Papers in Evolutionary Economic Geography

15.26

Same Place, Same Knowledge – Same People? The Geography of Non-Patent Citations in Dutch Polymer Patents

Dominik Heinisch, Önder Nomaler, Guido Buenstorf, Koen Frenken, Harry Lintsen

Utrecht University
Urban & Regional research centre Utrecht

Same Place, Same Knowledge – Same People?
The Geography of Non-Patent Citations in Dutch Polymer Patents

Dominik Heinisch [1]
Önder Nomaler [2]
Guido Buenstorf [1,3][*]
Koen Frenken [4]
Harry Lintsen [2]

[1] Department of Economics and International Centre for Higher Education Research (INCHER-Kassel), University of Kassel, Nora-Platiel-Strasse 5, 34109 Kassel, Germany, phone: +49-561-804 2506,fax: +49-561-804 2506, e-mail: buenstorf@uni-kassel.de
[2] School of Innovation Sciences, Eindhoven University of Technology, The Netherlands
[3] Leibniz Institute of Economic Research Halle (IWH), Germany
[4] Innovation Studies, Copernicus Institute of Sustainable Development, Utrecht University, The Netherlands & CIRCLE, Lund University, Sweden

[*] Corresponding author

Abstract

It has long been argued that geographic co-location supports knowledge spillovers. More recently, this argument has been challenged by showing that knowledge spillovers mainly flow through social networks, which may or may not be localized at various geographic scales. We further scrutinize the conjecture of geographically bounded knowledge spillovers by focusing on knowledge flows between academia and industry. Looking into citations to non-patent literature (NPL) in 2,385 Dutch polymer patents, we find that citation lags are shorter on average if Dutch rather than foreign NPLs are cited. However, when excluding individual and organizational self-citations, geographically proximate NPLs no longer diffuse faster than foreign NPLs. This suggests that knowledge is not "in the air" but transferred by mobile individuals and/or direct university-industry collaboration. Our findings moreover suggest an important role of international conferences in the diffusion of recent scientific knowledge.

Keywords: Non-patent literature, citation lags, knowledge spillovers, university-industry interaction, polymer industry.

JEL Codes: O 33, R 10, L 65.

1. Introduction

The past two decades have witnessed a rapidly growing interest in university-industry interaction as governments have aimed at increasing the economic returns to public research. A key issue in this context concerns the relevance of geographically bounded knowledge spillovers from public research to the private sector (Jaffe 1989, Fritsch and Slatchev 2007). This question is of direct policy relevance as knowledge spillovers provide an economic rationale for the public support of research activities, and their geographic range conditions the extent to which the benefits of support measures accrue in the jurisdiction that bears the costs of these measures. The more the diffusion of knowledge is constrained by geography, be it at the regional level or the level of a (small) country, the stronger are the incentives of policy makers in the respective region or country to invest in their domestic universities and public research organizations. In contrast, the rationale for (decentralized) support of public research is much weaker if unmediated knowledge spillovers are either of minor importance or turn out to be geographically unbounded.

In this study, we empirically analyze the relationship between co-location and the time lag between patented technology and non-patent literature (mostly scientific literature) cited by the respective patents. Our empirical approach is based on the insight by Jaffe et al. (1993) that as a piece of initially localized knowledge gets older, it should increasingly be diffused to more distant locations. Accordingly, co-located firms should be faster to absorb new scientific knowledge than more distant ones. In addition, a direct relationship between geographic proximity and the timely access to new knowledge need not be due to unmediated knowledge spillovers (Breschi and Lissoni 2001), but could also derive from other factors such as collaborative research or labor mobility that are favored by geographic proximity. We use several variants of our basic empirical design to probe into alternative mechanisms that also could cause citation lags to vary with geographic proximity.

The empirical context of our analysis is the Dutch polymer industry and the related polymer science. This is a particularly suitable empirical context because science-industry interaction is pronounced in polymer technology, where it has a history of at least 150 years (Walsh 1984, Murmann 2003, 2013). Looking at citation lags to non-patent literature (NPL for short) in all Dutch corporate polymer patents since 1960, we find that Dutch NPL is indeed cited more quickly than foreign NPL. However, this effect disappears once self-citations are controlled for. Faster citations of domestic NPL mostly reflect individual-level self-citations, and to a lesser extent intra-organizational self-citations. These findings suggest that innovation policies should be oriented towards direct interaction and mobility schemes of researchers as carriers of embodied knowledge, since co-location per se may have little effect on the diffusion of knowledge (Breschi and Lissoni 2009).

The remainder of this paper is organized as follows. Section 2 discusses the use of NPL citations as measures of knowledge flows from public research and derives hypotheses on the association of geography and the length of citation lags. Section 3 introduces the data used in the econometric analysis. Results of this analysis are presented in Section 4. Section 5 concludes.

2. Patent citations as a measure of knowledge flows from public research

The idea of geographically bounded knowledge spillovers as an explanation for the regional concentration of industries dates back to Alfred Marshall (1920). In his famous expression, in regions that have a high concentration of firms active in the same industry, industry-specific knowledge may be “as it were in the air" and provide producers in these regions with significant advantages over more isolated competitors. Following Marshall, localized knowledge spillovers have become a cornerstone of theories of industrial agglomeration – notwithstanding the fact that they are notoriously difficult to identify empirically.

Jaffe et al. (1993) first used patent citations as a proxy of knowledge spillovers. Patent applications contain citations to those earlier technologies that are most closely related to the citing patent. If localized knowledge spillovers exist, patent citations will be geographically concentrated. This is indeed found by Jaffe et al. (1993) at different geographic scales (U.S. statistical regions, U.S. states and entire countries), even after controlling for the uneven geographic concentration of firms within an industry. Also consistent with localized knowledge spillovers, these authors find that the localization of citations is reduced with an increasing time lag between the cited and the citing patent. In addition, there is stronger evidence of localization at smaller geographic scales as well as for the citation of university-owned patents.

Besides refining the methodology pioneered by Jaffe et al. (1993)[1], subsequent contributions have challenged the interpretation that patent citations reflect unmediated knowledge spillovers. Feldman (1999) raised early concerns about interpreting local (patent) citations as indicating localized knowledge spillovers. What appears as knowledge spillovers might be caused by other mechanisms at the regional level that are difficult to disentangle. In this context Feldman emphasizes the important role of individuals as carriers of knowledge. Breschi and Lissoni (2001) likewise argue that the localization of patent citations is difficult to explain by knowledge spillovers because patents are codified knowledge having public-goods characteristics and should therefore be available even at longer distances. They also suggest that the benefits of knowledge flows can often be internalized by the affected parties. For example, knowledgeable employees may be able to negotiate higher wages when changing jobs. In addition, social networks and interpersonal contacts play an important role in turning knowledge from a public good into a club good (see also Geroski 1995). Breschi and Lissoni (2001) suggest that the degree of tacitness may be deliberately increased by inventors to restrict the usability of the knowledge disclosed in their patents. In this case, only those agents (i.e., members of the "club") are able to utilize the knowledge who have access both to the patent itself and to the complementary insights that remain non-codified. The latter are personally bounded and shared only within the network. Localization of citations may then result from the localization of social networks, while mere co-location (without being member of the network or "club") does not allow outsiders to access the knowledge.

Using an extended version of the empirical design developed by Jaffe et al. (1993), Breschi and Lissoni (2009) show that substantial shares of patent citations are made within networks of inventors linked by co-inventorship, as well as by inventors having patents with different applicants. Since social networks are often geographically localized, the citation flows are also localized. However, this localization does not stem from co-location per se, but from the proximity of inventors in social networks. These social networks, in turn, are mainly the result of labor moving between companies, hereby connecting intra-organizational teams into larger social networks that span entire industries.

While Jaffe et al. (1993) mainly focused on intra-industry knowledge flows, the idea of geographically bounded knowledge spillovers has also been applied to the diffusion of knowledge from universities and other public research organizations. Again, patent citations can be used to proxy for these knowledge flows, as was for instance done by Jaffe and Trajtenberg (1996) in a follow-up paper analyzing the citations to university-owned patents. They assume that the citation likelihood of a given patent initially increases with time because the underlying knowledge gets increasingly well known. Subsequently the citation

[1] For instance, Thompson and Fox-Kean (2005) show that results are sensitive to the level of technological disaggregation in selecting uncited control patents. Alcacer and Gittelman (2006) highlight the role of citations added by patent examiners.

likelihood is predicted to decrease as the respective knowledge eventually becomes obsolete. Both effects are shown to be significant, resulting in a skewed inversely u-shaped distribution of the citation lag. The maximum number of citations is found for a lag of about three years. The authors also report a positive country-level localization effect on the citation rate. This effect is decreasing with longer time lags, consistent with spatial diffusion of knowledge over time. Jaffe and Trajtenberg (1999) confirm their previous findings on a national scale. The speed of knowledge diffusion is described as dependent on the geographic, institutional and technological distance inventors and inventions are apart from each other. The closer inventors are in these dimensions the higher their chance to profit from each other's newly created knowledge. Institutional self-citations (citations of same corporate organization) are found to have the fastest citation rates.

In the institutional framework of open science, publications rather than patents are the primary channel of disclosing new knowledge. As patents often contain citations to NPL, which indeed mostly consists of scientific publications (Callaert et al. 2006), knowledge flows from public research can be traced by studying the citations to NPL found in patent documents (Roach and Cohen 2013). A NPL document is cited in a patent if no relevant patent literature can be found and the cited document is considered either to contribute significantly to the state of the art or to be a relevant piece of knowledge making the invention of the patented technology possible (Tijssen 2002). Because citing patent documents is typically favored over citing NPL, finding an NPL citation indicates that the patent contributes to a technological field where no (patented) technology existed at the time when the patent application was filed. In addition, NPL represents citations of non-patentable knowledge such as scientific theories or discoveries (Grupp and Schmoch 1992). Since NPL is often scientific literature, NPL citations have been utilized as indicators of the extent to which a technology is science-based (Schmoch 1993).

Patent citations generally are a noisy indicator of (direct) technological impact and knowledge spillovers since citations do not necessarily indicate "intellectual debt" (Henderson et al. 1998). NPL citations may also be affected by strategic considerations at the firm and industry levels. In measuring the impact of science on industrial R&D, NPL citations nonetheless appear to be superior to citations of university patents (Roach and Cohen 2013). Prior research has found that cited scientific publications are not normally the origin of the patented invention (Meyer 2000). The citation is more likely indicating that the patent is in a field where no patent exists as prior art. The cited non-patent literature then is the closest related publicly available knowledge existing at the time of application. If NPL citations are interpreted in this way, it is of secondary concern who included them in the patent document. The final decision whether a document is cited or not is made by the patent examiner (Criscoulo and Verspagen 2008). Independently of where the citation originated, applicant or examiner, the cited literature is seen as relevant by the examiner (because it either proves the novelty of the patent or limits the claims). Accordingly, all NPL citations belonging to a patent application can be used for the analysis, independently of who introduced them.

Several prior studies using NPL citations to trace knowledge flows from public research have found that patents are often co-located with the non-patent literature they cite. Narin et al. (1997) report that U.S. inventors tend to cite scientific literature from the U.S. Cited U.S. authors are most often affiliated to universities, but researchers from some major companies (e.g., DuPont, IBM, or Merck) also receive substantial numbers of citations. The cited scientific literature is characterized as basic and recent. In related work on biotechnology and information technology, Verbeek et al. (2003) likewise find a strong national bias in patent citations to NPL, which is interpreted as evidence of geographically bounded knowledge spillovers.

Analogous to the discussion in Breschi and Lissoni (2001), a variety of factors other than spillovers could give rise to geographic localization of NPL citations. First, as they are exposed to a similar environment, both (publishing) university researchers and (patenting) R&D staff of private-sector firms may be induced to work in the same fields of research. This could for instance result from the support of specific fields of research as parts of regional or national R&D policy initiatives, coordination of universities' research profiles with the needs of regional firms, or even more simply a regional demand for solving specific problems. Second, localized R&D citations might reflect direct collaboration between universities and industries, or the mobility of researchers and graduates from public research to private-sector R&D, where the respective individuals continue to work on similar issues. These conjectures resonate with the findings by Breschi and Lissoni (2009) that social networks are the main carrier of patent citations. Finally, NPL citations might even reflect that the same individuals or research groups both publish and patent closely related results. [2]

Additional insights into the factors underlying the localization of patent citations can be gained from taking a closer look at the timing of citations, which will be in the focus of the empirical analysis presented in this paper. As noted above, it was an early insight by Jaffe et al. (1993) that the localization of spillovers should fade out over time as knowledge diffuses in space via direct interaction and/or is increasingly codified. The same rationale suggests that if scientific knowledge is "sticky" in terms of geography, then we should observe a direct relationship between geographic distance and the time lags of NPL citations: earlier citations of NPL should be more likely to be from geographically close patents than from more distant ones. This conjecture leads to our first hypothesis:

H1: Citations to geographically proximate NPL have shorter citation lags than citations to distant NPL.

Shorter citation lags of co-located patents as predicted by H1 need not be caused by unmediated knowledge spillovers, however. One indication of alternative channels of knowledge flows would be provided by finding extremely short citation lags for co-located NPL. Developing patentable technologies and producing publishable scientific results both require substantial amounts of time. Extremely short citation lags are therefore suggestive of a temporal overlap in private and public research efforts – the new technological and scientific knowledge may both have been generated at the same time, and the citing inventors may have known about the results of the cited research well before these were published.

Besides early familiarity, self-citations at the individual and organizational levels provide a direct measure of the extent to which NPL citations reflect direct interaction within networks or organizations, or even the identity of citing inventors and cited NPL authors, rather than by unmediated Marshallian knowledge spillovers. Various situations may lead to self-citations of NPL. Self-citations may be due to industry scientists who not only patent but also publish (Stephan 1996, Stern 2004), and also to patenting activities of researchers employed at universities and public research organizations (Azoulay et al. 2009). In both cases, we would expect to find self-citations both at the individual and the organizational level. Alternatively, self-citations may reflect that co-located firms and universities collaborate in knowledge production. For instance, we might then find university researchers listed as co-inventors of corporate patents, whereas their university affiliations would normally be listed on the cited NPL. In this case, the self-citation would only be picked up at the individual level, but not at the organizational level. The same would be expected if self-citations reflect labor mobility, for instance when university researchers migrate to the private sector and are subsequently involved in their new employers' patenting activities. Finally, citations may be made to NPL that originated within the same organization (firm or university) but by different

[2] Real patent-paper pairs in the sense of Murray and Stern (2007) are less likely to be observed in our empirical design because papers would normally be published after the patent application and thus not qualify as prior art cited in the patent.

individuals. For patent-to-patent citations Jaffe and Trajtenberg (1999) indeed find shorter citation lags when citing and cited patent originated from the same firm. In this case, it could be argued that the organization-level self-citation reflects Marshallian knowledge spillovers, particularly in large organizations where the respective individuals may not even know each other. It is therefore relevant to identify self-citations at the organizational as well as the individual level.

These considerations suggest that short citation lags of co-located patent-NPL-pairs may be driven by various kinds of processes resulting in self-citations at the individual and/or organizational level. Eliminating self-citations should then reduce the differences in citation lags between co-located and distant citations, which informs our second hypothesis:

H2: Shorter citations lags of geographically proximate NPL are due to self-citations at the individual and organizational level.

In prior studies, various geographic scales have been utilized to trace geographically bounded knowledge spillovers. For instance, the seminal paper by Jaffe et al. (1993) investigates localization at the level of U.S. statistical areas, U.S. states and entire economies. The work by Narin et al. (1997) on NPL citations focuses on the country level and finds a pronounced tendency of U.S. patents to cite domestic science. We follow the lead of these studies and test hypotheses H1 and H2 at the country level in the context of the Dutch polymer industry. In this context, the small size of the Netherlands is noteworthy. In terms of total size, the Netherlands are smaller than most U.S. states, and only about 50% larger than a small U.S. state such as Massachusetts or even an individual region such as the Bay Area.[3] Given the limited size of the country, almost the entire population lives less than two hours apart from each other. Considering the locations of the relevant players in polymer science and industry, travel times are even smaller. Except for the University of Groningen in the North, they hardly exceed 90 minutes.

3. Dataset

A new dataset with NPL citations in the Dutch polymer industry was developed to test the hypotheses formulated above. It originates from the PATSTAT patent database published by the EPO (version April 2008) and is restricted to patent applications by the major Dutch organizations active in polymer technology, including companies (AkzoNobel, Koninklijke DSM, Philips, Royal Dutch Shell, Stamicarbon, Unilever), universities, as well as TNO, the Dutch public research organization of applied scientific research.

A two-stage process was applied to retrieve the patent data from PATSTAT. First, all worldwide polymer patent applications were obtained by searching for patent documents with an international patent classification (IPC) related to polymer technology.[4] In the second step all patent documents associated to one of the selected Dutch applicants were filtered out. A total of 31,679 patent documents was obtained from the patent search. Patent documents were grouped according to DOCDB patent families,[5] which condensed the dataset to 9,003 DOCDB patent families. (In what follows, we will refer to these families as

3 The Netherlands cover an area of about 41,500 square kilometers. This compares to about 27,300 square kilometers for Massachusetts or about 27,600 square kilometers for the San Jose-San Francisco-Oakland CA Combined Statistical Area.

4 The following IPC were included in the query: C08F, C08G, C08K, C08L, C09D

5 The DOCDB patent families combine all patent documents related to the same priority patent applications. This proceeding is necessary to avoid double counting since worldwide patent documents are used. Duplicated patent applications arise if a patent is filed in several countries. For every country a separate patent application has to be submitted at the country's patent office. For the patent examination the priority date of the original filing will be used. If adjustments on the patent application are required new documents are included as well.

"patents".) Information corresponding to the patents (e.g., the year of filing) was obtained from the priority patent document. The further process of data assembly followed the basic procedure of Narin et al. (1997). First, information about patent references was retrieved from the patent documents. Second, citations were searched in a database of scientific literature and categorized by type of publication. Finally, characteristics of the cited patent document were related to those of the cited NPL.

The empirical analysis focuses on the patents' backward citations, using all citations of all members of the respective patent family. In our sample we found 1,088 patents referencing 3,104 NPL citations. NPL can include any kind of publicly available information source (other than patents). There is also no convention on citation rules for the inclusion of an NPL citation. As a result, NPL citations are included in a very unsystematic manner and substantial manual standardization efforts were required. Further information about the NPL was collected from Scopus and Web of Science. If an NPL was not listed in these databases a web-based search was undertaken.[6] By using DOCDB patent families duplicated patent-NPL-pairs might arise if a NPL is cited on several patent documents belonging to the same patent family. In our sample we identified 41 duplicates that had to be excluded from the sample.

To investigate the origin of NPL citations in more detail, the bibliographic information was categorized into several groups. According to document type, NPL was first grouped into (i) scientific journals [subsumed into the variable *journal*], (ii) handbooks [*handbook*], (iii) encyclopedias, (iv) symposia and conference proceedings [*symposia*], (v) company bulletins [*company bulletin*] and (vi) others [*other kind of publication*]. Scientific publication databases were used for the identification of journals. Conference proceedings and publications related to symposia were identified by the name of the respective documents, which commonly contain the word "conference" or "symposium". It was impossible to distinguish between scientific and non-scientific conferences. Therefore the category symposium refers to any kind of publication related to a meeting. Handbooks and encyclopedias were identified by a manual web-based search. Encyclopedias collect knowledge over a long time period and contain entries on knowledge created in different times. Encyclopedias are treated as a separate category because often several editions of the same encyclopedia are cited, which makes it almost impossible to calculate the correct citation lag for them. Because of these difficulties all 115 citations made to encyclopedias were excluded from the analysis. Company bulletins include direct publications of the companies as well as specific journals for defensive publishing.[7] All other documents such as PhD theses, reports of industrial and scientific associations, standards and internet sources were summarized in the category "other kind of publication". In case of doubt a web based search was undertaken. All NPL was manually checked after the categorization.

The NPL was also categorized according to the affiliation of its authors. Because many publications have authors affiliated to different organizations, a single piece of NPL may be associated with multiple affiliations. If an author was affiliated to multiple institutions, all of them were taken into account. Three types of organizations were distinguished: universities [*affiliation university*], research organizations (private and public) [*affiliation research institute*], and companies [*affiliation company*]. Some affiliations could not be identified unambiguously and were categorized as "no affiliation mentioned". In addition, it

---

[6] A similar search was undertaken for the chemical abstracts listed as NPL citations. Chemical Abstracts (CA) show summarized information about all kinds of chemical publications. They are built on a major information source in chemistry about chemical procedures and substances. 359 of the NPL listed are chemical abstracts. In many cases only the abstract identification number is cited. The SciFinder database was used to gain the information of the originating document. CA can refer to any kind of literature. Mostly scientific articles are cited, but also patent information is summarized. In many cases Japanese patents are found to be the disclosing source of the CA (this corresponds to the findings of Grupp and Schmoch (1992)).

[7] To some extent these journals serve similar purposes for companies as patents do. Companies publish their invention to prevent competitors from entering the market. They signal that they are already active in this field and prevent competitors from filing patents on the same technology (Grupp and Schmoch 1992).

proved impossible to clearly distinguish between private and public research organizations. We assume that the location of the affiliation coincides with the place where the knowledge was created. In most cases the country information was directly taken from the original document. If no further information was available the headquarter country for companies and research institutes was used. In addition to Dutch NPL [*affiliation NL*], we also identify NPL from the neighboring countries, Germany [*affiliation DE*] and Belgium [*affiliation BE*]. Because Shell and Unilever are partly located in the UK, another dummy is implemented for the UK [*affiliation UK*]. Regardless of where they originate from, almost all titles suggest that the respective NPL documents are written in English.

For a small number of NPL documents not all required information for the analysis could be retrieved. Because of the crucial importance of timing for this study, all citation pairs had to be excluded for which either no date of the NPL publication or no date of the patent filing was available. This reduced the dataset by another 88 pairs. The final sample contains 828 patents with 2,344 NPL citations. On average each patent contains 2.83 NPL citations. This is similar to the average of 2.68 citations found by Callaert et al. (2006) for European chemistry patents.

Descriptive statistics of the main variables are presented in Table 1[8]. A total of 181 NPL documents (7.72%) are Dutch. U.S. NPL [represented by the variable *affiliation US*] is dominant with 38.31% of all citations. In terms of affiliations, universities account for over 50% of the NPL while companies appear as author affiliations on almost 30% of the NPL. This shows that companies contribute a significant share of all scientific publications relevant to (patented) commercial applications. Dutch NPL is more often company affiliated than foreign NPL. The ratio of Dutch company publications to Dutch university publications (0.76) is substantially higher than that for foreign publications (0.45). With over 75% journal publications are the most prominent type of outlet. Regarding the mean citation lag, a strong variation can be observed between the different NPL categories. With 5.54 years the average citation lag of Dutch NPL is far below the overall average, while more distant locations such as the USA show above average citation lags. This seems to confirm our expectations. In general citation lags of university NPL are slightly below the average, while citations to company affiliated NPL are slightly above the overall average. Dutch universities [captured in the variable *affiliation NL university*], however, are found to have a very low average citation lag of 3.82 years. Some NPL citations date far back in time, reaching a maximum lag of 105 years. Dutch NPL is more recent with a maximum lag of 37 years.

[8] Pairwise correlations of explanatory variables, as well as further control variables introduces in Section 4.2 below, are shown in Table A1 in the appendix.

	sum	share	mean lag	median lag	lag stdev	min lag	max lag	mean patent filing year	mean NPL publication year
full sample	2344	100.00%	11.40	7	12	0	105	1995	1984
affiliation NL	181	7.72%	5.54	3	7	0	37	1996	1991
affiliation BE	17	0.73%	6.59	3	9	0	29	1993	1993
affiliation DE	312	13.31%	12.46	7	14	0	105	1995	1982
affiliation UK	193	8.23%	13.56	6	16	0	96	1982	1982
affiliation US	898	38.31%	12.37	8	13	0	91	1995	1983
affiliation university	1263	53.88%	10.28	6	12	0	105	1986	1986
affiliation company	652	27.82%	11.65	7	13	0	96	1995	1984
affiliation research institute	275	11.73%	9.52	7	9	0	43	1986	1986
affiliation NL university	106	4.52%	3.82	3	5	0	37	1995	1992
affiliation NL company	81	3.46%	7.42	4	9	0	35	1991	1991
affiliation non-NL univ.	1157	49.36%	10.87	6	12	0	105	1997	1986
affiliation non-NL comp.	520	22.18%	12.72	8	14	0	96	1983	1983
no affiliation mentioned	439	18.73%	14.19	11	13	0	91	1993	1979
journal	1776	75.77%	11.57	7	13	0	105	1984	1984
handbook	220	9.39%	14.60	12	13	0	91	1992	1977
symposia	122	5.20%	7.48	6	5	0	27	1990	1990
company bulletin	93	3.97%	8.14	6	8	0	35	1993	1985
other kind of publication	133	5.67%	9.68	7	10	0	58	1984	1984
same author inventor	142	6.06%	3.24	3	3	0	18	1996	1993
same applicant affiliation	60	2.56%	5.23	4	6	0	31	1994	1994

Table 1: Descriptive statistics of NPL subgroups

The distribution of the NPL citation lags is presented in Figure 1.[9] The maximum citation lag is 105 years. The use of priority dates also leads to a total of 49 NPL citations with negative citation lags. Negative lags may emerge for different reasons. Beside some erroneous entries that could not be detected it is possible that the cited literature was already publicly available when the patent application was filed (e.g., in form of a working paper) but the official publication was launched afterwards. In addition, some citations are listed in patents even though they were published after the priority date. Citations to unpublished work can be included if they are seen as important for the understanding of the technology that is claimed to be protected. Such citations do not affect the newness of the invention (Akers 2000). We correct the negative lags to zero based on the assumption that the cited work was existent but unpublished (or published as working paper) when the patent application was submitted. The solid line in Figure 1 visualizes the impact

9 Using the priority date of the patent family causes a slightly different distribution of the citation lag than the calculation on the basis of the patent documents date would provide. The cited literature is shifted closer to the application dates of the citing patents.

of correcting the negative lags. Only five NPL with a negative lag above four years were found. Because no plausible explanation for these observations could be found, they were eliminated from the dataset.

The lag distribution of all NPL is consistent with the conjecture developed by Jaffe and Trajtenberg (1996) that published knowledge becomes better known over time. In addition, older knowledge is also more likely to be obsolete and thus to be cited less often. However, the Dutch NPL shows a different distribution. Besides the shorter tail the high percentage of zero-lag citations is striking. These citations are included in the same year in which the patent application was filed. While this finding is consistent with our first hypothesis – Dutch affiliated NPL are cited more rapidly than foreign NPL – it may also be due to mechanisms other than unmediated knowledge spillovers.

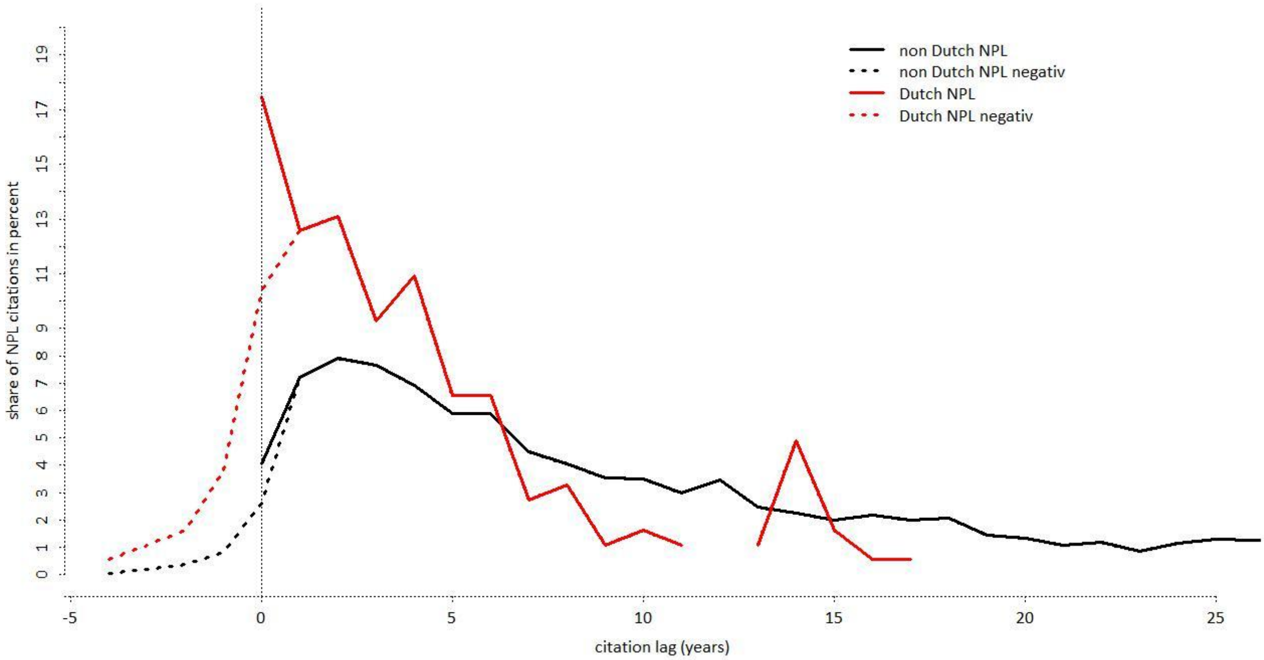

Figure 1: Distribution of the citation lags

4. Econometric analysis

4.1 Basic approach

We use the patent-NPL citation pairs identified in the PATSTAT data and employ hazard models to estimate whether Dutch NPL is cited systematically faster by Dutch polymer patents than NPL originating from other countries. Since citation lags as the outcome of interest are available only in annual intervals, we employ discrete-time hazard models. Specifically, we estimate complementary log-log hazard models, which are commonly used to analyze discrete-time data representing an underlying continuous-time process. These can be specified as follows (cf. Jenkins 1995, 2005):

$$h_j(X) = 1 - \exp\left[-\exp\left(\beta' X + \gamma_j\right)\right],$$

where h_j is the hazard rate for the j-th time interval, β is a vector of coefficients and X a vector of independent variables. We specify γ_i with a logarithmic link function, which corresponds to a continuous-time parametric hazard model in the Weibull specification (ibid.). As some patents cite multiple pieces of

NPL we calculate significance levels based on clustered standard errors in order to account for heteroscedasticity.

4.2 Control variables

In addition to the main explanatory variables described in Section 3 above, several further control variables are included in the analysis. First, the number of other NPL [*number of NPL cited*] and other (patent) citations [*number of patent citations*] was calculated for each patent. Many patents include more than one citation. The number of NPL cited in a patent may serve as an indicator for the relatedness to science of the patented technology (Branstetter 2005). The total number of patent citations reflects the complexity of the patented technology. According to Tijssen (2002), patents with many references have more claims and therefore more technical applications. They are more complex, consequently more literature needs to be included to define the state of the art. More complex and scientific technologies tend to go in line with more recent scientific findings.

Second, the dataset includes many patent families that contain an EPO application. In comparison with individual national applications the European patent application is more expensive (in terms of direct and indirect costs, e.g. for patent attorneys). Applicants might file only those applications at the EPO that are of a higher expected value or have a higher probability to be granted. Besides this potential bias for European patents national examiners might cite more domestic literature (Collins and Wyatt 1988). As noted above, inventors have been found to cite scientific literature from their home country more often than expected (Narin et al. 1997). This might be a result of a better and deeper knowledge of the domestic scientific environment. If examiners have better knowledge about the existing domestic literature they might also cite more recent literature than their colleges at foreign patent offices do. To control for such biases, variables indicating NPL citations originating from the EPO [*patent office EP*] and the Dutch patent office [*patent office NL*], respectively, are included in the model.

Third, the dataset includes all polymer patents filed by the major Dutch companies active in polymer production, by Dutch universities [*university patent*] and by the TNO [*TNO patent*]. University and TNO patents might systematically cite more recent scientific literature. The main objective of universities and TNO is to produce publishable findings rather than to file patents. We therefore expect university inventions to be more closely related to recent scientific findings.

Finally, van Vianen et al. (1990) point to differences in the age distribution of patent citations in different technological fields. Polymers can be used in many different technological fields. Even if the dataset is restricted to polymer patents there might be differences between the chemical subfields where polymers are used. Therefore control variables for the sub-disciplines in chemistry are introduced. The technology classes (IPC) of the patents are classified into industrial fields as suggested by Schmoch (2008). The following nine dummies [*IPC dummies*] are included (numbers in parentheses show how often the class is represented in the dataset, where a given patent can be a member of different subgroups): Organic fine chemistry (840), biotechnology (91), pharmaceuticals (76), food chemistry (14), basic materials chemistry (671), materials, metallurgy (198), surface technology, coating (143), micro-structural and nano-technology (2), chemical engineering (205) and environmental technology (6).

4.3 Localized knowledge spillovers

Regression results for the full sample are presented in Table 2. In Model 1 we find a significantly positive coefficient for the variable denoting Dutch NPL [*affiliation NL*]. In the search for related literature to a patent's technology, Dutch affiliated NPL face a higher "risk" to be cited, that is, are cited more rapidly. This result supports our first hypothesis. Compared to company NPL, NPL from universities [*affiliation university*] and other public research organizations [*affiliation research institute*] is cited more rapidly. In Model 2 we introduce four individual dummies to combine the distinction across types of affiliation with the distinction between Dutch and foreign NPL. To economize on degrees of freedom, NPL from universities and research institutes are aggregated in these dummies. The estimated coefficients indicate that NPL from Dutch public research [*affiliation NL univ./inst.*] are cited more rapidly than Dutch company NPL [*affiliation NL company*], NPL from public research outside the Netherlands [*affiliation non-NL univ./inst.*], and also foreign company NPL [omitted reference group]. Coefficient estimates are significantly different at the 1% level, suggesting that knowledge from domestic public research diffused most rapidly.

In order to treat all citations equally we adjust our dataset further. As Dutch publications have been relevant only since the 1960s (see Table 1) we limit all observed NPL citations to those published after 1960. While we have no real explanation why relevant Dutch publications are not cited before 1960 we decided to give both groups (Dutch and non-Dutch affiliated) the same chance to be cited. We further exclude all university and TNO patents because their citation behavior may differ from that of company patents. [10] Results for the adjusted dataset are presented in Models 3 and 4. Even though the coefficient of Dutch affiliated NPL is reduced, it is still positive and highly significant in Model 3. Also in line with Model 1, university affiliated NPL are found to have systematically shorter citations lags. Model 4 reproduces the finding that NPL from Dutch public research is cited most rapidly. In contrast, the coefficient for Dutch company NPLs is no longer significant at the five percent level.[11]

[10] Note that in Models 1 and 2 university patents were found to cite NPL earlier than companies, indicating a stronger dependence on very recent publications of university patent applications.

[11] The differences between Models 2 and 4 are mainly caused by excluding NPL published before 1960. Using two distinct datasets (one only excluding NPL published before 1960 and one excluding only TNO and university patents) revealed that publications made before 1960 causes the changes in significance by reducing the estimated coefficients. This was already suggested in Table 1 where the maximum citation lag for non-NL companies and universities is substantially higher. Results are available upon request.

	full sample						excluding NPL published 1960 and later as well as university and TNO patents					
	model 1			model 2			model 3			model 4		
(Intercept)	30.1903	*(9.5613)*	***	29.3262	*(9.5841)*	***	29.8954	*(10.8406)*	***	29.3222	*(10.8443)*	***
log(base hazard)	0.1581	*(0.0307)*	***	0.1598	*(0.0311)*	***	0.2885	*(0.036)*	***	0.2881	*(0.0361)*	***
affiliation NL	0.5423	*(0.0964)*	***				0.3070	*(0.1187)*	***			
affiliation university	0.2696	*(0.0578)*	***				0.2032	*(0.0625)*	***			
affiliation research institute	0.1950	*(0.0663)*	***				0.0646	*(0.0708)*				
affiliation NL univ. / inst.				0.9177	*(0.1491)*	***				0.6115	*(0.2067)*	***
affiliation NL company				0.4348	*(0.1298)*	***				0.2500	*(0.1456)*	*
affiliation non-NL univ. / inst.				0.3034	*(0.0636)*	***				0.1635	*(0.0686)*	**
affiliation BE	0.0983	*(0.2836)*		0.0959	*(0.2805)*	*	-0.1441	*(0.4623)*		-0.1331	*(0.4661)*	
affiliation DE	-0.1313	*(0.075)*	*	-0.1222	*(0.0777)*	*	-0.0937	*(0.0742)*		-0.0778	*(0.077)*	
affiliation UK	-0.1989	*(0.0916)*	**	-0.1757	*(0.0923)*	*	-0.1311	*(0.1065)*		-0.1115	*(0.1054)*	
affiliation US	-0.2570	*(0.0487)*	***	-0.2395	*(0.0487)*	***	-0.2265	*(0.0524)*	***	-0.2083	*(0.0517)*	***
handbook	-0.2589	*(0.071)*	***	-0.2528	*(0.0707)*	***	-0.3155	*(0.0688)*	***	-0.3153	*(0.0683)*	***
symposia	0.5430	*(0.1009)*	***	0.5762	*(0.103)*	***	0.3741	*(0.1033)*	***	0.3759	*(0.1063)*	***
company bulletin	0.4451	*(0.1479)*	***	0.4634	*(0.1484)*	***	0.2465	*(0.1636)*		0.2381	*(0.1633)*	
no affiliation mentioned	-0.2330	*(0.0663)*	***	-0.1984	*(0.0693)*	***	-0.2359	*(0.0674)*	***	-0.2401	*(0.0707)*	***
chemical abstract	-0.2568	*(0.0789)*	***	-0.2647	*(0.08)*	***	-0.2693	*(0.1049)*	**	-0.2687	*(0.1062)*	**
other kind of publication	0.1756	*(0.1019)*	*	0.1590	*(0.1037)*	*	0.1727	*(0.1104)*		0.1723	*(0.1097)*	
number of patent citations	-0.0010	*(0.0025)*		-0.0005	*(0.0025)*	*	-0.0009	*(0.0029)*		-0.0006	*(0.0029)*	
number of NPL cited	-0.0055	*(0.0029)*	*	-0.0056	*(0.0029)*	**	-0.0011	*(0.0037)*		-0.0013	*(0.0037)*	
patent filing year	-0.0165	*(0.0048)*	***	-0.0160	*(0.0048)*	***	-0.0164	*(0.0054)*	***	-0.0161	*(0.0054)*	***
patent office NL	-0.1141	*(0.0766)*		-0.1154	*(0.0758)*	*	-0.0076	*(0.0906)*		-0.0022	*(0.0905)*	
patent office EP	0.0730	*(0.0582)*		0.0743	*(0.0576)*	*	0.0810	*(0.0635)*		0.0893	*(0.0633)*	
TNO patent	0.2268	*(0.1592)*		0.2294	*(0.1576)*	*						
university patent	0.4926	*(0.1243)*	***	0.4972	*(0.1241)*	***						
IPC dummies	TRUE			TRUE			TRUE			TRUE		
n	2344			2344			2006			2006		
logLik	-7936.3152			-7933.5965			-6560.6582			-6561.0162		
(p > chi2)	0.0000			0.0000			0.0000			0.0000		

*: p < 0.1 regression coefficient
**: p < 0.05 (standard errors in brackets)
***: p < 0.01

Table 2: Estimation of the citation lag

In Hypothesis 2 we suggested self-citations at the organizational and individual level as a possible driver of shorter citation lags. To probe into this conjecture, we took a closer look at the NPL citations in our restricted dataset limited to company patents and post-1960 NPL. In particular we searched for NPL authors having the same affiliation as the citing patent [subsumed into the variable *same applicant affiliation*] and also identified self-citations at the individual level (at least one inventor is also listed among the authors of the NPL) [*same author inventor*]. To this purpose, family names of inventors and authors were matched by patent. Inventor names were first cleaned. Where positive matches were found, a manual check for false positive matches was made by comparing initials. As most names in the dataset are Dutch and frequently contain several initials, this method allowed for reliable identification of individuals. Affiliations of the authors of Dutch NPL were manually standardized and matched with patent applicant names. For non-university affiliations that were not in the list of applicants we checked whether the affiliation is a branch or subsidiary of one of the patent applicants in the dataset. Descriptive statistics of these additional variables are presented in Table 1. For individual self-citations very short average citation lags are found (mean: 3.24 years). Average citation lags of organizational self-citations are also short even if they are a bit longer than those of the individual ones.

129 Dutch NPL are cited by the companies included in our sample. For 52 NPL one of the authors' affiliations corresponds to the patent applicant; i.e. they are based on organization-level self-citations. For

97 NPL or about three quarters of all citations made to Dutch affiliated NPL, at least one author is listed as inventor on the citing patent; i.e. these citations reflect individual-level self-citations. The overlap of the two sets is 29 NPL sharing the affiliation and at least one author. Altogether, these numbers indicate that 120 of the 129 citations made to Dutch affiliated NPL are either affiliated to the patenting company or have at least one inventor and author in common. It is remarkable that among the 97 NPL patent pairs with a common author, 65 NPL have at least one author listed who is affiliated to a university. This is especially true for citations with short citation lags, suggesting that the respective NPL citations may reflect direct university-company collaboration or recent labor mobility of university researchers to the private sector.

The large numbers of correspondences at the individual and organizational level seem to confirm our second hypothesis suggesting that self-citations may be responsible for the shorter time lags of co-localized NPL citations. To test this hypothesis we re-estimate various variants of the earlier models with one or both types of self-citations excluded from the dataset.[12] In this way we hope to estimate only "real" localized knowledge spillover effects, since all internal knowledge sources (same inventor and author and/or same affiliation) are excluded. The results of these estimations are presented as Models 5-10 in Table 3. In Models 5 and 6, all self-citations at the organizational level are excluded, in Models 7 and 8 all individual-level self-citations. Models 9 and 10 exclude both types of self-citations.

Consistent with Hypothesis 2, the coefficient of Dutch affiliated NPL [*affiliation NL*] is reduced and no longer significant when self-citations are excluded. This holds for all three tested variants of the dataset (taking values close to zero in Models 7 and 9). What appeared to be the result of localized knowledge spillovers in the earlier models is thus found to depend strongly on citations that are closely related to the patenting company. Classifying NPL both by type and location, no significant difference in citation lags is found for Dutch public research [*affiliation NL univ./inst.*] when individual self-citations are excluded (Models 8 and 10).[13] In contrast, we still find significant coefficients for NPL from Dutch public research in Model 6 that only excludes organizational self-citations. This suggests that not controlling for all types of intermediated knowledge sources could lead to faulty conclusions about the existence of local knowledge spillovers. It also suggests the importance of university-industry collaboration and/or individual labor mobility in the diffusion of knowledge from Dutch public research to the Dutch polymer industry. Among foreign NPL, publications from public research are cited significantly more rapidly than those originating from companies.

[12] We decided to exclude self-citations in favor of including dummy variables for reasons of simplicity. In the Appendix, we provide results from two models using the alternative approach of controlling for self-citations. Model A1 corresponds to Model 3 but includes controls for individual-level and organization-level self-citations. Model A2 is a variant of Model 4 that distinguishes affiliation groups for observations with or without self-citations (at both levels). The results of these models are consistent with those discussed in the text.

[13] We verified that our results are not dependent on the zero corrected negative citation lags by repeating all estimations with a sample excluding observations with negative lags. Results were unaffected by this modification.

	excluding self citations (same affiliation)						excluding self citations (same author)						excluding self citations (same author and same affiliation)					
	model 5			model 6			model 7			model 8			model 9			model 10		
(Intercept)	30.7105	*(10.8461)*	***	30.6475	*(10.8139)*	***	33.8412	*(11.1277)*	***	33.4379	*(11.1208)*	***	32.7878	*(11.1571)*	***	32.7048	*(11.1398)*	***
log(base hazard)	0.3040	*(0.0355)*	***	0.3027	*(0.0356)*	***	0.3207	*(0.0363)*	***	0.3196	*(0.0364)*	***	0.3195	*(0.0362)*	***	0.3183	*(0.0363)*	***
affiliation NL	0.1990	*(0.144)*					0.0377	*(0.121)*					0.0664	*(0.1426)*				
affiliation university	0.1868	*(0.0631)*	***				0.1904	*(0.0634)*	***				0.1712	*(0.0639)*	***			
affiliation research institute	0.0605	*(0.0714)*					0.0716	*(0.0726)*					0.0634	*(0.0729)*				
affiliation NL univ. / inst.				0.4583	*(0.221)*	**				0.3242	*(0.2155)*					0.2374	*(0.2279)*	
affiliation NL company				-0.0698	*(0.2254)*					-0.0828	*(0.154)*					-0.1136	*(0.2226)*	
affiliation non-NL univ. / inst.				0.1508	*(0.0697)*	**				0.1543	*(0.0695)*	**				0.1392	*(0.0704)*	**
affiliation BE	-0.2511	*(0.4748)*		-0.2400	*(0.4757)*		-0.4885	*(0.4236)*		-0.4853	*(0.4247)*		-0.4981	*(0.4247)*		-0.4956	*(0.4255)*	
affiliation DE	-0.0874	*(0.0756)*		-0.0614	*(0.0794)*		-0.0607	*(0.0763)*		-0.0468	*(0.0784)*		-0.0637	*(0.0766)*		-0.0480	*(0.0799)*	
affiliation UK	-0.1292	*(0.1021)*		-0.0879	*(0.1018)*		-0.1644	*(0.1033)*		-0.1425	*(0.1039)*		-0.1620	*(0.104)*		-0.1354	*(0.1058)*	
affiliation US	-0.2231	*(0.0532)*	***	-0.2013	*(0.0534)*	***	-0.2238	*(0.0532)*	***	-0.2078	*(0.0526)*	***	-0.2225	*(0.0536)*	***	-0.2074	*(0.0537)*	***
handbook	-0.3290	*(0.0706)*	***	-0.3405	*(0.0702)*	***	-0.3174	*(0.0702)*	***	-0.3195	*(0.0699)*	***	-0.3241	*(0.0711)*	***	-0.3298	*(0.0709)*	***
symposia	0.3870	*(0.1065)*	***	0.3794	*(0.1095)*	***	0.3833	*(0.1086)*	***	0.3780	*(0.1117)*	***	0.3836	*(0.1104)*	***	0.3758	*(0.1133)*	***
company bulletin	0.2335	*(0.1633)*		0.2240	*(0.1634)*		0.2507	*(0.1656)*		0.2390	*(0.1653)*		0.2383	*(0.1645)*		0.2272	*(0.1648)*	
no affiliation mentioned	-0.2516	*(0.0666)*	***	-0.2550	*(0.0702)*	***	-0.2261	*(0.0677)*	***	-0.2367	*(0.0707)*	***	-0.2404	*(0.0674)*	***	-0.2489	*(0.0711)*	***
chemical abstract	-0.2772	*(0.1108)*	**	-0.2800	*(0.1118)*	**	-0.2898	*(0.1093)*	***	-0.2931	*(0.1106)*	***	-0.2909	*(0.1105)*	***	-0.2944	*(0.1116)*	***
other kind of publication	0.1703	*(0.1134)*		0.1738	*(0.1126)*		0.1752	*(0.1129)*		0.1775	*(0.1123)*		0.1740	*(0.1144)*		0.1765	*(0.1139)*	
number of patent citations	-0.0009	*(0.003)*		-0.0004	*(0.003)*		-0.0001	*(0.003)*		0.0001	*(0.003)*		-0.0003	*(0.0031)*		0.0000	*(0.0031)*	
number of NPL cited	-0.0005	*(0.0038)*		-0.0009	*(0.0038)*		-0.0002	*(0.0037)*		-0.0004	*(0.0037)*		-0.0002	*(0.0037)*		-0.0004	*(0.0038)*	
patent filing year	-0.0168	*(0.0054)*	***	-0.0168	*(0.0054)*	***	-0.0184	*(0.0056)*	***	-0.0182	*(0.0056)*	***	-0.0179	*(0.0056)*	***	-0.0178	*(0.0056)*	***
patent office NL	-0.0428	*(0.0912)*		-0.0239	*(0.0911)*		-0.0054	*(0.0902)*		0.0039	*(0.0905)*		-0.0172	*(0.091)*		-0.0052	*(0.0912)*	
patent office EP	0.0914	*(0.0647)*		0.1036	*(0.0645)*		0.0919	*(0.0646)*		0.1015	*(0.0646)*		0.0965	*(0.0651)*		0.1060	*(0.0651)*	
IPC dummies	TRUE			TRUE			TRUE			TRUE			TRUE			TRUE		
n	1954			1954			1909			1909			1886			1886		
logLik	-6408.7153			-6410.4676			-6295.2248			-6296.5892			-6221.4928			-6223.0793		
(p > chi2)	0.0000			0.0000			0.0000			0.0000			0.0000			0.0000		

*: p < 0.1 regression coefficient
**: p < 0.05 (standard errors in brackets)
***: p < 0.01

Table 3: Estimation of the citation lag (reduced sample)

4.4 Publication outlets

Finally, we look at how the different types of publication outlets are associated with the citation lags of NPL. All the above models include dummy variables denoting handbook publications [*handbook*] containing highly codified knowledge, symposia and proceedings [*symposia*] containing the most recent knowledge, and company bulletins mainly used for a company's own publications. Journal articles are the omitted reference category throughout.

Irrespective of the specific model specification and sample, we consistently find that handbooks have significantly lower citation hazards, i.e. they are cited less rapidly than journal articles. Most likely this reflects that handbooks contain general, highly codified knowledge that sets the standard for longer time periods. More interesting is the significantly positive coefficient that is consistently obtained for conferences and symposia proceedings, indicating that they are an important publication medium in polymer science. The knowledge published in those proceedings is not only related to the patented technology but also recent referring to state-of-the-art research findings. This is especially striking since only six of the 122 symposia were held in The Netherlands. Hence, it seems that the temporary co-location of researchers, both from academia and industry, at international symposia provides a key channel of knowledge diffusion (cf. Torre 2008). In contrast, no significant differences are found between journal publications and company bulletins [*company bulletin*] in Models 3-10. This is unexpected since company bulletins may be used for defensive publishing and could thus provide firms with a faster and cheaper alternative to patents in protecting against patenting by competitors.

5. Conclusion - Is there still knowledge in the air?

Studying citation lags of NPL citations in Dutch polymer patents, we found that NPL with Dutch affiliation tend to be cited earlier than geographically distant NPL. While this seems to provide *prima facie* evidence of localized knowledge spillovers, the lag distribution of co-located NPL indicates that different mechanisms give rise to the shorter citation lag. A large share of localized NPL citations has a lag of zero, which is indicative of simultaneous production of scientific and technological knowledge within a region. A closer look into the nature of co-located NPL citations revealed an important role of organizational and individual self-citations. Excluding these self-citations, significant differences in the citation lags of Dutch and foreign NPL are no longer obtained, which suggests a minor role of unmediated localized knowledge spillovers for the access of innovators to recent scientific knowledge. Consistent with earlier findings based on patent citations (e.g. Breschi and Lissoni 2009), knowledge flows primarily appear to be based on direct collaboration and/or labor mobility. As regards different types of publication outlets, we found that proceedings of conferences and symposia are cited more rapidly than journal articles, indicating the importance of these "temporary clusters" (Maskell et al. 2006) of experts in the diffusion of scientific knowledge.

Provided that these findings generalize beyond our specific empirical setting, policy measures targeting the mere co-location of public research and innovative firms do not seem to be focused enough. Instead, direct interaction and mobility across organizational contexts are more promising objects of policy initiatives. Such policies have indeed been enacted in a number of jurisdictions, and university leaderships have also provided stronger incentives to their researchers to engage in private-sector collaboration in recent years. At the same time, the finding that localized knowledge flows are mostly not due to substantial unmediated knowledge spillovers hints at a more circumscribed role of policy interventions (cf. Breschi and Lissoni

2001). To the extent that the involved parties are able to appropriate the returns to their knowledge, it is not clear whether, and to what extent, their interaction needs to be induced by policy support. In addition, the relevance of individual channels of knowledge transfer and their interplay warrant closer attention. As we noted above, science-industry interaction has a long tradition in polymer technology, where it anteceded most of the measures of present-day innovation policy. Hiring recent graduates was a key channel of knowledge transfer in the 19th century. It is still widespread today but has received much less policy attention than other transfer channels, e.g., university patents and licensing. Conferences and symposia, suggested above as another relevant channel of knowledge flows, have likewise been all but neglected in recent discussions of science-industry interaction. More research is required whether the present focus on specific transfer channels, at the expense of others, is efficient.

Given that our study focused on a single technological context at one geographic scale and only one country, we are hesitant to derive far-reaching policy implications from our results. It is conceivable that the polymer industry with its long history of university-industry interaction and inter-sectoral labor mobility differs from other innovative industries in aspects that are relevant for the issues discussed in this article. It could also be that specificities of the Dutch innovation system drive our results. Further work would be required to rule out these possibilities, but our results resonate with those of related work by others. They add to the accumulating body of evidence from a variety of empirical contexts and study designs suggesting that less knowledge may be “in the air” than Alfred Marshall conjectured almost 100 years ago.

Acknowledgement

We would like to thank the Foundation for the History of Technology (Stichting Historie der Techniek, SHT) for providing data for this project. The dataset was originally created for a project on the history of Dutch polymer science and financed by the Dutch Polymer Institute (DPI). We also thank Stefan Krabel for helpful feedback on an early draft.

References

Akers, Noël J. 2000. "The Referencing of Prior Art Documents in European Patents and Applications." World Patent Information 22 (4): 309–315.

Alcacer, Juan, and Michelle Gittelman. 2006. "Patent Citations as a Measure of Knowledge Flows: The Influence of Examiner Citations." Review of Economics and Statistics 88 (4): 774–779.

Azoulay, Pierre, Waverly Ding, and Toby Stuart. 2009. "The Impact of Academic Patenting on the Rate, Quality and Direction of (public) Research Output." Journal of Industrial Economics 57 (4): 637-676.

Branstetter, Lee. 2005. "Exploring the Link Between Academic Science and Industrial Innovation." Annals of Economics and Statistics / Annalesd'Économieet de Statistique 79/80: 119–142.

Breschi, Stefano, and Francesco Lissoni. 2001. "Knowledge Spillovers and Local Innovation Systems: A Critical Survey." Industrial and Corporate Change 10 (4): 975–1005.

Breschi, Stefano, and Francesco Lissoni. 2009. "Mobility of skilled workers and co-invention networks: an anataomy of localized knowledge flows." Journal of Economic Geography 9: 439-468.

Callaert, Julie, Bart Van Looy, Arnold Verbeek, Koenraad Debackere, and Bart Thijs. 2006. "Traces of Prior Art: An Analysis of Non-patent References Found in Patent Documents." Scientometrics 69 (1): 3–20.

Collins, Peter, and Suzanne Wyatt. 1988. "Citations in Patents to the Basic Research Literature." Research Policy 17 (2): 65–74.

Criscuolo, Paola, and Bart Verspagen. 2008. "Does It Matter Where Patent Citations Come from? Inventor Vs. Examiner Citations in European Patents." Research Policy 37 (10): 1892–1908.

Feldman, Maryann P. 1999. "The New Economics Of Innovation, Spillovers And Agglomeration: A review Of Empirical Studies." Economics of Innovation and New Technology 8 (1-2): 5–25.

Fritsch, Michael, and Viktor Slavtchev. 2007. "Universities and innovation in space." Industry and Innovation 14: 201-218.

Geroski, Paul A. 1995. "Markets for Technology: Knowledge, Innovation and Appropriability." In Handbook of the Economics of Innovation and Technological Change, edited by Paul Stoneman, 90–131. Oxford: Blackwell.

Grupp, Hariolf, and Ulrich Schmoch. 1992. "Perceptions of Scientification of Innovation as Measured by Referencing Between Patents and Papers: Dynamics in Science-Based Fields of Technology." In Dynamics of Science-Based Innovation, edited by Grupp, Hariolf. 73–128. Berlin and Heidelberg: Springer.

Henderson, Rebecca, Adam B. Jaffe, and Manuel Trajtenberg. 1998. "Universities as a Source of Commercial Technology: A Detailed Analysis of University Patenting, 1965-1988." Review of Economics and Statistics 80 (1): 119–127.

Jaffe, Adam B. 1989. "Real effects of academic research." American Economic Review 79 (5): 957-970.

Jaffe, Adam B., Manuel Trajtenberg, and Rebecca Henderson. 1993. "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations." The Quarterly Journal of Economics 108 (3): 577–598.

Jaffe, Adam B., and Manuel Trajtenberg. 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 93 (23): 12671–12677.

Jaffe, Adam B., and Manuel Trajtenberg. 1999. “International Knowledge Flows: Evidence From Patent Citations.” Economics of Innovation and New Technology 8 (1-2): 105–36.

Jenkins, Stephen P. 1995. “Easy Estimation Methods for Discrete-Time Duration Models.” Oxford Bulletin of Economics and Statistics 57 (1).

Jenkins, Stephen P. 2005. Survival Analysis. Unpublished manuscript, Institute for Social and Economic Research, University of Essex, Colchester, UK.

Marshall, Alfred. 1920. Principles of Economics. 8th ed. London.

Maskell, Peter, Harald Bathelt, and Anders Malmberg. 2006. “Building global knowledge pipelines: the role of temporary clusters.” European Planning Studies 14 (8): 997-1013.

Meyer, Martin. 2000. “Patent Citations in a Novel Field of Technology — What Can They Tell About Interactions Between Emerging Communities of Science and Technology?” Scientometrics 48 (2): 151–178.

Murmann, J. Peter. 2003. Knowledge and Competitive Advantage: The Coevolution of Firms, Technology, and National Institutions. New York: Cambridge University Press.

Murmann, J. Peter. 2013. “The co-development of industrial sectors and academic disciplines.” Science and Public Policy 40: 229–246

Murray, Fiona, and Scott Stern. 2007. “Do Formal Intellectual Property Rights Hinder the Free Flow of Scientific Knowledge?: An Empirical Test of the Anti-commons Hypothesis.” Journal of Economic Behavior & Organization 63 (4): 648-687.

Narin, Francis, Kimberly S. Hamilton, and Dominic Olivastro. 1997. “The Increasing Linkage Between U.S. Technology and Public Science.” Research Policy 26 (3): 317–330.

Schmoch, Ulrich. 1993. “Tracing the Knowledge Transfer from Science to Technology as Reflected in Patent Indicators.” Scientometrics 26 (1): 193–211.

Roach, Michael, and Wesley M. Cohen. 2013. “Lens or prism? Patent citations as a measure of knowledge flows from public research.” Management Science 59 (2): 504-525.

Schmoch, Ulrich. 2008. “Concept of a Technology Classification for Country Comparisons.” Final Report to the World Intellectual Property Organization (WIPO), Fraunhofer Institute for Systems and Innovation Research, Karlsruhe.

Stephan, Paula E. 1996. “The Economics of Science.” Journal of Economic Literature 34 (3): 1199–1235.

Stern, Scott. 2004. “Do scientists pay to be scientists?” Management Science 50 (6): 835-853.

Thompson, Peter, and Melanie Fox-Kean. 2005. “Patent Citations and the Geography of Knowledge Spillovers: A Reassessment.” American Economic Review 95 (1): 450-460.

Tijssen, Robert JW. 2002. “Science Dependence of Technologies: Evidence from Inventions and Their Inventors.” Research Policy 31 (4): 509–526.

Torre, André. 2008. “On the Role Played by Temporary Geographical Proximity in Knowledge Transmission.” Regional Studies 42 (6): 869–889.

Van Vianen, B. G., H. F. Moed, and A. F. J. Van Raan. 1990. “An Exploration of the Science Base of Recent Technology.” Research Policy 19 (1): 61–81.

Verbeek, Arnold, Koenraad Debackere, and Marc Luwel. 2003. “Science Cited in Patents: A Geographic‘ Flow’ Analysis of Bibliographic Citation Patterns in Patents.” Scientometrics 58 (2): 241–263.

Walsh, Vivien. 1984. “Invention and Innovation in the Chemical Industry: Demand-pull or Discovery-push?” Research Policy 13 (4): 211–234.

Appendix

	affiliation BE	affiliation DE	affiliation UK	affiliation US	affiliation university	affiliation company	affiliation research institute	affiliation NL university	affiliation NL company	affiliation non-NL university	affiliation non-NL company	no affiliation mentioned	journal	handbook	symposia	company bulletin	other kind of publication	same author inventor	same applicant affiliation	chemical abstract	number of patent citations	number of NPL cited	patent filing year	patent office EP	patent office NL	university patent	TNO patent
affiliation NL	0.11	-0.05	0.00	-0.17	0.12	0.12	-0.04	0.75	0.65	-0.19	-0.14	-0.11	0.06	-0.05	0.02	-0.05	-0.02	0.51	0.56	-0.06	-0.04	-0.04	0.03	0.02	0.13	0.20	0.05
affiliation BE		-0.03	-0.01	-0.06	0.05	0.01	-0.03	0.13	0.01	0.00	0.01	-0.03	0.05	-0.03	-0.02	-0.02	-0.02	0.15	0.02	-0.03	-0.05	-0.02	0.05	0.02	0.03	0.14	-0.02
affiliation DE			-0.04	-0.18	0.05	0.02	0.01	-0.09	0.02	0.08	0.00	-0.03	-0.09	0.17	-0.05	0.00	-0.01	-0.08	-0.04	-0.03	-0.07	-0.03	-0.05	0.08	0.07	-0.06	0.01
affiliation UK				-0.14	0.06	0.05	-0.03	-0.06	0.06	0.09	0.01	-0.04	0.00	0.04	-0.06	0.03	-0.01	0.05	0.00	-0.02	-0.07	0.03	-0.01	0.04	0.00	0.00	-0.03
affiliation US					0.05	0.22	-0.04	-0.16	-0.07	0.12	0.28	-0.21	-0.09	0.09	0.10	0.00	-0.05	-0.09	-0.09	-0.09	0.12	0.03	-0.02	-0.01	0.02	-0.03	0.01
affiliation university						-0.36	-0.11	0.20	0.01	0.91	-0.33	-0.52	0.31	-0.10	-0.16	-0.20	-0.13	0.11	0.01	0.02	-0.23	0.02	0.17	0.17	0.22	0.15	0.01
affiliation company							-0.16	-0.04	0.30	-0.34	0.86	-0.30	-0.14	0.00	0.11	0.27	-0.08	0.04	0.21	-0.06	0.17	0.03	-0.01	-0.05	-0.03	-0.04	-0.06
affiliation research institute								-0.01	-0.07	-0.11	-0.12	-0.18	0.09	-0.03	-0.01	-0.07	-0.06	-0.01	0.01	-0.02	-0.04	-0.01	0.01	-0.04	0.00	0.00	0.09
affiliation NL university									0.15	-0.21	-0.10	-0.10	0.07	-0.04	-0.04	-0.04	0.01	0.49	0.22	-0.04	-0.09	-0.03	0.00	-0.01	0.11	0.29	0.07
affiliation NL company										-0.05	-0.10	-0.09	0.05	-0.03	0.01	-0.03	-0.04	0.27	0.74	-0.05	-0.02	-0.03	0.07	0.04	0.13	-0.03	-0.02
affiliation non-NL university											-0.29	-0.47	0.28	-0.08	-0.14	-0.18	-0.13	-0.09	-0.08	0.04	-0.19	0.03	0.17	0.17	0.18	0.03	-0.01
affiliation non-NL company												-0.26	-0.07	0.03	0.12	0.05	-0.06	-0.06	-0.09	-0.03	0.17	0.05	-0.02	-0.06	-0.07	-0.02	-0.05
no affiliation mentioned													-0.26	0.18	0.08	-0.06	0.24	-0.11	-0.08	0.00	0.10	-0.05	-0.16	-0.10	-0.16	-0.10	-0.01
journal														-0.57	-0.41	-0.36	-0.43	0.06	0.02	0.10	-0.29	0.06	0.14	0.06	0.19	0.12	0.02
handbook															-0.08	-0.07	-0.08	-0.07	-0.01	-0.09	0.00	-0.10	-0.17	0.11	-0.07	-0.08	0.02
symposia																-0.05	-0.06	0.05	0.02	-0.01	0.44	0.12	0.07	-0.12	-0.08	-0.05	-0.04
company bulletin																	-0.05	-0.04	-0.03	-0.05	0.07	-0.02	-0.06	-0.04	-0.10	-0.04	-0.02
other kind of publication																		-0.04	-0.02	-0.03	0.06	-0.08	-0.05	-0.11	-0.11	-0.03	0.00
same author inventor																			0.38	-0.04	-0.06	-0.03	0.03	0.02	0.08	0.19	0.06
same applicant affiliation																				-0.04	-0.01	-0.03	0.09	0.03	0.09	-0.04	0.09
chemical abstract																					0.02	0.37	0.10	0.09	0.18	-0.05	-0.01
patent filing year																						0.36	0.08	-0.30	-0.13	-0.14	-0.05
number of patent citations																							0.22	0.00	0.24	0.10	0.00
number of NPL cited																								0.14	0.26	0.06	0.04
patent office EP																									-0.02	0.10	-0.05
patent office NL																										0.13	0.08
university patent																											-0.05

all correlations r >= |0.06| on p < 0.01 significant

Table A1: Pairwise correlations between variables

	model A1			model A2		
(Intercept)	29.9978	*(10.6348)*	***	27.8279	*(10.8399)*	**
log(base hazard)	0.3102	*(0.0355)*	***	0.2964	*(0.036)*	***
affiliation NL	0.1408	*(0.1348)*				
affiliation university	0.2069	*(0.0631)*	***			
affiliation research institute	0.0788	*(0.0713)*				
affiliation NL univ. / inst. (only self cit.)				1.3255	*(0.2655)*	***
affiliation NL univ. / inst. (excl. self cit)				0.2731	*(0.2258)*	
affiliation NL company (only self cit.)				0.3116	*(0.1993)*	
affiliation NL company (excl. self cit.)				-0.1763	*(0.2255)*	
affiliation non-NL univ. / inst.				0.1658	*(0.069)*	**
affiliation BE	-0.1931	*(0.4075)*		-0.1291	*(0.4656)*	
affiliation DE	-0.0665	*(0.0745)*		-0.0472	*(0.0789)*	
affiliation UK	-0.1415	*(0.1016)*		-0.0592	*(0.1054)*	
affiliation US	-0.2247	*(0.053)*	***	-0.1971	*(0.0524)*	***
handbook	-0.3012	*(0.0703)*	***	-0.3337	*(0.0688)*	***
symposia	0.2898	*(0.1111)*	***	0.3614	*(0.1068)*	***
company bulletin	0.2713	*(0.1654)*		0.2362	*(0.1644)*	
no affiliation mentioned	-0.2158	*(0.0696)*	***	-0.2312	*(0.0701)*	***
chemical abstract	-0.2573	*(0.1046)*	**	-0.2736	*(0.107)*	**
other kind of publication	0.1859	*(0.1109)*	*	0.1445	*(0.1115)*	
number of patent citations	-0.0005	*(0.003)*		-0.0003	*(0.0029)*	
patent office NL	0.0016	*(0.0894)*		-0.0039	*(0.0898)*	
patent office EP	0.0684	*(0.0636)*		0.0948	*(0.0634)*	
same author inventor	0.9251	*(0.1441)*	***			
same applicant affiliation	0.0954	*(0.2131)*				
IPC dummies	TRUE			TRUE		
n	2006			2006		
logLik	-6532.7570			-6551.1974		
(p > chi2)	0.0000			0.0000		

*: $p < 0.1$ regression coefficient
**: $p < 0.05$ (standard errors in brackets)
***: $p < 0.01$

Table A2: Estimation of the citation lag (dummies for self-citations included)