted over this period, as confirmed by many authors, including Johnson himself (Ceh, 2001; Co, 2002; Hicks et al, 2001; Johnson and Brown, 2004; Ó hUallacháin and Leslie, 2005), the average distance between current patenting activities and past patenting activities will automatically have increased. Thus, all other things equal, it is probable that the average distance between a citing patent and a cited patent will also have increased.

All in all, then, the jury is still out on whether the importance of geographical proximity to innovation is increasing or decreasing.

3 Theoretical perspectives: why might the effect of geographical proximity increase?

Theory provides a number of reasons to believe that geographical proximity remains essential to certain dimensions of the innovation process. Globalization and reductions in spatial transport and communication costs seem to be a double-edged sword: in some ways, they reduce the importance of geographical proximity; but in others, they maintain or even strengthen the role of geographical proximity.

First, improvements in telecommunications and transportation make codified knowledge nearly ubiquitous; but this has the additional effect of increasing the role of tacit knowledge in being on the technological frontier. Tacit knowledge is often geographically bounded, because its production, exploitation, and circulation occur through dense, costly, and culturally embedded interactions between economic entities (Gertler, 2003). Such interactions between economic entities are more easily developed informally at the national scale through the institutional setting of the 'national system of innovation' (Lundvall, 1992; Nelson, 1993).(1) To a certain degree, this institutional support can be developed at the metropolitan or regional scale (Cooke and Morgan, 1998).

Second, increasing organizational flexibility within the business organizations that carry out knowledge production may encourage them to colocate. For example, flexible organizational strategies in knowledge production, such as strategic alliances and joint R&D arrangements are increasingly common. Spin-offs of large firms often subcontract from the large, parent firm, thereby forming another alternative to the hierarchical R&D organization. R&D workers in these flexible, network-type organizations have less 'organizational proximity' than do those in traditional big, hierarchically organized firms. Geographical proximity may be a way to reconcile flexibility and coordination, by reducing the costs of interorganizational knowledge flows. Flexibility in local labor markets, through high turnover, may also be important in local knowledge flows (Almeida and Kogut, 1999). Job turnover is often facilitated by geographical clustering because it facilitates searching and matching. When scientists and engineers change jobs, they take with them to their new workplaces knowledge acquired in their previous employment. Increasing employment flexibility can therefore have the unintended consequence of increasing local knowledge flows.

Third, by shortening product lifecycles, globalization erodes cost and quality advantages more rapidly through faster imitation and cross-penetration of markets. To cope with this situation, innovation must be timelier, often making impossible the investments in codification and standardization that would, in turn, make it possible to transfer knowledge efficiently over long distances (Leamer and Storper, 2001). Alternatively, a heightened pace of innovation might mean a heightened pace of obsolescence, so that there would not be time for patented knowledge to diffuse spatially and hence to be cited over widely spread innovator communities, before it is rendered obsolete. Baptista and Swann (1998) and Baptista (2000) lend support to this notion, showing that companies in industrial clusters innovate more actively.

Short product lifecycles also mean that R&D workers face greater technological uncertainty, which is reflected in their use of noncodified knowledge characterized by a high level of conceptual complexity (Levy and Murnane, 2004). When information is uncertain because it is in development, moreover, planned contact among remotely located economic agents may be insufficient to move mutual understanding forward (Amin and Cohendet, 1999; Leamer and Storper, 2001). In such relationships, there

---

(1) The preference for intranational economic interactions is affirmed by the finding of persistent 'home bias' in trade patterns (McCallum, 1995).

are also moral hazards that can be overcome only in the presence of sustained direct contact (Storper and Venables, 2004). In all these cases, a combination of face-to-face contact and colocation may be efficient solutions to communication relationships.

All of this suggests that both long-distance and local interactions could intensify in the knowledge-based economy, with even multinational companies increasingly embedding themselves in localized innovation systems and acting as long-distance links between such nodes in global networks (Verspagen and Schoenmakers, 2004).

4 Data and methods

4.1 Patent data

Patent citations are regarded as 'noisy' but useful indicators of interactions that lead to invention and innovation (Breschi et al, 2003; Hall et al, 2001). Before we can use them, their limitations must be fully understood. First, some patents are not commercialized and thus do not have any direct impact on economic output. However, Acs et al (2002) found that patent data represent the geography of commercial innovation rather well. Moreover, when the topic of research is knowledge production, rather than its direct economic impacts, noncommercialized patents are relevant.

Second, sectors differ in their propensity to patent. In particular, knowledge production in the service sector, which is very likely to have increasing significance in today's economy, is less likely to be patented than that in manufacturing (Hipp and Grupp, 2005). Within manufacturing, firms in the pharmaceutical and chemical sectors are more likely to patent because they tend to be larger (Scherer, 1983). It should also be noted that knowledge production in the software engineering, biotechnology, and service sectors is increasingly likely to be patented than before, due to case law from the 1980s and 90s (see Jaffe, 2000). Thus, there is an increase in the coverage of technological innovation by the patent system, but the inconsistency caused by this change creates certain difficulties in determining the economic value of patents and R&D. These issues should not limit the use of patent data in research on knowledge production, at least at the aggregate level and for the analysis of long-term trends; but they do call for caution about sectoral composition, firm size, and coverage issues to be exercised in any use of the data.

The third problem concerns whether patent citations represent knowledge flows between innovators. The skeptical position is based on the fact that patent examiners in the US Patent and Trademark Office (USPTO) can add citations to past patents when they think the necessary citations are missing from the inventor's original list of citations (Breschi et al, 2003 page 4). If added by the examiners, a citation does not represent a conscious knowledge flow. In response to this skepticism, a number of studies have been conducted to see if this problem changes the results of empirical studies. Through inventor surveys, Jaffe et al (1998) and Jaffe and Fogarty (2000) concluded that citations added by the examiners correspond closely to the inventors' expertise. Thompson and Fox-Kean (2005) took advantage of changes in patent records in 2001, in which the USPTO indicates who adds each citation. They found that inventors are more likely than examiners to cite other inventors in the same country in their citations. Although they did not extend the research to the subnational level, the findings suggest that earlier measures of the localization of innovation are most likely to have been underestimates. Thus, even though some imprecision remains, patent citations can be regarded as a conservative proxy of local knowledge flows.

Address information for individual inventors and patent assignees provides unique opportunities for geographical analysis. Because of this advantage, an increasing number of geographers and urban economists are using patents in their analyses

(Ceh, 2001; Co, 2002; Johnson and Brown, 2004; Ó hUallacháin, 1999; Ó hUallacháin and Leslie, 2005). Therefore, in spite of a few limitations, Griliches's conclusion that "Nothing else even comes close in the quantity of available data, accessibility, and the potential industrial, organization, and technological detail" (1990, page 1702) as patent citations, remains valid, and motivates the use of these data for the present research.

The main dataset for our research comes from *The NBER Patent Citations Data Files* (Hall et al, 2001). It contains information about patents, their inventors, and citation records as well as the addresses of inventors and assignees. We used data from application years 1975 to 1997. We adopt the Office of Management and Budget's 1993 definition of Consolidated Metropolitan Statistical Areas and Metropolitan Statistical Areas of metro areas. We used the city and the state of the inventors' addresses to assign each US inventor to a metropolitan area. To do so, we used ZipList, a commercial concordance table that connects US zip codes to cities and counties. Among the 4 330 491 inventors in the database, 2 147 894 inventors are based in the US, of whom 2 082 226, or 96.94%, were successfully geocoded at the metro level. Because it was not feasible to do this for foreign inventors, our subnational analysis is limited to US inventors.

4.2 Testing for changes in the importance of proximity

Following the experimental design used by Jaffe et al (1993), when N inventors are involved in a patent, each of the inventors is assigned a $1/N$ fraction of that patent. Accordingly, a citation between a cited patent with N_1 inventors and a citing patent with N_2 inventors was considered a collection of N_2 citations, each of which carries a weight of $1/(N_1 N_2)$. To illustrate, assume there are two patents called patent A and patent B and the former cites the latter. Assume further that patent A, the citing patent, was developed by two inventors, one in Milwaukee, NY, and the other in the UK, and that patent B, the cited patent, was developed by three inventors, one in Los Angeles, CA, another in France, and the other in China. In this example we regarded 1/2 of patent A as coming from Milwaukee, NY, and the other 1/2 as coming from the UK. In the same way, because patent B has three inventors, the locations of Los Angeles, France, and China are each assigned 1/3 of the patent count. The citation from patent A to patent B is regarded as a collection of six fractional citations, each of which carries a weight of 1/6, that is: 1/6 citation from Milwaukee to Los Angeles; another 1/6 citation from Milwaukee to France, another 1/6 citation from Milwaukee to China, another from the UK to Los Angeles; and so forth.

We sum the fractional counts where citing and cited patents were from the same geographical unit, or what we call the total local citation (TLC). We then divide the TLC by the total number of citations to derive the total local citation percentage (TLCP). We exclude citations in which the citing patent and the cited patent were assigned to the same assignee (self-citation), on the assumption that this is most likely the result of an intraorganizational, rather than external but geographically proximate, flow of knowledge.

As noted, only when citations are more highly localized than is patenting activity as a whole can it be said that there is localization in knowledge flows. This requires that we control for the geographical unevenness of knowledge-production activities. To this end, we employed Jaffe, Trajtenberg, and Henderson's control technique (the JTH control technique: Jaffe et al, 1993).(2) The essence of the JTH control technique is the

---

(2) For further explanation, see Jaffe et al (1993, pages 581–583). It should also be noted that the JTH method was originally claimed, by its authors, to measure 'localized knowledge spillovers'. In reality, it measures something more restricted: localization of knowledge flows. JTH's method cannot detect all localized technological spillovers, as shown by Breschi and Lissoni (2003), and Thompson and Fox-Kean (2005). We consider it to be useful in this latter, more restricted sense.

construction of the control variable, which connects each patent with a 'control patent' from the same patent class, application year, and closest grant date to the citing patents. It then counts the number of cases where the geographical location of a cited patent coincides with that of the control patent (JTH control-matching frequency). The frequency of this coincidence is a proxy of the preexisting distribution of knowledge-production activities. From the JTH control-matching frequency, we calculated the JTH control-matching percentage (JTHCMP) by dividing it by the number of total citations.

The net local citation percentage (NLCP) was then calculated according to the following formula:

$$\text{NLCP} = \text{TLCP} - \text{JTHCMP} .$$

By definition, the NLCP increases by either an increase in the TLCP or a decrease in the JTHCMP. An increase in the TCLP would mean an increasing propensity to cite locally, regardless of the distribution of knowledge-production activities. On the other hand, a decrease in the JTHCMP would mean dispersion of knowledge-production activities which in turn, would mean increased opportunities for nonlocal citations. In this last case (that is, when knowledge-production activities become more evenly distributed), the NLCP increases without a change in the TLCP. This means that the NLCP indirectly measures the local flow of knowledge by measuring the unrealized opportunities of nonlocal citations.

If the NLCP at a certain geographical scale is greater than zero, it suggests that inventors disproportionately cite past patents from the same geographical unit after controlling for the existing distribution of knowledge-production activities. In other words, in the process of creating new technology, scientists and engineers depend more on knowledge created by other people in the same geographical unit than on knowledge created in other units, assuming that existing inventive activities are evenly distributed across geographical space; a negative score suggests that local knowledge flows are not important. By measuring the evolution of NLCPs over time, which has not been done before, we determine whether this localization effect in knowledge flows is becoming stronger or weaker.

Whereas Jaffe et al (1993) grouped patents according to the application year of *citing* patents, we grouped them according to the application year of *cited* patents because we are interested in determining whether today's inventors are citing nearby patents more than inventors did in the past. We departed from the original Jaffe et al (1993) method in one other way. They selected a control patent that matched the year and primary patent class of the citing patent application. They then compared geographical overlap between control patent – cited patent pairs and citing patent – cited patent pairs. The former represents a hypothetical matching percentage, where the cited patents are cited by inventors who are local. By contrast, we select a control patent that matches the application year and three-digit primary patent class of the *cited* patent. We compare geographical matching of control patent – citing patent pairs and that of cited patent – citing patent pairs. Jaffe et al (1993) do not clarify why they choose to control for the distribution of citing patents instead of that of cited patents, but either serves their purpose since they want to determine whether the real citations are different from geographically neutral citation behavior. For our research, controlling for cited patents makes more sense because it allows us to take the distribution of knowledge-production activities (cited patents) as given and find out whether citations in a certain year are more localized than those in other years. As such, we built a dataset in which citations are grouped according to the application year of the citing patent.

In our dataset, however, there are truncation problems at both ends. The address of the inventor at the metropolitan level is known only for patents granted between 1 January 1975 and 31 December 1999. Since the review process may take many years, approval decisions have not yet been made for a considerable portion of patent applications made in the years just prior to 1999. Therefore, we had to remove patents applied for in 1998 and 1999. At the opposite end of the dataset, truncation is caused by citation lag. Jaffe et al (1993) noted that the time lag between a citing patent and a cited patent varies significantly. There are only a few citations of certain patents in the first two years after application. From the third to eleventh years, relatively stable numbers of citations are made each year (see Jaffe et al, 1993, figure 1, page 587). However, we are not able to estimate citation lags because patents will continue to be cited in the future. Moreover, the dataset contains address information of patents granted only since 1975. This means that addresses of patents granted prior to 1975 cannot be found even if the citations to those patents were made between 1975 and 1997. Because of this problem, for earlier citing cohorts, substantial portions of citations are made to patents granted prior to 1975. For example, for 1975 citing cohorts, out of 335 499 citations made, 109 943, or 32.8%, were citations to patents granted prior to 1975. These figures are 31.7% for 1976, and 30.0% for 1977, but fall to 8.0% in 1997 and 7.8% in 1996. To compensate—at least partially—for this problem, we removed citations from the database for which the lag is more than seven years. We chose seven years because this is safely after the peak in the distribution of citation lags: more than half of citations occur within seven years after application. In this way, datasets for 1982 to 1997 have citations for which the lags are zero to seven years. However, this measure is only partially effective for citations between 1975 and 1981. The longest citation lag in 1981 is six years rather than seven because 1974 patents are not in the database. The longest lag for 1980 citations will be five years, for 1979 citations four years, and so forth. Therefore, in the first years covered by the database, an increasingly smaller proportion of the citations actually made are in the database.

4.3 Potential and real problems with our method

At this point, it is important to address some of the potential problems with our method. One possible concern regards potential changes in patent classification. If a new class is created, or a class is divided into two or more, patents that would have been in the same class would now belong to different classifications. This possibility may have an unexpected effect on our analysis because our method relies heavily upon patent class. Fortunately, all of the three-digit classes were created before 1975, the year our test period began, and thus we can avoid this problem by using the three-digit classification system.

There are, nonetheless, other issues that call for caution in interpretation of the NLCP results. As mentioned previously, citations with smaller time lags are more likely to be local. This tendency affects our analysis in two ways. Firstly, because we removed citations with lags of more than seven years, our NLCPs could be overestimated. Therefore, we should not regard the NLCP as an 'absolute measure' of local flow of information but, rather, changes in NLCPs as an indicator of the trend in local/non-local knowledge flows. Secondly, because cohorts of the period spanning 1975–81 include only citations in which time lags are less than seven years, NLCPs from this period could be subject to greater overestimation than those of later years. Note, however, that such a bias would make it more likely that we would find decreasing NLCPs in subsequent periods. In light of this, if we still find increasing NLCPs, it should give us more confidence that the importance of proximity really has increased over time.

Another reason why the NLCP should not be treated as an absolute measure of local flow of information is the inconsistency that exists in the legal definition of primary class. The 'primary class' of patent is supposed to represent the 'broadest claim' of that patent but that is not the case for all industries. The semiconductor industry is a prominent exception. If a patent uses semiconductor technology among many input technologies, that patent will be classified as 'semiconductors'. Even if this patent is only weakly related to semiconductor technology, it is still classified as a semiconductor patent, thus associating it with the semiconductor control patent. Because such a patent is only weakly related to the semiconductor industry, it will be cited often by patents in other industries, which will tend to reduce its NLCP. More seriously, because the semiconductor industry has become more important over time, it is likely that the NLCP of later years will be more significantly reduced than that of earlier years. This tendency constitutes another reason why the NLCP should not be interpreted as a measure of the level of local flows of information. But, once again, this feature of the data would only make it easier to reject our hypothesis. If we still find evidence in favor of the hypothesis of increasing proximity effect, such evidence can be considered all the more reliable.

Finally, as Thompson and Fox-Kean (2005) precisely point out, the NLCP would decrease were we to employ a tighter control method. In defining the control patent we assume that two patents in the same class have similar qualities. The question is: how similar is similar enough? The finer the subclass used, the more similar will be the control to the cited patent. Thompson and Fox-Kean (2005) do this and find that the NLCP is significantly reduced. The aforementioned tendency reinforces the point that the NLCP estimates we present are not measures of the levels of local information flows. But as long as we employ the NLCP consistently over time, it sheds reliable evidence on the trend.

All in all, if anything, the results reported below can be considered very conservative indicators of change in local knowledge flows.

4.4 Difference among inventor types

We expect different types of inventors to have different levels of local knowledge flow. In particular, we expect larger companies to be less dependent upon localized knowledge flows—but they are likely to have both local and long-distance flows (Verspagen and Schoenmakers, 2004). Companies with branches in different parts of the world have the resources to send their researchers and engineers to different locations to acquire new information through occasional face-to-face relations. These technological and organizational investments in long-distance communication help them create dispersed 'epistemic communities' (Amin and Cohendet, 1999; Hakanson, 2005; Lissoni, 2003).

To test this hypothesis, we decomposed the cohorts into five categories: 'individuals', 'federal government', 'universities', 'top 500 most innovative companies', and 'other companies'. To identify inventor types, we drew on assignee information in the patent records. An 'assignee' is a natural or legal person to which the legal ownership of the intellectual property rights has been transferred from the inventor. If the inventor worked for a company, a university, or the federal government, the intellectual property rights are automatically transferred to the institution by which he or she is employed. 'Individual' here refers to intellectual property rights transferred from the inventor to an individual. Those patents that are lacking assignees have very likely been developed either by an independent inventor or an owner–inventor. Thus, we classified 'unassigned' patents into the 'individual' category along with those that were originally classified as 'individual'.

To build the list of university assignees, we first selected assignees whose name contains 'univ', 'college', or 'institute' from the list of nongovernment organizational assignees. We then manually removed those that were clearly not universities (for example, Universal Valve Company, Industrial Technology Research Institute), and manually checked all assignees with more than fifty patents between 1975 and 1997, and found additional university-related assignees (for example, Wisconsin Alumni Research Foundation). This procedure may miss some non-US universities, as well as certain independent entities set up by universities to manage their patents. However, this procedure does capture the great bulk of universities. Universities in non-European language countries often have English names and use English names in their US patent applications (for example, University of Tokyo and Seoul National University), which will be also captured by our search for 'univ'. And when a university sets up a separate entity for the management of research, its name usually includes the name of the university (for example, Iowa State University Research Foundation, Research Foundation of State University of New York), which will also be detected by our procedure. Any overlooked universities would be grouped into the 'top 500 companies' or the 'other companies categories because in the original dataset, universities and firms were in the same group. The number of patents from universities comprises only 1 – 3% of these two categories. Thus, even if our university-patent omission is as high as 10%, which it should not be, only 0.1 – 0.3% of changes are made to these two categories.

This method yields 1279 university assignees. Among them, the University of California is ranked first, with 2768 patents between 1975 and 1997, followed by MIT and the University of Texas with 2151 and 1007 patents, respectively, in the same period. Among non-US universities, the University of British Columbia is ranked first with 229 patents and ranked 26th overall. The Hebrew University of Jerusalem is ranked 2nd among foreign universities and 41st overall (see appendix A).

The 500 companies that had the most patents assigned to them in the period spanning 1975 through 1997 were categorized as 'top 500 companies'; the rest were categorized as 'other companies'. This method may undercount the number of patents by the largest companies because some firms in the 'other companies' group may be subsidiaries of 'top 500 companies'. Maurseth and Verspagen (2002) argue that all citations inside metropolitan areas should be removed from the sample, assuming that a company and its subsidiaries are usually colocated and thus intrametropolitan area citations are mostly between a big firm and its subsidiaries. However, since our purpose is to measure changes in the importance of proximity, change in the geography of parent – subsidiary citations would be a potentially important component of the overall pattern of change we seek to identify. Furthermore, the impact of incorrect classification is likely to be minimal, consisting mostly of a possible underestimation of the activity of the 'top 500 companies'. The biggest companies will still be in the 'top 500 companies' category, not in 'other companies'.

In the 'top 500 companies' category, IBM is ranked first, with 26 340 patents, and General Electric second, with 25 868 patents. Hitachi is ranked first among non-U.S. companies and third overall. As expected, Canon, Toshiba, Eastman Kodak, AT&T, Du Pont, Motorola, Mitsubishi, Siemens, and other well-known innovators are ranked highly (see appendix B). Montecatini Edison SPA occupies the 500th spot, with 467 patents. Companies that have less than 467 patents were grouped as 'other companies'. Many companies in this category had only a single patent.

Figure 1 shows that the share of 'top 500 companies' in overall citations decreased from 61% in 1975 to 24% in 1997. This result is consistent with the frequently advanced notion that the innovation process has been reorganized since the 1970s, away from the so-called 'linear innovation model' in which large companies were expected to internalize many steps in the knowledge-production/innovation process, toward an innovation process involving external, or interorganizational relationships between many different actors, and involving a back-and-forth movement between science, knowledge production, and commercialization (Lundvall, 1992). It may equally be the case that externalization of knowledge production by big companies during the period under examination exhibited a local bias.(3) Along these lines, Patel and Pavitt (1991) argued that, while globalization accelerates the exchange of products, firms have historically kept essential knowledge-production activities in the home country where they have privileged access to information sources and dense relational networks.

Other noticeable changes include a rise of 'university' patenting (0.4% in 1975 to 1.4% in 1997) and decline both of the 'federal government' (2.4% to 0.5%) and 'individual' (15.7% to 12.0%). These changes may stem from the increasing commercialization of knowledge production. Universities are becoming more and more active in commercializing their inventions and thus attempt to protect their intellectual property rights by patenting their inventions (Jaffe, 2000). Independent inventors should attempt the same thing. However they cannot always transform their inventions into commercial products, so they are more likely to sell their intellectual property rights to an entrepreneurial entity who can do so. It is also possible

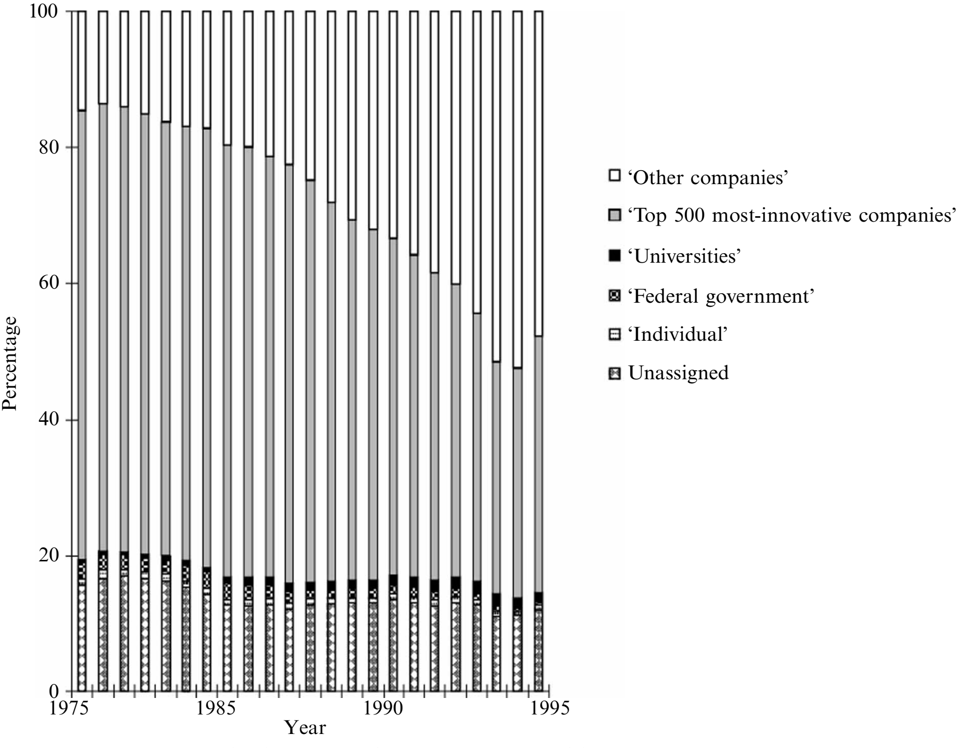

Figure 1. Number of citations by assignee category of citing patents.

---

(3) Some authors argue that knowledge-intensive projects are more likely to be assigned to teams within companies because incentives for knowledge and data production are better aligned in the company's internal labor market (Azoulay, 2003). Yet in basic innovation, where projects are ill defined, external contacts may be necessary as companies search for the right project team (Storper and Venables, 2004). Local bias in outsourcing R&D may represent the best compromise of externalization with the benefits of proximity for coordination.

that the increasing investments required for successful knowledge production may leave less room for individual independent inventors.(4)

5 Findings

5.1 The importance of proximity increases over time

Figure 2 shows increases in the NLCP at three geographical scales: at the country scale, the NLCP increases from 3.0% in 1975 to 7.4 % in 1997; at the state level, it rises from 2.7% to 3.7%; and at the metropolitan level, it begins the period at 0.6% and ends it at 3.1%.

The clear increase in citation activity at the national level supports the argument of Lundvall (1992) and Nelson (1993) that national innovation systems tie together innovative actors, so as both to make them more innovative and to specialize them in fields of innovation.(5) But our national level analysis is influenced strongly by changes in the United States because US inventors account for about half of all US patents. Therefore, it is especially important to determine whether other countries exhibit similar trends. Figure 3 reveals that the increase in localized knowledge flows at the national level is stronger in the US than the other countries' average. That might be because it was in this period that the US regained leadership in the technological development realm thanks to its impressive telecommunications and biotech industry development.

To follow up on this, we examined the NLCP changes within individual countries. For the twenty-nine countries that had more than 500 patents lodged between 1975 and 1997, and that had increased patenting during that same time period, there is a positive correlation between an increase in national NLCP and the log of national patent growth.

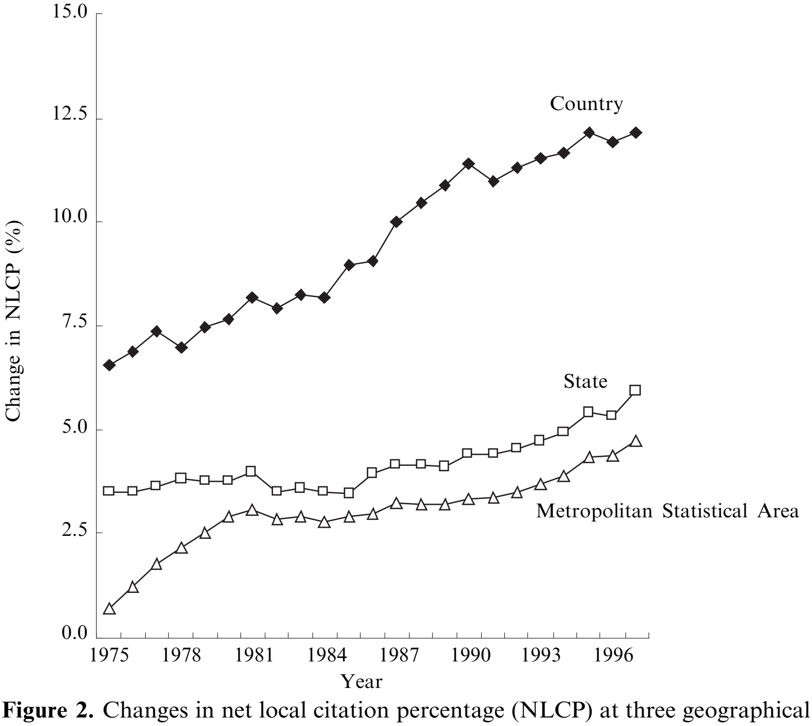

Figure 2. Changes in net local citation percentage (NLCP) at three geographical scales, 1975–97.

---

(4) There is a debate over whether this represents a capital market failure, however—especially in those countries which have weak venture capital sectors.

(5) Cantwell and Iammarino (2001) show that MNEs interact with national strengths and weaknesses in this regard, adapting their strategies to such local economies.

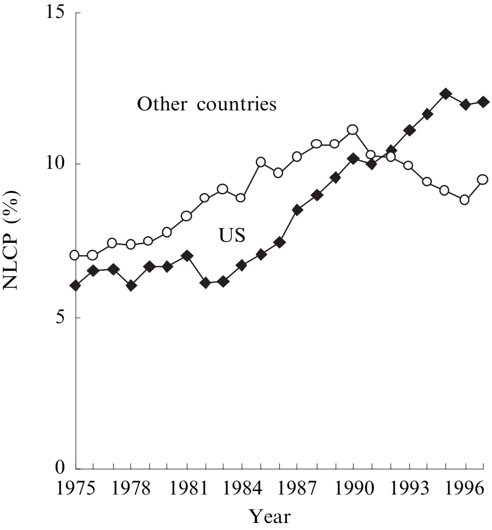

Figure 3. Comparison of the US with other countries in net local citation percentage (NLCP), 1975 – 97.

We used the log of increase in national patent growth because some of the fast-growing countries such as South Korea (20 141.5% growth), China (5333.3% growth), Taiwan (4154.7% growth), and Singapore (3335.7% growth) distort the linear correlation test results. In our test, the Pearson's correlation coefficient is 0.45, which is significant at the 5% level (figure 4). Although it would be premature, without a more systematic causal analysis, to conclude that a country's increasing dependency on local flows of knowledge is the cause of that country's success in patenting, we can say that the successfully patenting countries are also the ones that increasingly exploited domestically created knowledge.

The increasing importance of proximity at the national level seems to also exist at the metropolitan level in the US (although we cannot extend this finding to other countries due to data limitations). This tendency may reflect the outsourcing of innovation by big companies to local inventors, the intensification of science/university – firm relationships, the increasing importance of spin-offs, or a more decentralized form in company – company innovative networks. Unlike at the national level, a correlation test of patent growth and NLCP change for twenty metropolitan areas that had more than 10 000 patents in our test period did not generate any statistically significant results (table 1). This might be because patent growth in metropolitan areas is more strongly affected by nontechnological factors, such as migration patterns, than it is at the national level. More research is necessary to tackle such questions of causality.

Our findings are consistent with prior research on the existence of knowledge flows at the US state level, but it remains unclear whether the state is actually a significant geographical unit of knowledge flow. As Breschi and Lissoni (2001) rightly point out, the state is not a natural geographical scale of innovation, because its borders coincide neither with the boundaries of economic interactions, like the metropolitan area, nor of important policy interventions, such as those of the federal government. To complicate matters, the boundaries of urban areas sometimes extend across state boundaries. As is shown in figure 2, the difference between metropolitan and state NLCPs is small and varied over time. If we could eliminate interaction between geographically proximate metropolitan areas, the remaining NLCPs at state level would be even smaller.

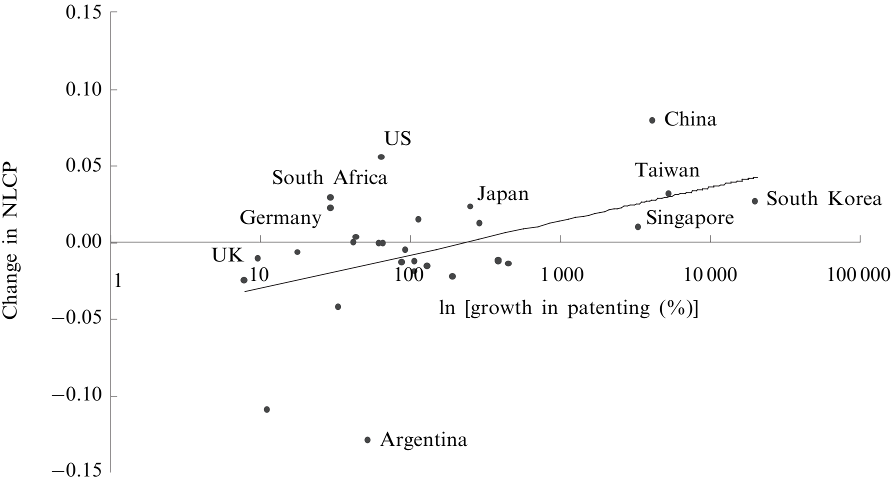

Figure 4. Change in net local citation percentage (NLCP) and national growth in patenting. Note: Not all the countries are labelled because of lack of space. The line is the correlation line (Pearson's correlation coefficient 0.45, significant at 5% level).

One of the possible causes behind the seemingly increasing importance of proximity is a shortening technology lifecycle—suggested by theory (section 3). It is well established in the literature that the share of local citations is higher in the earlier years, soon after the approval of a patent. Therefore, if the technology lifecycle has generally become shorter, the share of local citations would be expected to increase and the average citation lag would be expected to decline. Figure 5 shows the average citation lag of all citations, at the country, state, and metro levels. At all three levels, the average citation lag of local citations in our dataset is systematically shorter

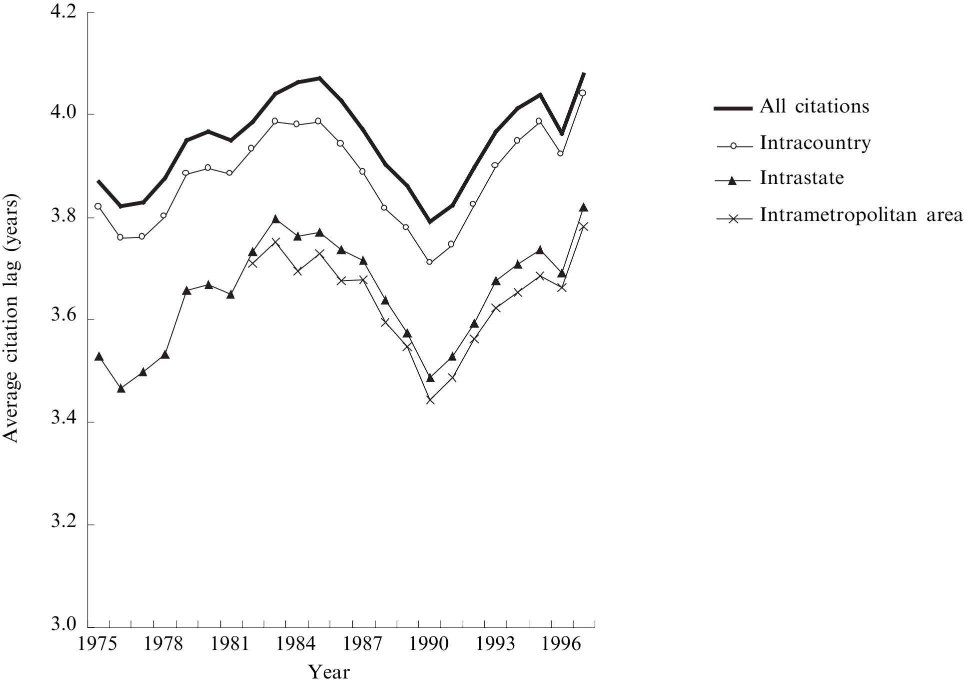

Figure 5. Changes in average citation lag.

Table 1. Patent growth and net local citation percentage (NLCP) changes in twenty metropolitan areas.

Metropolitan area in rank order[a]	Patents				NLCP		
	1975–97	1975–79	1993–97	growth (%)	1975–79	1993–97	change
New York–Northern New Jersey–Long Island CMSA	122 803	27 498	31 327	13.9	0.0200	0.0299	0.0099
San Francisco–Oakland–San Jose CMSA	72 919	9 226	31 410	240.5	0.0221	0.0545	0.0323
Los Angeles–Riverside–Orange County CMSA	66 522	13 904	18 317	31.7	0.0130	0.0381	0.0252
Chicago–Gary–Kenosha CMSA	53 785	12 621	13 209	4.7	0.0405	0.0211	−0.0193
Boston–Worcester–Lawrence CMSA	46 917	8 198	14 611	78.2	0.0379	0.0255	−0.0124
Philadelphia–Wilmington–Atlantic City CMSA	37 635	8 035	9 908	23.3	0.0114	0.0276	0.0162
Detroit–Ann Arbor–Flint CMSA	35 785	6 543	10 691	63.4	0.0283	0.0542	0.0260
Houston–Galveston–Brazoria CMSA	25 393	4 112	7 093	72.5	0.0545	0.0474	−0.0070
Washington–Baltimore CMSA	25 324	4 796	7 735	61.3	0.0383	0.0270	−0.0113
Minneapolis–St Paul MSA	24 868	3 537	8 610	143.5	0.0123	0.0176	0.0054
Dallas–Fort Worth CMSA	20 635	2 822	7 487	165.3	0.0156	0.0172	0.0016
Rochester MSA	19 892	2 729	6 971	155.4	0.0077	0.0014	−0.0063
Cleveland–Akron CMSA	19 820	4 743	4 441	−6.4	0.0299	0.0391	0.0092
San Diego MSA	16 290	2 173	6 349	192.2	0.0367	0.0292	−0.0075
Pittsburgh MSA	15 665	4 077	3 133	−23.2	0.0107	0.0359	0.0252
Seattle–Tacoma–Bremerton CMSA	13 916	1 792	5 284	194.9	0.0292	0.0199	−0.0093
Denver–Boulder–Greeley CMSA	12 862	2 079	4 541	118.5	0.0252	0.0328	0.0076
Phoenix–Mesa MSA	12 825	1 839	4 563	148.1	0.0125	0.0112	−0.0013
Cincinnati–Hamilton CMSA	11 793	2 019	4 044	100.3	0.0103	0.0139	0.0036
St Louis MSA	10 441	2 251	3 059	35.9	0.0082	0.0942	0.0860

[a] CMSA—Consolidated Metropolitan Statistical Area; MSA—Metropolitan Statistical Area.

than that of all citations. However, we could not find any evidence of an overall shortening of product lifecycles. The average citation lag of all citations in our dataset increases until around 1985; it then shifts downward until around 1990, when it begins to rise again. It should be noted that because our dataset excludes citations for which the time lag is longer than seven years, this finding does not allow for definitive rejection of the shorter lifecycle hypothesis. For the same reason, neither does this fluctuation reflect variations in the average citation lag of the entire dataset. However, we can safely say that the technology lifecycle does not explain our finding of increasing NLCP.

The other two theories we suggested in section 3 (namely, increasing flexibility, and increasing importance of tacit knowledge) might contribute to an explanation of our findings, but the present data do not allow testing of them.

5.2 Dependence on geographically proximate knowledge flows increases for most assignee categories

Figure 6 shows that the NLCP has increased for most of the inventor categories. As expected, the 'top 500 companies' depend less on local flows of information than do the other assignee categories at all three geographical levels. As we also anticipated, 'individuals' depend more on local flows of information than do the other assignee categories for most of the test period.

To smooth random fluctuations, we calculated weighted averages of the NLCP of the first five years and the last five years; we then conducted *t*-tests with these two ratios in each assignee category at three geographical levels. Twelve out of the fifteen series of NLCP display statistically significant increases at the 99% confidence level (see table 2). Among the other three, two are at the state level and one at the metropolitan-area level. At the state level, the NLCPs of 'individuals' and 'federal government' decrease. At the metropolitan level, the 'federal government' NLCP did not change.

None of these three exceptions detracts from our general finding of increasing NLCP. This can be seen by considering each in turn. First of all, the relative stability of 'individuals' (small decrease) is likely due to the high level of in-state knowledge flows in 1975. The NLCPs of 'individuals' are consistently higher than any of the other categories for the majority of the test period; therefore, these categories are closer to the 'saturation' point of in-state knowledge flows. This interpretation is supported by the NLCP changes of 'individuals' at the country and metropolitan scales (3.4% increase at country level and 1.6% increase at metropolitan level). They are lower than those of any other categories except the 'federal government' at the metro level.

The stability of flows for 'federal government' at the metropolitan level reflects the fact that the federal government's research activities are less subject to market forces than are other economic actors, and less so even than universities (who must compete for their research funding). Federal government research is allocated through national political processes and bureaucracies, and to some extent it is insulated from the time and price pressures which apply to the other assignee types. Military R&D, which comprises an important part of the federal government's research, has also at times been deliberately geographically isolated—as in the location of national laboratories such as those at Los Alamos.

In sum, saturation theory and the specificities of federal government research policies lead us to conclude that none of the three outliers represents a significant exception to the general finding, for twelve of the fifteen categories, that geographically proximate knowledge flows appear to be increasing. However, the outliers do open up interesting questions for further research.

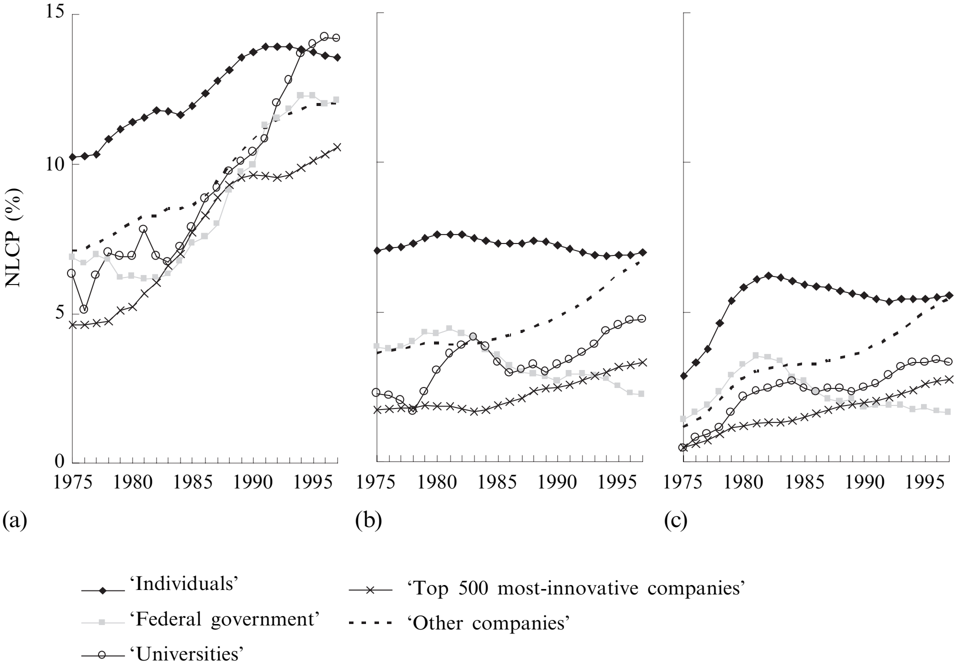

Figure 6. Net local citation percentage (NLCP) by assignee type at three geographical scales (five-year moving averages): (a) country, (b) US state, (c) US metropolitan area.

6 Some implications for research and policy

The geography of knowledge production—like the geography of economic activities more generally—is undoubtedly a complex tapestry of relations at several different scales. This research suggests that some of the pieces may have to fit together in relatively restricted geographical spaces. One of the most important and earliest empirical claims, that multinational enterprises appeared to be deeply linked to their home territories for their most innovative activities, was made by Patel and Pavitt in their widely cited 1991 paper. Since then, there has been considerable empirical confirmation of the effect of geographical proximity in innovation. Until now, however, there has been no indication of whether this phenomenon has increased or decreased with the changing organization of the innovation process and the possibility that innovative activities could be geographically fragmented by the use of advanced information technologies. Our research adds this critical missing element, suggesting that inventors cite local patents *increasingly over time*. This was confirmed for all assignee categories except the 'federal government'. Moreover, even though individual inventors depend more on locally created knowledge than do firms, the national-level 'Patel–Pavitt' effect is still present for large companies and it is increasing over time. This finding thus calls for more investigation of the extent to which very large businesses are able to internalize complex knowledge flows or otherwise manage them over large distances.

The present results leave open most of the issues of causality. Further research on the extent and causes of increasing proximity effects is needed in such areas as: the nature of information, coordination, and contracting issues in a geographical context; the geography of the institutional bases of innovation; the pace of innovation and product cycles; and the timing and geography of innovation diffusion. Such research

Table 2. Changes in net local citation percentage (NLCP) by assignee type.

	1975 – 78		1993 – 96		Change (%) (A − B)	*t*-statistics of (A − B)
	NLCP (%) (A)	number of citations	NLCP (%) (B)	number of citations		
Country						
Individual	10.4	127 775	13.7	348 286	3.4	30.95*
Federal government	6.9	13 475	12.4	15 668	5.4	15.53*
Universities	6.9	4 317	14.1	47 949	7.2	13.30*
Top 500 most-innovative companies	4.7	257 493	10.2	875 890	5.5	85.74*
Other companies	7.3	222 257	12.0	1 071 198	4.6	63.19*
Total	7.1	625 317	11.9	2 358 991	4.8	109.33*
State						
Individual	7.2	122 982	6.9	336 594	−0.3	3.51*
Federal government	3.8	13 208	2.6	15 390	−1.2	5.64*
Universities	1.9	4 277	4.6	47 653	2.7	8.30*
Top 500 most-innovative companies	1.8	239 706	3.2	731 786	1.4	34.89*
Other companies	3.7	208 796	6.3	1 013 372	2.6	45.76*
Total	3.7	588 969	5.3	2 144 795	1.6	51.35*
Metropolitan area						
Individual	3.8	120 496	5.4	328 165	1.6	22.31*
Federal government	1.8	13 016	1.8	15 205	0.0	0.03
Universities	1.0	4 224	3.4	46 628	2.4	8.51*
Top 500 most-innoative companies'	0.7	237 226	2.6	725 649	1.9	54.77*
Other companies	1.6	206 778	5.1	1 000 360	3.5	69.60*
Total	1.7	581 739	4.2	2 116 006	2.5	91.00*

* *t*-statistics valid at 99% confidence level (critical point: 2.58).

could also test the alternative explanation that with increasing patenting activity in general, existing patents are rendered obsolete more quickly—before they are cited nonlocally. To get at the empirical dimensions of these causes, further disaggregation of the geography of patenting will have to be carried out—possibly at the commodity-chain or 'technology-field' levels. Such research would make a major contribution to our understanding both of globalization and of technological knowledge production.

Acknowledgements. The authors would like to thank the participants at the Sussex conference in honour of Keith Pavitt for their helpful comments, as well as Andrés Rodríguez-Pose, Gilles Duranton, Mike Hibbard, Jun Koo, and three anonymous referees for remarks on an earlier version of this paper. During this research, Sonn was supported by the Benjamin Stevens Fellowship for Regional Science and the Leverhulme Early Career Fellowship.

References

Acs Z J, 2002 *Innovation and the Growth of Cities* (Edward Elgar, Cheltenham, Glos)

Acs Z J, Audretsch D B, Feldman M P, 1992, "Real effects of academic research: comment" *American Economic Review* **82** 363 – 367

Acs Z J, Anselin L, Varga A, 2002, "Patents and innovation counts as measures of regional production of new knowledge" *Research Policy* **31** 1069 – 1085

Aharonson B S, Baum J A C, Feldman M P, 2004, "Industrial clustering and the returns to inventive activity: Canadian biotechnology firms, 1991 – 2000", WP 04-03, Danish Research Unit for Industrial Dynamics, Copenhagen

Almeida P, Kogut B, 1999, "Localization of knowledge and the mobility of engineers in regional networks" *Management Science* **45** 905 – 918

Amin A, Cohendet P, 1999, "Learning and adaptation in decentralised business networks" *Environment and Planning D: Society and Space* **17** 87 – 104

Archibugi D, Lundvall B (Eds), 2001 *The Globalizing Learning Economy* (Oxford University Press, Oxford)

Audretsch D B, Feldman M P, 1996, "R&D spillovers and the geography of innovation and production" *American Economic Review* **86** 630 – 640

Azoulay P, 2003, "Acquiring knowledge within and across firm boundaries: evidence from clinical development", WP 10083, National Bureau of Economic Research, Cambridge, MA

Baptista R, 2000, "Do innovations diffuse faster within geographical clusters?" *International Journal of Industrial Organization* **18** 515 – 535

Baptista R, Swann P, 1998, "Do firms in clusters innovate more?" *Research Policy* **27** 525 – 540

Bathelt H, Graft A, 2006, "Internal and external dynamics of the Munich film and TV industry cluster and the international entertainment sector" *Spaces* 1 – 31

Braczyk H-J, Cooke P, Heidenreich M, 1998 *Regional Innovation Systems : The Role of Governances in a Globalized World* (UCL Press, London)

Breschi S, Lissoni F, 2001, "Knowledge spillovers and local innovation systems: a critical survey", LIUC Papers in Economics, Cattaneo University, Castellanza

Breschi S, Lissoni F, 2003, "Mobility and social networks: localised knowledge spillovers revisited", working paper series, Centre for Research on Innovation and Internationalisation, Università Bocconi, Milan

Breschi S, Malerba F, 2003, "Networks of knowledge through patent citations: an analysis of EPO patent citations data", paper presented at DRUID Summer Conference, Copenhagen, http://www.druid.dk/conferences/

Breschi S, Lissoni F, Malebra F, 2003, "Knowledge networks from patent citations: methodological issues and preliminary results", paper presented at DRUID Summer Conference, Copenhagen, http://www.druid.dk/conferences/

Cantwell J, Iammarino S, 2001, "EU regions and multinational corporations: change, stability, and strengthening of technological comparative advantage" *Industrial and Corporate Change* **10** 1007 – 1037

Ceh B, 2001, "Regional innovation potential in the United States: evidence of spatial transformation" *Papers in Regional Science* **80** 297 – 316

Co C, 2002, "Evolution of the geography of innovation: evidence from patent data" *Growth and Change* **33** 393 – 523

Cooke P, Morgan K, 1998 *The Associational Economy : Firms, Regions, and Innovation* (Oxford University Press, Oxford)

Gertler M S, 2003, "Tacit knowledge and the economic geography of context, or the undefinable tacitness of being (there)" *Journal of Economic Geography* **3** 75 – 99

Glaeser E L, 1999, "Learning in cities" *Journal of Urban Economics* **46** 254 – 277

Griliches Z, 1990, "Patent statistics as economic indicators: a survey" *Journal of Economic Literature* **28** 1661 – 1707

Hakanson L, 2005, "Epistemic communities and cluster dynamics: on the role of knowledge in industrial districts" *Industry and Innovation* **12** 433 – 463

Hall B H, Jaffe A B, Trajtenberg M, 2001, "The NBER patent citations data file: lessons, insights and methodological tools", WP 8498, National Bureau of Economic Research, Cambridge, MA

Hicks D, Breitzman T, Olivastro D, Hamilton K, 2001, "The changing composition of innovative activity in the US: a portrait based on patent analysis" *Research Policy* **30** 681 – 703

Hipp C, Grupp H, 2005, "Innovation in the service sector: the demand for service-specific innovation measurement concepts and typologies" *Research Policy* **34** 517 – 535

Jaffe A B, 1989, "Real effects of academic research" *American Economic Review* **79** 957 – 970

Jaffe A B, 2000, "The US patent system in transition: policy innovation and the innovation process" *Research Policy* **29** 531 – 557

Jaffe A B, Fogarty M S, 2000, "The meaning of patent citations: report on the NBER/Case-Western Reserve Survey of Patentees", WP 763,1 National Bureau of Economic Research, Cambridge, MA

Jaffe A B, Trajtenberg M, 1996, "Flows of knowledge from universities and federal laboratories: modeling the flow of patent citations over time and across institutional and geographic boundaries" *Proceedings of the National Academy of Sciences of the United States* **93** 12671 – 12677

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

Jaffe A B, Fogarty M S, Banks B A, 1998, "Evidence from patents and patent citations on the impact of NASA and other Federal labs on commercial innovation" *Journal of Industrial Economics* **XLVI** 183 – 205

Johnson D K N, Brown A, 2004, "How the West has won: regional and industrial inversion in US patent activity" *Economic Geography* **80** 241 – 260

Johnson D K N, Siripong N, Brown A, 2006, "The demise of distance? The declining role of physical proximity for knowledge transmission" *Growth and Change* **37** 19 – 33

Kelly M, Hageman A, 1999, "Marshallian externalities in innovation" *Journal of Economic Growth* **4** 39 – 54

Leamer E E, Storper M, 2001, "The economic geography of the Internet age" *Journal of International Business Studies* **32** 641 – 665

Levy F, Murnane R J, 2004 *The New Division of Labor: How Computers Are Creating the Next Job Market"* (Princeton University Press, Princeton, NJ)

Lissoni F, 2003, "Knowledge codification and the geography of innovation: the case of Brescia mechanical cluster" *Research Policy* **30** 1479 – 1500

Lundvall B-A (Ed.), 1992 *National Systems of Innovation: Towards a Theory of Innovation and Interactive Learning* (Pinter, New York)

Lundvall B, Johnson B, 1994, "The learning economy" *Journal of Industrial Studies* **1** 23 – 42

McCallum J, 1995, "National borders matter: Canada – US regional trade patterns" *American Economic Review* **85** 615 – 623

Maurseth P B, Verspagen B, 2002, "Knowledge spillovers in Europe: a patent citation analysis" *Scandinavian Journal of Economics* **104** 531 – 545

Nelson R R (Ed.), 1993 *National Innovation Systems: A Comparative Analysis* (Oxford University Press, Oxford)

Ó hUallacháin B, 1999, "Patent places: size matters" *Journal of Regional Science* **39** 613 – 640

Ó hUallacháin B, Leslie T F, 2005, "Spatial convergence and spillovers in American invention" *Annals of the Association of American Geographers* **95** 866 – 886

Patel P, Pavitt K, 1991, "Larger firms in the production of the world's technology: an important case of 'non-globalization' " *Journal of International Business Studies* **22** 35 – 54

Porter M E, 2000, "Location, competition, and economic development: local clusters in a global economy" *Economic Development Quarterly* **14** 15 – 34

Saxenian A, 1994 *Regional Advantage: Culture and Competition in Silicon Valley and Route 128* (Harvard University Press, Cambridge, MA)

Scherer F M, 1983, "The propensity to patent" *International Journal of Industrial Organization* **1** 107 – 128

Scott A J, 1993 *Technopolis: High-technology Industry and Regional Development in Southern California* (University of California Press, Berkeley, CA)

Storper M, 1997 *The Regional World: Territorial Development in a Global Economy* (Guilford Press, New York)

Storper M, Venables A J, 2004, "Buzz: face-to-face contact and the urban economy" *Journal of Economic Geography* **4** 351 – 370

Thompson P, Fox-Kean M, 2005, "Patent citations and the geography of knowledge spillovers: a reassessment" *American Economic Review* **95** 450 – 460

Verspagen B, Schoenmakers W, 2004, "The spatial dimension of patenting by multinational firms in Europe" *Journal of Economic Geography* **4** 23 – 42

Zook M A, 2000, "The web of production: the economic geography of commercial internet content production in the United States" *Environment and Planning A* **32** 411 – 426

Zook M A, 2002, "Hubs, nodes, and bypassed places: a typology of E-commerce regions in the United States" *Tijdschrift voor Economische en Sociale Geografie* **93** 509 – 521

Zucker L G, Darby M R, Armstrong J, 1998, "Geographically localized knowledge: spillovers or markets?" *Economic Inquiry* **36** 65 – 86

Appendix A

Table A1. Top universities in patenting 1974 – 97 (source: authors' estimation based on Hall et al, 2001).

Rank	University	Number of patents
1	University of California, The Regents of	2 768
2	Massachusetts Institute of Technology	2 151
3	University of Texas	1 007
4	Stanford University	961
5	California Institute of Technology	852
6	Wisconsin Alumni Research Foundation	849
7	Iowa State University Research Foundation Inc.	667
8	Johns Hopkins University	590
9	University of Minnesota, The Regents of	556
10	University of Pennsylvania	439
11	University of Michigan	406
12	University of Florida Board of Regents	385
13	Research Foundation of State University of New York	374
14	Columbia University	328
15	Harvard College, President and Fellows	327
16	Michigan State University	320
17	Georgia Tech. Research Corporation	283
18	Ohio State University	276
19	Washington University	273
26	University of British Colombia	229
		186
41	Yissum R&D Cc.—Hebrew University of Jerusalem	181

Appendix B

Table B1. Top 500 patenting companies 1974 – 97 (source: authors' estimation based on Hall et al, 2001).

Rank	Company	Number of patents
1	IBM Company	26 340
2	General Electric Company	25 868
3	Hitachi Ltd	19 055
4	Canon Kabushiki Kaisha	18 771
5	Toshiba Company	16 881
6	Eastman Kodak Company	16 032
7	AT&T Corporation	14 836
8	US Philips Company	14 573
9	EI Du Pont De Nemours and Company	13 735
10	Motorola Incorporation	13 681
11	Mitsubishi Denki Kabushiki Kaisha	13 407
12	Siemens Aktiengesellschaft	13 324
13	NEC	12 461
14	Bayer Aktiengesellschaft	12 188
15	Westinghouse Electric Corporation	11 970
16	Matsushita Electric Industrial Company	11 776
17	General Motors Company	11 659
18	Xerox Company	11 638
19	Fuji Photo Film Company Ltd	11 401
20	Sony Company	10 774
500	Montecatini Edison SPA	467