Science and Public Policy, 2017, 1–13
doi: 10.1093/scipol/scx001
Article

Inventor mobility and the geography of knowledge flows: evidence from the US biopharmaceutical industry

Zafer Sonmez*

College of Urban Planning and Public Affairs, University of Illinois-Chicago, 412 South Peoria, Suite 400, Chicago, IL 60607, USA

*Corresponding author. Email: zsonme2@uic.edu

Abstract

This article investigates the role of labor mobility and geographical proximity in the knowledge diffusion process in the US biopharmaceutical industry. The application of social network analysis to patent authorship reveals that labor mobility and co-inventorship are responsible for a large portion of knowledge flows. This finding provides support for recent studies that called into question the notion that technical and commercially valuable knowledge ubiquitously disseminates in high-technology industrial agglomerations, indicating instead that such an explanation is only partially true. Results also suggest that high quality inventions draw (proportionally) more from nonlocal knowledge sources and that network connections are more important for the transmission of knowledge for high quality patents than for low quality patents. The substantial concentration of local knowledge flows suggests that industrially targeted public financial support for research and development activities at the regional and state levels can be considered as supportive of firm performance and by extension economic development.

Key words: knowledge flows; spatial proximity; inventor mobility; social network analysis; US biopharmaceutical industry.

1. Introduction

Understanding the learning and innovation enhancing properties of industrial agglomerations has been subject to a large number of empirical analyses in the last two decades or so. To explain the forms and spatial dimensions of the organizational and individual learning occurring in regional industrial agglomerations, studies in the extant empirical literature focused on (1) the spatial extent of knowledge flows (Jaffe et al. 1993; Peri 2005; Breschi and Lissoni 2009; Singh and Marx 2013); (2) the importance of weak ties versus strong ties (Lissoni 2001; Fritsch and Kauffeld-Monz 2010); (3) the importance of informal ties versus formal ties (Agrawal et al. 2006; Steiner and Hartmann 2006; Breschi and Lissoni 2009; Jenkins and Tallman 2010); and (4) the importance of local ties versus extra-local ties (Owen-Smith and Powell 2004; Gertler and Levitte 2005; Gallié 2009). Overall, by considering the spatial, social, organizational, and knowledge-based elements more explicitly than previous studies, these studies provide a more coherent picture of the knowledge production and diffusion processes. A quick methodological survey of the current empirical literature reveals that theoretical insights and analytical approaches introduced by the critical assessments of dominant concepts (e.g. proximity) or their reinterpretations (embeddedness and cluster) have been very instrumental in the undertaking of most of these empirical studies (Malmberg and Maskell 2002, 2006; Markusen 2003; Hess 2004; Bathelt et al. 2004; Boschma 2005). For example, Boschma's call for analytically isolating the effects of geographical proximity from other forms of proximity (social, organizational, cognitive, and institutional) to determine whether spatial closeness really matters in the process of knowledge production and flows has been instrumental in the growth of studies that found that spatial proximity is neither a necessary nor a sufficient condition, but a facilitator (e.g. Breschi and Lissoni 2009; Gallié 2009).

In the context of knowledge flows, more recent conceptual and empirical contributions highlight the interrelatedness of proximities and coevolution of proximity dimensions (Balland et al. 2015; Broekel 2015); balance between local and global linkages in accessing knowledge (Bathelt and Cohendet 2014; Crespo and Vicente 2016); and the role and changing importance over time of the impact of geographical distance and other proximities in the knowledge diffusion process (Cappelli and Montobbio 2016; Breschi and Lissoni 2009; Sonn and Storper 2008). The present investigation is considered to be part of this latter type of research efforts. Examining the temporal evolution of the impact of geographical distance in the case of biopharmaceutical industry is expected to be

© The Author 2017. Published by Oxford University Press. All rights reserved. For Permissions, please email: journals.permissions@oup.com

Downloaded from https://academic.oup.com/spp/article-abstract/doi/10.1093/scipol/scx001/2964537/Inventor-mobility-and-the-geography-of-knowledge
by lib-electronic@uic.edu user
on 02 October 2017

particularly insightful because knowledge production and sharing plays a relatively important role in this industry (Dimasi and Grabowski 2007).

Studies quantifying the extent to which knowledge flows are spatially concentrated have brought substantial precision to our understanding of the effects of spatial proximity on knowledge spillovers, which were a very strong assumption as a subset of agglomeration economies for a long time. Today there is a general agreement on the localized nature of knowledge flows (Feldman and Kogler 2010; Singh and Marx 2013). However, studies have yet to produce a consensus on how this localization occurs. The identification of specific mechanisms by which commercially and technologically valuable knowledge flows across organizational boundaries is particularly important for regional economic development policy making, given that an increasing number of studies demonstrate a positive relationship between knowledge flows and innovation (Aharonson et al. 2004; Miguelez and Moreno 2014). Moreover, although previous studies analyzing the effect of inventor mobility and inventor networks on knowledge flows generally find a positive influence, it is not clear how (if any) this positive impact varies depending on the quality of inventions. This gap is particularly applicable to those studies that use patent statistics because there is great amount of variation in the quality of patents (Nagaoka et al. 2010).

The main purpose of our investigation is to sort out nonmarket-based social ties from market-based channels in explaining localized knowledge spillovers. Following Breschi and Lissoni (2009) (hereafter, BL), we refer to nonmarket-based social ties as those that are a product of repeated meetings made possible by spatial proximity (e.g. those developed by local social events). And we consider market-based channels of knowledge flows as an outcome of the labor market (Almeida and Kogut 1999) and formal collaborative networking (Mowery and Ziedonis 2004). The first type of these connections are informal channels of knowledge flows in the form of knowledge spillovers (pure externalities) while the second type of connections (formal) functions as a qualitatively different channel of knowledge flows, the extent of which would be limited to exchanging parties and consequently resulting in limited pure externalities for nonparticipating parties. We operationalize formal collaborative networking through co-inventorship ties and measure it by applying social network methods to historical patent data.

The application of social network analysis in examining the spatial extent of knowledge flows is an example of one way to theorize interrelations between organizations, space and knowledge production. In such an analysis, the network is used as a mediator of geographical effects on knowledge flows (Gluckler 2007; Bouba-Olga et al. 2015). Going beyond the conceptualization of knowledge production and dissemination activities as locally situated, with this approach the analyst is able to explore the spatial implications of such conceptualization. By precisely categorizing and examining all possible relations between the origin and the destination of knowledge flows, he/she avoids taking the logical short cut of assuming that tacitness necessarily implies the localization of knowledge flows, as had been often the case in the literature (Boschma and Frenken 2006; Breschi and Lissoni 2001 in Breschi et al. 2010).

This article addresses four questions. First, w*hat is the extent of the localization of knowledge flows? Second, has there been a change in the extent of the localization of knowledge flows over time?* These questions are investigated through an experimental design in which the geographic location of citations (citing patents) is compared with that of their originating (cited) patents, while controlling for the concentration of industrial activity and the progress of technology in the industry (Jaffe et *al.* 1993; Sonn and Storper 2008). Third, *what is the role of inventors' mobility across firms and in space in the geography of knowledge flows?* Four, *does the importance of geography, mobility and networks vary in the knowledge diffusion process depending on the quality of inventions (patents)?* Answering these questions necessitates a more specific estimation procedure. In this procedure, subsequent patent citations are considered as a function of labor mobility[1] and social ties established through previous co-inventorships (Singh 2005; Breschi and Lissoni 2009).

This study contributes to the empirical literature on geography and knowledge flows by first showing that high quality inventions draw (proportionally) more from nonlocal knowledge sources and that network connections are more important for the transmission of knowledge for high quality patents than for low quality patents. Second, it demonstrates that the importance of spatial distance remained relatively stable over time in the diffusion of knowledge. While there has been a slight increase in the localized knowledge flows over time at the regional level, no significant change is observed at the state level. In addition, it verifies previous results in a different time frame. This study further clarifies the role of geography in the knowledge diffusion process by considering the effects of geography and networks simultaneously. It suggests that distance impedes knowledge transmission because geographical co-location (physical proximity) largely shapes inventor networks and these networks in turn drive diffusion of knowledge. Finally, this study contributes to a better understanding of how different degrees of network connections among inventors influence knowledge flows between organizations. Social ties established through inventor co-authorship are found to transmit knowledge most often at close distances. That is, the frequency of knowledge flows steadily declines with increased (social) distance between inventors. This reinforces the existing empirical knowledge about the interrelatedness of proximities—geographical and social (Broekel 2015; Breschi and Lissoni 2009). Overall, these empirical results have important implications for science and technology policies aimed diffusing innovation and R&D outcomes.

The rest of the article is organized as follows. The second section presents the case industry. The third section discusses the sample construction and the localization test. The fourth section presents the analyses and results. A detailed comparison with previous results is made in these sections. The final section concludes with some policy implications and directions for future research.

2. Case: biopharmaceutical industry

There are strong analytical and practical reasons for focusing on the biopharmaceutical industry. First, it is one of the most research and knowledge intensive sectors, measured in both industrial R&D spending and its reliance on academic scientific advancement for growth (Cooke 2007; Dimasi and Grabowski 2007). Second, the emergence of biotechnology as a new paradigm for drug discovery has been revolutionizing the pharmaceutical industry since 1970s by redefining the fundamentals of firm location choices, the relationships of firms with one another, and the internal organization of firms (Powell and Sandholtz 2012). Third, in the case of the USA, at the local level policy-makers view the biotechnology sector as a cornerstone of knowledge-based economic growth. And, at the national level it is regarded as one of the industries that will compensate American firms' loss of competitiveness in more traditional

Downloaded from https://academic.oup.com/spp/article-abstract/doi/10.1093/scipol/scx001/2964537/Inventor-mobility-and-the-geography-of-knowledge
by lib-electronic@uic.edu user
on 02 October 2017

sectors to Asian and European rivals. This, at least partially, explains the legal reforms initiated by the US Congress to create a favorable institutional environment for biotechnology in the early 1980s and the fact that the National Institutes of Health (NIH) has been the largest financial supporter of biomedical research in the world. Finally, the selection of this industry was driven by a crucial analytical necessity: the validity of empirical findings. Being a highly regulated industry, organizations in this industry largely rely on patenting activity for the protection of intellectual property. Therefore, patent statistics should reflect the invention activities in this industry fairly well (Fornahl et al. 2011). These data are available to researchers as a nearly complete historical record.[2] The relatively large extent of the traceability of knowledge production activities in this industry enhances the internal validity of research findings and the applicability of the conclusions to the issues observed in the industry in general.

3. Sample construction and localization test

The initial sample included all patent applications at USPTO from 2000 to 2010. From this sample, patents that report at least one US inventor and were assigned to a US organization in one of either drug, bio-affecting and body treating compositions (USPTO Technology Class 424) or chemistry: molecular biology and microbiology (USPTO Technology Class 435) were selected. Such a sampling generated 23,120 patent applications, which have been signed by 30,509 different inventors and have been assigned to 3,546 different organizations. Finally, since the validity of the analysis crucially depends on the correct identification of individual inventors, assignees and their locations, a thorough work of finding unique inventors' and cleaning organizations' names and addresses was carried out. USPTO does not require consistent and unique identifiers for inventors, and thus researchers face formidable obstacles in examining inventor career paths and geographies. For the identification of unique inventors, Harvard University's Patent Network Dataverse was utilized.[3] When organization names were available but not their IDs, organization IDs were identified by searching the database for other patents assigned to the same organization. In cases where there was no existing organization ID information in the database, new organization IDs were assigned. Among these, there were eighty-nine instances where both the organization's name and ID were missing. These patents were dropped from the sample because it was not possible to assign a new organization to such cases in a reliable way. In the final sample, there were 24,423 patent-assignee instances and 67,674 patent-inventor instances.

To measure the extent of localized knowledge flows in the form of patent citations, a test developed by Jaffe et al. (1993) (hereafter, JTH) was employed (see Agrawal et al. 2006 and BL, for recent applications). This test can be described as a quasi-experiment in which the geographic location of citations (citing patents) is compared with that of originating (cited) patents that they cite, while controlling for the concentration of industrial activity and the progress of technology in the industry.

As a first step in the test, from the sample of 23,120 patents, three cohorts of originating patents, consisting respectively of the 2000, 2001, and 2002 patent applications that received at least one citation by the end of 2010 were selected. Those were designated as the cited patents. In the next step, patents that referenced one of the cited patents were selected. Those were designated as the citing patents. All observations in which the citing and cited patents have been assigned to the same organization were found and designated as self-citations (later eliminated from the test). There were 4,722 cited patents versus 6,261 citing patents. The number of citing patents is greater than the number of cited patents as cited patents may receive more than one citation. There were 14,895 citing-cited patent pairs, of which 5,760 were self-citations. As a final step, a control sample was constructed. For each citing patent, a control patent with approximately the same application date and exactly the same technological class of the citing one was selected. In all, 73 percent of the citing patents were matched on the basis of the same day/month/year while 27 percent of the citing patents was matched on the basis of the same month and year. The control sample 'mimics the distribution of patent production and thereby is used as a benchmark against which one can evaluate the fraction of spatially co-located citing-cited patents' (BL 441).

Another major issue in measuring the geographic dispersion of patents and patent citations is the way patents are assigned to locations (Breschi et al. 2010). In the case of multiple inventors, following BL, each citing-cited patent pair was designated as 'matching geographically if they have in common at least one location among those reported in the inventors' addresses (BL: 452). Similar logic applies to the designation of self-citations: in the case of multiple assignees, a citing-cited patent pair was designated as a self-cite if the pair has at least one assignee organization in common. This logic is in line with BL and differs from the approach taken by JTH. In the case of multiple inventors, JTH assign each patent to the region/state in which pluralities of inventors reside, and then compared the assigned locations across patent pairs. The decision to choose a relatively low threshold for the geographical matching is driven by a concern of capturing the full extent (impact) of geographical co-location on the citation probability. The JTH approach implicitly treats each potential tie between inventors of cited and citing patent as qualitatively same by simply limiting the co-location to the majority of authors. However, it is highly probable that fewer connections between inventors of patent pairs that JTH consider as geographically nonmatching are socially stronger than the connection between inventors of patent pairs that they consider as geographically matching. Therefore, the approach in this study better captures the differentiated impact of space on the formation and strength of social ties among inventors.

The geographic locations of citing patents were compared with the geography of patents that they cite (on the basis of the location of inventors). First, the fraction of spatially co-located citing-cited patents, called the citation-match frequency, was calculated. The same procedure was applied to the control-cited pair: the geographic location of the control patents was compared with the geography of cited patents and then the fraction of spatially co-located control-cited patents was calculated. The final step in the analysis was simply to compare the two fractions. If the first fraction (citation-match frequency) is greater than the second fraction (control match frequency), this implies that citations are localized after controlling for (1) timing and (2) technology. The estimated fractions of all citing-cited and control-cited patent pairs that match geographically are presented in Table 1. The fourth column of the tables reports the percentage of citing patents that are co-located with the cited ones; the sixth column reports the same percentage for the control sample. Finally, the last column reports the odds ratios (see Appendix for z-test).

These results confirm the findings of previous studies (BL; Sonn and Storper 2008; Singh and Marx 2013). Citing patents are significantly more likely than control patents to come from the same

Downloaded from https://academic.oup.com/spp/article-abstract/doi/10.1093/scipol/scx001/2964537/Inventor-mobility-and-the-geography-of-knowledge
by lib-electronic@uic.edu user
on 02 October 2017

Table 1. Geographical matching at the state and regional level (without self cites, 2000s cohort).

	Number of observations	Citing-cited		Control-cited		*z*-test (P > *z*)	Odds ratio
		Frequency	%	Frequency	%		
State	9,135	2,696	29.51	1,607	17.59	18.99 (0.000)	1.96
Region	9,135	2,030	22.22	1,017	11.13	20.10 (0.000)	2.28

Table 2. Geographical matching at the state and regional level (without self cites, 1990s cohort).

	Number of observations	Citing-cited		Control-cited		*z*-test (P > *z*)	Odds ratio
		Frequency	%	Frequency	%		
State	3,411	799	23.42	456	13.37	10.72 (0.000)	1.98
Region	3,411	680	19.94	365	10.70	10.59 (0.000)	2.08

geographical area of cited patents. Excluding self-citations, at the state level, 29 percent of citations are localized, compared to 17 percent of controls; similarly, at the regional level,[4] 22 percent of citations are localized, compared to 11 percent of controls. The odds ratio indicates that citing patents are 96 percent more likely than controls to come from the same state as the cited patents; the same value at the regional level is more than 128 percent. The greater odds ratio for the regional level indicates that knowledge flows are more intense at this level than they are at the state level.

Has there been a change in the extent of the localization of knowledge flows over time? The period of observation in Table 1 is the 2000s. There might be an overall increase in the number of innovating actors locally due to outsourcing of innovation from big companies to small firms or individual inventors and an increasing number and importance of spin-offs in the innovation process (Sonn and Storper 2008). Note that the former does not necessarily mean increased concentration of industry, which is controlled for in the test and cannot be responsible for the observed differences. Holding the employment and patent production in the industry constant, the higher the number of innovating actors, the higher the probability that citing and cited invention pairs would be co-located and higher localized knowledge flows would be observed. To check if there has been an increase in localized knowledge flows over time, we repeated the geographical matching analysis using patents applied for in the 1990s (Table 2). Comparison of results reported in Tables 1 and 2 indicates that while there has been a slight increase in the localized knowledge flows over time (from 1990s to 2000s) at the regional level, there is no change at the state level.

3.1 Comparison with earlier results and implications

For these findings, one directly comparable study is the BL study of the US biopharmaceutical patents filed at EPO. The absolute geographic matching rates of the present study are higher than those reported in their study: approximately 41 percent more at the state level, and 27 percent at the regional level. The odds ratio at the state level in this study is 21 percent higher than the previous study and 31 percent higher at the regional level. There are a number of explanations for the differences observed between this study's findings and those of BL. First, the patent applications of US companies at EPO are probably high quality/impact patents as they have been deemed worthy of extending to Europe (through a costly procedure of international application at EPO) (BL). These patents may be selective on the basis of assignee, they are perhaps filed primarily by multinational US companies that have business operations in Europe or have some products designed specifically for European markets. Being different in impact, authors of these patents might have access to extra-regional knowledge sources that might not be available to authors of other patents.

To check the validity of this explanation, we tested whether geography matters more or less depending on the quality of (citing) patents. The citing-cited patent pairs were divided into two groups on the basis of whether or not the citing patent received a citation, and then the test of localization was repeated for both groups. The count of forward citations is a widely used indicator of inventive quality in terms of the technological or economic value of patents in the literature (Hall et al. 2005; Von Wartburg et al. 2005; Gambardella et al. 2008; Nagaoka et al. 2010; Petruzzelli et al. 2015). The quality of a patent is reliably approximated by the number of times it is cited subsequently because citations signify the importance of the invention as a foundation for subsequent patents.

Geographical matching frequencies and odds ratios reported in Table 3 confirm the proposition that authors of high quality patents have access to extra-regional knowledge sources and therefore are less likely to rely on local knowledge sources. The odds ratios indicate that both low and high quality citing patents are more likely than controls to come from the same state or region as the cited patents, but smaller odds ratios for high quality citing patents suggest that high quality inventions draw (proportionally) more from nonlocal knowledge sources. These results suggest that while geographical proximity is more important for low quality patents, it is less important for high quality patents in facilitating knowledge flows.

Another explanation for the difference between two studies is related to the citation rules at the respective patent offices. There are substantial differences between USPTO and EPO regulations and practices with regard to citations. In terms of referring to prior art, applicants at USPTO are required to list all inventions regardless of their degree of relevance to the patentability of the invention in question. This is called the 'duty of candor'. Because failing to comply with this criterion of full disclosure results in serious penalties, including the revocation of a granted patent in the future, applicants in the USA provide a long list of references. For applicants at EPO, there is no such rule and in fact applicants are asked to *limit* their references to prior inventions that are essential to assess the present invention's claim of novelty (Harhoff et al. 1999; Von Wartburg et al. 2005). Given this fundamental difference in citation

Downloaded from https://academic.oup.com/spp/article-abstract/doi/10.1093/scipol/scx001/2964537/Inventor-mobility-and-the-geography-of-knowledge
by lib-electronic@uic.edu user
on 02 October 2017

Table 3. Geographical matching based on the quality of citing patent.

	Number of observations	Citing-cited		Control-cited		z-test (P > z)	Odds ratio
		Frequency	%	Frequency	%		
Citing Patents with without forward citation (low quality citing patents)							
State	6,653	2,019	30.35	1,196	17.98	16.67 (0.007)	1.99
Region	6,653	1,546	23.24	753	11.32	18.18 (0.006)	2.37
Citing Patents with forward citation (high quality citing patents)							
State	2,482	677	27.28	411	16.56	9.13 (0.011)	1.89
Region	2,482	484	19.50	264	10.64	8.73 (0.010)	2.04

procedures, compared to USPTO, EPO patent citations are less numerous, and more directly relevant to prior art. This means that numerous citations in patents filed at the USPTO appear in the analysis to come disproportionally from the same region/state. It might be the case that co-location exposes some patents to high impact patents (cited) and leads them to cite those patents for reasons such as a desire to be associated with a particular area of technical development, invention, or discovery. A note by BL (25) supports both this and the first explanation:

> We must point out that the importance attached by our results to inventors' mobility and connectedness may suffer of a selection bias effect, due to our choice of using EPO patents (over U.S. inventions) instead of USPTO ones .... EPO patents do not include low-quality patents which are not worth extending to Europe. It is possible that low-quality patents come from relatively isolated inventors, who may nonetheless cite the originating patents when co-located with the latter's inventors, thanks to contacts different from the co-invention network (e.g. conferences and local press). As a result, by using EPO patents, we would end up underestimating the co-location proportion of citing-cited patents and also overestimating the importance of the co-invention network in explaining the geographical co-location of citing-cited patents, relatively to the control sample.

Before moving to the next section, it is important to assess the degree to which the localization results reported above might be driven by technological differences between citing and control patents (one of two factors controlled for in the test of localization). In a critical assessment of JTH's method, Thompson and Fox-Kean (2005) argue that selecting controls from three-digit technology classes is methodologically problematic and causes spurious evidence of localized knowledge spillovers. They claim that controls sampled at this level do not really serve as controls for the concentration of industrial activity due to the widely observed within-class heterogeneity at this level. Because of the lack of industrial relatedness between citing patents and controls in such cases, they argue, by extension there would likely be a significant degree of technological distance between cited and control patents as well. Therefore controls cannot be expected to make a citation to cited patents, as assumed in the Jaffe et al. experiment. To address this issue, in this study the set of control patents were resampled to force a narrower technological matching between citing and control patents. Specifically, control patents were resampled to match citing patents not only at the three-digit primary class level but also at the five-digit secondary class level. Then the percentage of co-located citing-cited and control-cited patents were calculated for the new samples. As seen in Table 4, the share of co-located control patents increases considerably while changes slightly for citing ones. Consequently, the difference between the two proportions narrows, as suggested by the smaller odds ratios. This result partially confirms the critique raised by Thompson and Fox-Kean; however, the statistically significant odds ratio still indicates the localization of citations.

4. Labor mobility, geographical proximity, and knowledge flows

Can the localized knowledge 'spillovers' discussed in the preceding section be product of individual chance encounters? That is, do citing inventors learn about new inventions and discoveries through informal encounters, gatherings, and the exposure to local press that are solely facilitated by the spatial proximity? The purpose of the analysis in this section is to quantify the extent to which inventor mobility and inventor social networks constructed through patent co-authorship function as a mechanism of knowledge flows between different organizations. It is argued that there is a social structural explanation for the observed outcomes (the hypothesis of nonrandomness of localized knowledge spillovers).

Labor mobility is one of four mechanisms of knowledge flows in regional industrial agglomerations.[5] This mechanism operates through the flexible labor market dominant in high technology industries. High job turnover is supported by spatial clustering as a critical mass of possible employers facilitates job searching and matching for engineers and scientists. When these highly skilled workers change jobs, they carry experience and social networks from previous employment to their new employers. Increasing labor mobility means a higher possibility of local knowledge flows through the direct movement of workers and the expansion of their social networks (Almeida and Kogut 1999; Sonn and Storper 2008). This suggests that knowledge flows in the form of the patent citations examined in the preceding section might be structured by such labor mobility and the resulting expansion of social networks. To uncover the role of these social links in facilitating knowledge flows, one needs to examine the extent to which the same localization patterns reported above (Table 1) could still be found once these links are controlled for. For this purpose, the method developed by BL was applied. The BL method essentially finds three types of connections between pair of patents from different organizations based on past co-inventorships: (1) mobility-linkages, (2) co-inventor network-linkages, and (3) no linkage.

Mobility pairs (type 1): Patents share at least one common inventor, who has moved from the organization of cited patent to the organization of citing patent, either as an employee or a consultant. The social distance between such patent pairs is zero, as patent pairs are personally connected.

Socially connected pairs (type 2): Among the authors of the citing patent, there is at least one individual that is socially connected

Downloaded from https://academic.oup.com/spp/article-abstract/doi/10.1093/scipol/scx001/2964537/Inventor-mobility-and-the-geography-of-knowledge
by lib-electronic@uic.edu user
on 02 October 2017

Table 4. Geographical matching at the state and regional level (without self cites).

	Number of observations	Citing - cited		Control-cited		z-test ($P > z$)	Odds ratio
		Frequency	%	Frequency	%		
State	1,739	532	30.59	379	21.79	5.90 (0.015)	1.58
Region	1,739	419	24.09	266	15.30	6.52 (0.013)	1.76

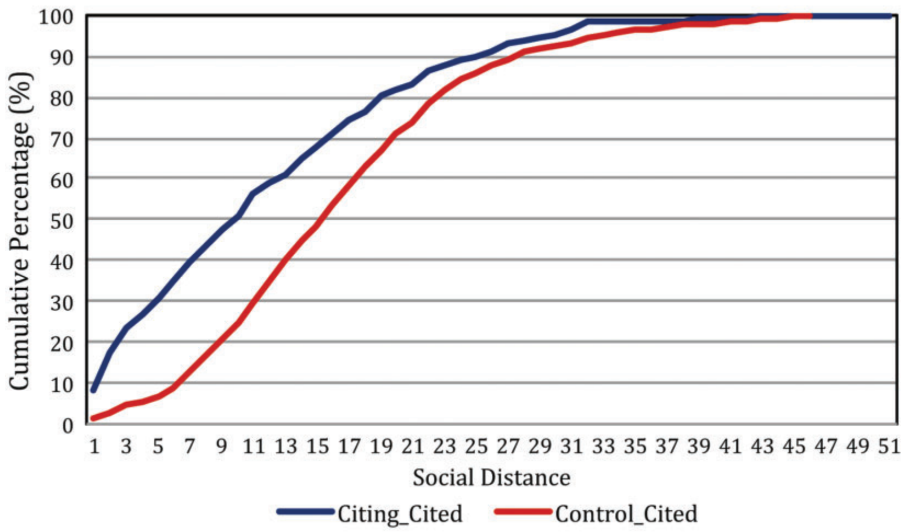

Figure 1. Distribution of social distances between connected patent pairs.

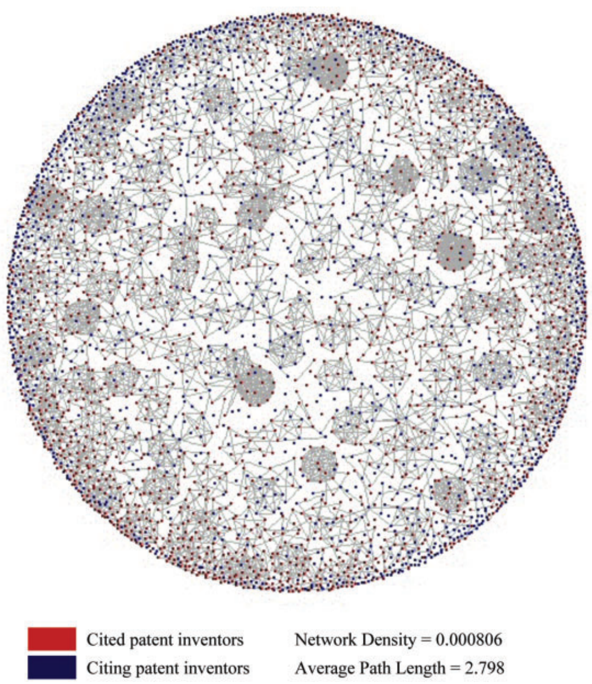

Figure 2. Co-invention network (citing and cited patent inventors).

to one of the authors of the cited patent. For instance, authors of citing and cited patent worked together on a different patent in the past. Or, one of the inventors listed on the citing patent knows somebody (a previous collaborator) who knows one of the inventors listed on the cited patent. In such cases, the distance between patent pairs is measured by geodesic distance. The minimum social distance between such patent pairs is 1 and the maximum extends to any positive discreet value.

Unconnected pairs (type 3): Among the authors of the citing patent, there is no individual that is socially (formally) connected to one of the authors of the cited patent. In other words, the inventors of the citing patent and the inventors of the cited patent belong to distinct network components. The social distance between such patent pairs equals infinity. Note, however, that it cannot be completely ruled out that some of these patent pairs are linked through some form of informal relations that exist among inventors not captured by patenting activity.

To examine the degree of social proximity among patent pairs connected via the co-inventor network in more detail, the geodesic distance between them was computed. Geodesic distance is defined as the minimum number of edges that separate two nodes in a network. For example, the geodesic distance between two inventors who have worked together is one. On the other hand, the geodesic distance between two inventors who have not worked before, but share a common collaborator is two. While the minimum value for geodesic distance is 1, the maximum extends to any positive discrete value. The results are reported in Table 5. It is shown that citing-cited patent pairs are considerably more likely to be connected by mobile inventors than control-cited pairs (8.9% versus 0.2%). It is also shown that citing-cited patent pairs are more likely to be linked via a co-invention network (10.7% versus 6.3%). In general, the citing-cited patent pairs present not only overall higher connectedness but also closer social proximity than the control-cited patent pairs (relatively

Downloaded from https://academic.oup.com/spp/article-abstract/doi/10.1093/scipol/scx001/2964537/Inventor-mobility-and-the-geography-of-knowledge
by lib-electronic@uic.edu user
on 02 October 2017

more control-cited pairs are connected at greater distances; see Fig. 1). As an example, we present network graphs for the inventors of one patent cohort in Figs 2 and 3. Year 2005 is selected because it represents the mid-point of the time period analyzed. In consistent with the patterns observed in Table 5 and Fig. 1, the citing-cited inventor network is denser than the cited-control inventor network and on average inventors in the former network are connected to each other at closer distances (see network density and average path length measures).[6]

These comparative figures suggest two things. First, mobile inventors are responsible for a substantial portion of knowledge flows between different organizations. Second, social ties established through co-authorship might function to transmit knowledge mostly at close distances (see Fig. 4). To exactly what extent these are so is the subject of the following analysis.

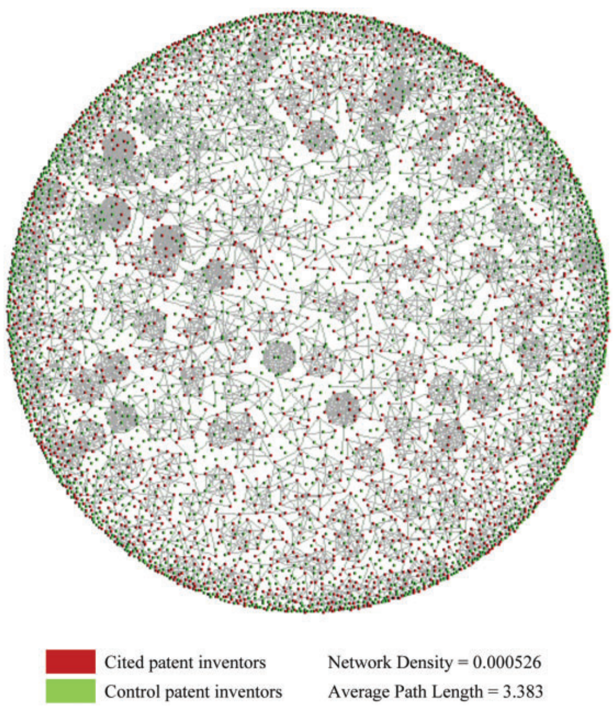

Figure 3. Co-invention network (control and cited patent inventors).

4.1 The re-analysis of localization patterns based on mobility trends and co-invention ties

To find out to what extent labor mobility and social ties are responsible for local knowledge flows, all patent pairs (citing-cited and the corresponding control-cited) linked by a specific mechanism were removed from the whole sample successively and the rate of geographical matching for both sets of pairs was recalculated (Table 6). The first row for both the state and regional panels presents matching rates for all patent pairs for comparison purposes (presented earlier in Table 1). The second row in both panels reports the localization rates and corresponding odds ratios after all patent pairs connected by mobile inventors were removed from the sample. The proportions of citing patents that are co-located with the cited ones decline substantially. At the state level, the share declines from 29.5 percent to 23.2 percent and at the regional level it declines from 22.2 percent to 15.3 percent. However, the corresponding decline for the control-cited pairs is almost negligible, much less than 1 per cent. As indicated by the z-tests and odds ratios, however, the differences in co-location between citing-cited and control-cited pairs are still statistically significant and substantial. Citing patents are still 43 percent more likely than the controls to come from the same state as the cited patents; the corresponding value at the

Table 5. Status of social connectedness for the patent pairs (%).

Connectivity	Citing-cited	Control-cited
Connected through mobility (social distance = 0)	8.92	0.22
Connected through co-invention network (distance ≥ 1)	10.70	6.36
Not Connected (distance = ∞)	80.38	93.42
Total	**100**	**100**

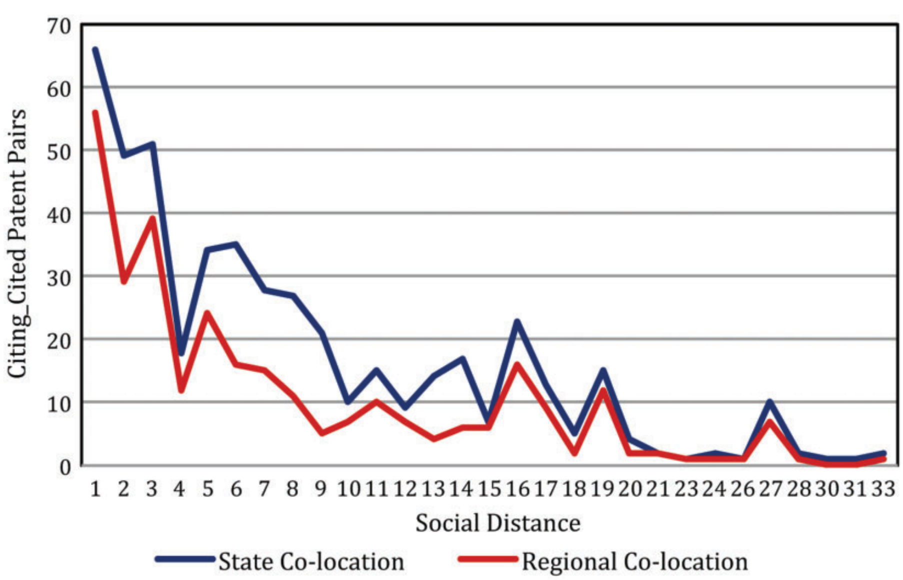

Figure 4. State and regional co-location by social distances.

Downloaded from https://academic.oup.com/spp/article-abstract/doi/10.1093/scipol/scx001/2964537/Inventor-mobility-and-the-geography-of-knowledge
by lib-electronic@uic.edu user
on 02 October 2017

Table 6. Social connections and geographical matching at the state and regional levels.

	N[a]	Citing-cited		Control-cited		z-test (P > z)	Odds ratio
		Frequency	%	Frequency	%		
State level							
All patent pairs	9,135	2,696	29.5	1,607	17.6	18.99 (0.001)	1.96
All pairs except those linked by mobility	8,381	1,944	23.2	1,458	17.4	9.33 (0.001)	1.43
All pairs except those linked by SN[b] distance ≤5	8,017	1,726	21.5	1,376	17.2	7.00 (0.001)	1.32
All pairs except those linked by SN distance ≤10	7,825	1,605	20.5	1,323	16.9	5.78 (0.001)	1.27
All pairs except those linked by SN distance ≤15	7,658	1,543	20.1	1,282	16.7	5.44 (0.001)	1.25
Only unconnected pairs	7,343	1,461	19.9	1,216	16.6	5.24 (0.001)	1.25
Regional level							
All patent pairs	9,135	2,030	22.2	1,017	11.1	20.10 (0.001)	2.28
All pairs except those linked by mobility	8,381	1,286	15.3	922	11.0	8.31 (0.001)	1.47
All pairs except those linked by SN distance ≤5	8,017	1,126	14.0	873	10.9	6.05 (0.001)	1.34
All pairs except those linked by SN distance ≤10	7,825	1,072	13.7	843	10.8	5.59 (0.001)	1.31
All pairs except those linked by SN distance ≤15	7,658	1,039	13.6	822	10.7	5.37 (0.001)	1.31
Only unconnected pairs	7,343	984	13.4	779	10.6	5.20 (0.001)	1.30

[a]N: number of observations.
[b]SN: Social network.

regional level is 47 percent. The big decline in the odds ratio after mobility linked patent pairs were removed from the sample suggests that the movement of inventors across organizations accounts for a large fraction of patent citations/knowledge flows in the biopharmaceuticals industry.

In the next step, in addition to mobility linked patent pairs, all patent pairs (both citing-cited and control-cited) linked by a co-invention network at a distance less than 6 were removed from the sample and the rate of geographical matching for both sets of pairs was recalculated. The third row in both panels of Table 6 reports the localization rates and the corresponding odds ratios after all patent pairs linked at distance of 5 or shorter were removed, in addition to those connected through mobility. The proportion of citing patents that are co-located with the cited patents declines again, but not as substantially. At the state level, it declines from 23.2 percent to 21.5 percent and at the regional level it declines from 15.3 percent to 14.0 percent. Not surprisingly, similar to the removal of mobility linked pairs, the proportion of geographically co-located control-cited patent pairs decline only slightly (explained by the observation that fewer corresponding control-cited patent pairs are linked via mobility or connected at such close social distances). As indicated by the z-test, the differences in co-location between citing-cited and control-cited pairs are still statistically significant. Citing patents are now 32 percent more likely than control ones to come from the same state of the cited patents while citing patents are 34 percent more likely than control ones to come from the same region of the cited patents.

As a final step, all citing-cited patent pairs linked by a network tie at any distance were removed from the sample and the geographical matching rates were recalculated for the remaining observations. The results are reported in the very bottom row of each panel in Table 6. Despite the relatively large number of pairs connected at distance greater than 15, the localization rates for both state and regions decline only slightly (see also Fig. 2). Similar to preceding test results, these findings suggest that being connected remotely is relatively less effective at conveying knowledge among network nodes (inventors). That is, social ties distant to the core of the network explain a much smaller fraction of localization effects. Overall, these findings are similar to previous findings.

Because it accounts for the largest portion of localized knowledge flows, inventor mobility deserves further examination. Interesting labor mobility patterns emerge. First, the majority of inventors in this period signed patents for just one organization (77.8%). Most mobile inventors signed patents for just two organizations throughout their inventive career (72.9% of mobile inventors, compared with 16.2% of all inventors). Very few inventors signed patents for five or more assignees. A mobility rate of 22.2 percent for all inventors indicates that the movement of inventors across organizations is a very limited phenomenon in the biopharmaceutical industry during the period under examination. What is more relevant for economic development planning is that the mobility of all inventors across regions and states is even more limited: 16.9 percent regionally and 15.2 percent at a state level. Although the propensity to change location is higher for inventors that move across organizations, only 16.9 percent of mobile inventors and 5.6 percent of all inventors has been active in more than one region. A similar pattern is observed for the state level; not surprisingly, a smaller percentage of inventors move across states than regions. Intuitively, as the number of assignees that an inventor has worked or signed patents for increases, the number of regions and states in which he or she has been active increases as well. Overall, these figures suggest that a large portion of job mobility (approximately 85%) occurs within a given state and region. These observations not only confirm the role of mobility in the localization of patent citations as noted earlier, but also have implications for economic development policies designed to attract skilled workers and professionals.

BL present directly comparable evidence to the present analysis. First, similar to the results reported above (Table 6), these authors find that mobility and network ties at relatively short distances account for the largest portion of localized citations. They observe approximately 4 percent declines in the localization rates at both the state and MSA levels after removing mobility-linked citing-cited pairs from their sample. The corresponding figures in the present analysis are 6 percent for the state level and 7 percent for the regional level. These differences may be explained by the fact that a larger fraction of citing-cited patents are connected through labor mobility in the present analysis than the previous one (9% versus 5%).

Downloaded from https://academic.oup.com/spp/article-abstract/doi/10.1093/scipol/scx001/2964537/Inventor-mobility-and-the-geography-of-knowledge
by lib-electronic@uic.edu user
on 02 October 2017

Table 7. Social connections and geographical matching at the state and regional levels (citing patents without forward citations).

	N[a]	Citing-cited		Control-cited		*z*-test (P > *z*)	Odds ratio
		Frequency	%	Frequency	%		
State level							
All patent pairs	6,653	2,019	30.3	1,196	17.9	16.67 (0.007)	1.99
All pairs except those linked by mobility	6,015	1,439	23.9	1,075	17.8	8.16 (0.007)	1.45
All pairs except those linked by SN	5,380	1,099	20.4	905	16.8	4.80 (0.007)	1.27
Regional level							
All patent pairs	6,653	1,546	23.2	753	11.3	18.18 (0.006)	2.37
All pairs except those linked by mobility	6,015	974	16.2	674	11.2	7.95 (0.006)	1.53
All pairs except those linked by SN	5,380	764	14.2	576	10.7	5.49 (0.006)	1.38

[a]N: number of observations.

Second, mobility, in combination with social ties, at any distance accounts for nearly 6 percent of the localization rates for both state and MSA levels in BL's study while it explains approximately 9 percent of the localization rates in the present study. Again the difference is explained by the fact that in the present study a larger proportion of patent pairs are connected by mobile inventors. Notably, however, social ties in both studies account for a very similar portion of localized knowledge diffusion patterns at both spatial levels of analysis. The reason that a relatively large number of mobility-linked citing-cited patent pairs is observed in the present study can be related to the qualitative differences between patents filed at EPO and USPTO. BL examine US patents filed at EPO and the present study examines US patents filed at USPTO. As mentioned earlier, if EPO patents are higher quality patents, given the cost and legal complexity associated with extending them to Europe, they are perhaps prepared by a relatively small number of high impact scientists and prolific inventors at prominent organizations that do not move across organizations as frequently. Or it may be that the employees of these organizations or inventors who assigned their patents to such organizations move relatively less frequently than those inventors whose patents are not worth extending to Europe.

To test whether mobility and network connections matter more or less depending on the quality of (citing) patents, the major steps in Table 6 were repeated for two groups of citing-cited patent pairs (one group of citing patents with citations and another group of citing patents without citations; see Section 3 for how the quality of a patent is approximated by the number of times it is cited subsequently). Looking at the change in odds ratios after mobility pairs were removed from the samples, at the state level mobility linked patent pairs seem to explain a very similar portion of localized knowledge flows for both groups of patents (the decline from 1.99 to 1.45 in Table 7 is slightly greater than the decline from 1.89 to 1.41 in Table 8). Similarly, at the regional level, mobility linked patent pairs seem to explain a very similar portion of localized knowledge flows for low quality and high quality patents (the decline from 2.37 to 1.53 in Table 7 is only slightly less than the decline from 2.04 to 1.30 in Table 8). Based on these results, it is difficult to conclude whether or not labor mobility facilitates knowledge flows more for high quality patents than low quality patents.

To test whether network connections matter more or less depending on the quality of the (citing) patents, in addition to mobility linked patent pairs all citing-cited patent pairs linked by a network tie at any distance were removed from the sample and the geographical matching rates were recalculated for the remaining observations. The results are reported in the very bottom row of each panel in Table 7 and Table 8. Comparing the change in odds ratios, network linkages seem to explain a slightly larger proportion of localized knowledge flows for high quality patents at the state level (the decline from 1.41 to 1.20 in Table 8 is slightly greater than the decline from 1.45 to 1.27 in Table 7). The magnitude of difference at the regional level is larger. Network linked patent pairs seem to explain a larger proportion of localized knowledge flows for high quality patents (the decline from 1.30 to 1.09 in Table 8 is greater than the decline from 1.53 to 1.38 in Table 7). These results indicate that network connections are more important for the transmission of knowledge for high quality patents than they are for low quality patents.

A third point of comparison between the two studies is the statistically significant odds ratios based on the fraction of co-located citing-cited patent pairs that are left unexplained after all mobility and socially connected pairs were removed from the test. BL report 1.24 for the state level and 1.21 for the Metropolitan Statistical Areas (MSA) level. In the present analysis, the corresponding figures are 1.25 and 1.30 (Table 6). The relatively large difference between regional-level odds ratios between the two studies indicates that citing-cited U.S. patents filed at USPTO are more regionally concentrated than their counterparts filed at EPO (this outcome is also implied in Table 1). Another potential explanation is the difference in the definition of regions in the two studies. BL designate MSAs as regions, whereas this study uses the US Bureau of Economic Analysis's functional economic areas (EAs). Because EAs are larger than MSAs, different regional designations cannot explain the observed difference between odds ratios (the larger the region the greater the likelihood that control patent would be co-located with the cited patent, and the smaller the odds ratio for citing-cited versus control-cited patent pairs). In fact, this makes the difference between the two studies substantively greater.

Finally, the two studies differ in the explanation given for the unaccounted portion of localized knowledge flows. Note that the odds ratios reported in the very bottom row of each panel in Table 6 are still highly statistically significant (after the removal of all socially linked cases reduced their magnitude substantially). It appears that other interactions among inventors such as those derived from formal organizational contractual agreements (e.g. R&D, licensing, commercialization) or local informal encounters and meetings also play an important role in knowledge flows at the local level. It is argued that BL (460-465) overly downplay the role of such mechanisms when the authors say:

> . . . the logical room left to more informal social ties, conventionally thought to be responsible for the localized diffusion of tacit knowledge, appears to be greatly reduced. In the absence of localized movements of inventors and the ensuing creation of

Downloaded from https://academic.oup.com/spp/article-abstract/doi/10.1093/scipol/scx001/2964537/Inventor-mobility-and-the-geography-of-knowledge
by lib-electronic@uic.edu user
on 02 October 2017

Table 8. Social connections and geographical matching at the state and regional levels (citing patents with forward citations).

	N[a]	Citing-cited		Control-cited		z-test (P > z)	Odds ratio
		Frequency	%	Frequency	%		
State level							
All patent pairs	2,482	677	27.2	411	16.5	9.13 (0.011)	1.89
All pairs except those linked by mobility	2,305	505	21.9	383	16.6	4.56 (0.011)	1.41
All pairs except those linked by SN	1,963	362	18.4	311	15.8	2.16 (0.001)	1.20
Regional level							
All patent pairs	2,482	484	19.5	264	10.6	8.73 (0.010)	2.04
All pairs except those linked by mobility	2,305	312	13.5	248	10.7	2.89 (0.012)	1.30
All pairs except those linked by SN	1,963	220	11.2	203	10.3	0.88 (0.009)	1.09

[a]N: number of observations.

> closed networks of formally collaborating agents (co-inventors), informal linkages (such as those we leave out from our co-invention networks) are likely to explain *only a minor fraction* of the observed phenomenon (emphasis added).

It is true that in both studies once social connections and mobility are together accounted for, a smaller portion of localization remains to be explained. However, this remaining portion is nearly as large as what is explained by the social connections alone (when the geographical matching rate for both patent pairs becomes equal or the odds ratio becomes 1, there is no geographical localization of knowledge flows).

5. Policy implications and future research

Ascertaining the extent of localized knowledge flows is one way of empirically establishing the micro-foundations of industrial agglomeration (Ellison et al. 2010). The application of social network analysis to examine knowledge diffusion patterns helps disentangle spatial proximity effects from social proximity effects. The contribution of this study is that it reinforces the previous empirical findings by showing that they are replicable using different observations and a different time frame. In addition, the results indicate that network connections are more important for the transmission of knowledge for high quality patents than they are for low quality patents.

The substantial concentration of local knowledge flows found in this study suggests that industrially targeted public financial support for research and development activities at the regional and state levels can be considered as supportive of firm performance and by extension economic development. For example, when the state government contributes to local firms and universities' scientific research and technological efforts, either directly by making financial contributions or indirectly through tax reductions, the outcomes of such supported research ultimately benefit other local firms and research institutions as well as the targeted organizations. To the extent that there are tangible positive outcomes for local organizations from 'a targeted intervention' by a regional economic agency or a state government, the next relevant question becomes how to maximize such positive externalities or the number of local beneficiaries. First, given that regional and state level labor mobility explains the largest fraction of localized knowledge flows, legal barriers that limit or prevent movement of labor across organizations locally should be reconsidered. In fact, Belenzon and Schankerman (2013) show that noncompete statutes limit not only labor mobility, but also the knowledge diffusion that labor mobility generates. Similarly, Marx et al. (2014) report that employee noncompete agreements[7] not only reduce within-state mobility among firms but also induce inventors to exit the state. Their results show that exiting inventors migrate specifically to states where employee noncompete agreements are unenforceable. Furthermore, in terms of the potential economic impact of such movement for the 'inventor-losing' states, researchers find migration to be the most prevalent among workers who are more collaborative and whose work has greater impact. The policy regarding employee noncompete agreements is implementable. A number of states, including Texas, Louisiana, Florida, Idaho, New York, and New Hampshire have reversed their noncompete employment policies within the last 20 years (Marx et al. 2014).

This study is not without limitations. First limitation is related to its research setting. Our study is only focused on the US biopharmaceutical industry, and the interpretations of our findings may be limited by the characteristics of this particular industry. However, taking an industry-based approach is in line with the most recent contributions (Massard and Autant-Bernard 2015) that note that the variation across industries in terms of the spatial dimension of knowledge flows and the need to produce industry specific empirical evidence to better inform science and technology policy making. Second limitation is related to patent data. There are two major shortcomings of using patents as measures of inventions and knowledge-based interactions. First, their representativeness of the type of knowledge production processes and inventions dominant in the industry is limited by the nature of technological progress and products in the industry and the strategic decisions of innovators. Patents represent only a subset of actual inventions and technological progress occurring in the industry. Not all useful research outcomes are patentable and not all research outcomes that are patentable are patented. The basic nature of some research outcomes simply do not meet the practicality or social utility criteria of the patent office and are not considered for the patentability, hence are not observed. From the point of view of firms, patenting is ultimately a business decision and for strategic reasons firms may choose secrecy over the limited legal rights that come with patents (Jaffe and Trajtenberg 2002). Although relative to firms in other industries, firms in the biopharmaceutical industry largely depend on patenting for intellectual property protection (citation), there is probably still some level of secrecy in the industry. The second shortcoming of using patent data concerns the content of patents. Patent documents by their definition include only the process and results of successful development efforts; they do not contain information about failed experiments. But, 'knowledge of what does not work, what approaches have previously been tried, and led to dead ends, are part of local knowledge [base]' which could be equally

Downloaded from https://academic.oup.com/spp/article-abstract/doi/10.1093/scipol/scx001/2964537/Inventor-mobility-and-the-geography-of-knowledge
by lib-electronic@uic.edu user
on 02 October 2017

valuable for future production and technological development efforts (Feldman and Kogler 2010: 386).[8] Despite these limitations, with the wealth of information they contain, patent records are among the few data sources that enable large-scale research on the economics and geography of innovation.

There are several lines of investigation for future research that could likely improve the findings reported in this study. First, when accounting for the role of established social connections in knowledge flows in Section 4, in order not to deviate too much from the core focus of the research question, it is assumed that social connections established in any of the last 5 years (among patent authors before the invention year) would function the same way in facilitating knowledge flows. It is conceivable that more currently established, fresh connections convey knowledge more efficiently. In the future, a temporal decay specification of social relations could be a useful extension for a more precise understanding of the impact of social connections in explaining knowledge flows. Specifically, one can impose a decay function over time so that ties established more recently are presumed to invoke closer social links than older ties. Second, citation frequency and inventor mobility may vary based on the specialty of inventors or subfields of the biopharmaceutical industry in which inventors perform research and development activity. For example, patents granted for the invention of general purpose technologies or biotechnology research tools are likely to receive relatively more citations given their wider applicability and/or relevance for future inventions. Thus, taking into account the presence of different research fields could produce more precise and nuanced understanding of inventor interaction and knowledge diffusion patterns dominant in the industry.[9] Third, one can examine the extent to which the localization patterns observed in this study are explained by recent university graduates' decision to stay where they graduate. It might be that labor market stickiness for this group of workers is accounting for a portion of localized knowledge flows, such that frequent local hiring of recent graduates functions as a mediator between local university scientists and their industrial counterparts, resulting in higher tendency of local firms to cite local university research outcomes more frequently. This could be ascertained by determining inventors' graduating school affiliations and then repeating the same localization test conducted in this study. But, to the best of our knowledge, these data are not available on a large scale. Data limitations such as this one specifically and the challenge of local innovation policy design with regard to the facilitation of knowledge flows in general imply new roles for economic development agencies and local governments that seek to promote knowledge-intensive industrial development in their jurisdictions.

Notes

1. Unless otherwise mentioned, labor mobility refers to the movement of inventors across different organizations (companies).
2. Nagaoka et al. (2010) argue that because there is no way to invent around new chemical entities in drugs (as protected by a compound patent), patents are a relatively effective mechanism for appropriating rents from innovation in the biopharmaceutical industry.
3. The authors of this project (Li et al. 2014) apply a disambiguation algorithm to the entire USPTO patent database that clusters patents by the same inventor and distinguishes them from patents by other inventors with the same or a similar name. The general approach consists of developing a set of attributes for each ambiguous inventor-patent and then computing similarity profiles for each instance. The likelihood values are then used to determine the level of evidence that helps researchers declare a pair as a match or a nonmatch.
4. 'Region' refers to the US Bureau of Economic Analysis (BEA)-designated functional EAs.
5. There are four mechanisms of knowledge diffusion among co-located firms. The first mechanism is labor mobility as discussed in this article. The second mechanism relates to informal encounters and interactions among local workers and entrepreneurs. The third mechanism concerns direct inter-firm links in cooperation networks. Finally, the generation of spin-offs is a mechanism of knowledge transfer that tends to be spatially concentrated (Capello 1999; Ter Wal and Boschma 2009).
6. Network density is the ratio of the number of present connections to the number of all possible connections. Average path length is the mean of the shortest distance between each pair of nodes in the network.
7. 'Non-compete agreements are employment contracts that place restrictions on the sorts of jobs ex-employees may take after leaving the firm, usually for a term of 1–2 years. Although companies frequently ask employees to sign non-disclosure agreements that bar them from sharing trade secrets, violations can be difficult to detect whereas it is more straightforward to determine whether an ex-employee joined a competing firm' (Marx et al. 2014: 1).
8. Note, however, that one cannot rule out completely that inventors do not exchange such knowledge when they are socially connected. In fact, the social network approach to explaining localized knowledge spillovers not only views citation as a function of maintained social relations, but also assumes that knowledge besides that revealed in the patent document is also conveyed, such as details of experimentations, scientific deliberations, etc. (Breschi et al. 2010).
9. We thank an anonymous reviewer for pointing this distinction.

References

Agrawal, A., Cockburn, I., and McHale, J. (2006) 'Gone but not Forgotten: Knowledge Flows, Labor Mobility, and Enduring Social Relationships', *Journal of Economic Geography*, 6: 571–91.

Aharonson, B., Baum, J. A. C., and Feldman, M. P. (2004) 'Industrial Clustering and Returns to Inventive Activity: Canadian Biotechnology Firms, 1991-2000', Working Paper No. 04-03, Danish Research Unit for Industrial Dynamics, Copenhagen.

Almeida, P., and Kogut, B. (1999) 'Localization of Knowledge and the Mobility of Engineers in Regional Networks', *Management Science*, 45: 905–18.

Balland, P. A., Boschma, R., and Frenken, K. (2015) 'Proximity and Innovation: From Statics to Dynamics', *Regional Studies*, 49/6: 907–20.

Bathelt, H., and Cohendet, P. (2014) 'The Creation of Knowledge: Local Building, Global Accessing and Economic Development—Toward an Agenda', *Journal of Economic Geography*, 14: 869–82.

Downloaded from https://academic.oup.com/spp/article-abstract/doi/10.1093/scipol/scx001/2964537/Inventor-mobility-and-the-geography-of-knowledge
by lib-electronic@uic.edu user
on 02 October 2017

——, Malmberg, A., and Maskell, P. (2004) 'Clusters and Knowledge: Local Buzz, Global Pipelines and the Process of Knowledge Creation', *Progress in Human Geography*, 28/1: 31–56.

Belenzon, S., and Schankerman, M. (2013) 'Spreading the Word: Geography, Policy, and Knowledge Spillovers', *The Review of Economics and Statistics*, 95/3: 884–903.

Boschma, R. A. (2005) 'Proximity and Innovation: A Critical Assessment', *Regional Studies*, 39: 61–74.

——, and Frenken, K. (2006) 'Why is economic Geography not an Evolutionary Science? Towards an evolutionary Economic Geography', *Journal of Economic Geography*, 6/3: 273–302.

Bouba-Olga, O., Carrincazeaux, C., Coris, M., and Ferru, M. (2015) 'Editorial: Proximity Dynamics, Social Networks and Innovation', *Regional Studies*, 49/6: 901–6.

Breschi, S., and Lissoni, F. (2001) 'Knowledge Spillovers and Local Innovation Systems: A Critical Survey', *Industrial and Corporate Change*, 10/4: 975–1005.

——, and —— (2009) 'Mobility of Skilled Workers and Co-Invention Networks: An Anatomy of Localized Knowledge Flows', *Journal of Economic Geography*, 9/4: 439–68.

——, Lenzi, C., Lissoni, F., and Vezzulli, A. (2010) 'The Geography of Knowledge Spillovers: The Role of Inventors' Mobility Across Firms and in Space', in R Boschma and R Martin (eds) *The Handbook of Evolutionary Economic Geography*, pp. 353–69. London: Edward Elgar.

Broekel, T. (2015) 'The Co-evolution of Proximities – A Network Level Study', *Regional Studies*, 49/6: 921–35.

Cappelli, R., and Montobbio, F. (2016) 'European Integration and Knowledge Flows Across European Regions', *Regional Studies*, 50/4: 709–27.

Capello, R. (1999) 'Spatial Transfer of Knowledge in High Technology Milieux: Learning Versus Collective Learning Processes', *Regional Studies*, 33/4: 353–65.

Cooke, P. (2007) 'European Asymmetries: A Comparative Analysis of German and UK Biotechnology Clusters', *Science and Public Policy*, 34/7: 454–74.

Crespo, J., and Vicente, J. (2016) 'Proximity and Distance in Knowledge Relationships: From Micro to Structural Considerations Based on Territorial Knowledge Dynamics (TKDs)', *Regional Studies*, 50/2: 202–19.

Dimasi, J. A., and Grabowski, H. G. (2007) 'The Cost of Biopharmaceutical R&D: Is Biotech Different?', *Managerial and Decision Economics*, 28: 469–79.

Ellison, G., Glaeser, E., and Kerr, W. (2010) 'What Causes Industry Agglomeration? Evidence from Co-Agglomeration Patterns', *American Economic Review*, 100/3: 1195–213.

Feldman, M. P., and Kogler, D. F. (2010) 'The Stylized Facts in the Geography of Innovation', in BH Hall and N Rosenberg (eds) *Handbook of the Economics of Innovation*, pp. 381–410. London: North-Holland.

Fornahl, D., Broekel, T., and Boschma, R. (2011) 'What Drives Patent Performance of German Biotech Firms? The Impact of R&D Subsidies, Knowledge Networks and their Location', *Papers in Regional Science*, 90/2: 395–418.

Fritsch, M., and Kauffeld-Monz, M. (2010) 'The Impact of Network Structure on Knowledge Transfer: An Application of Social Network Analysis in the Context of Regional Innovation Networks', *Annals of Regional Science*, 44: 21–38.

Gallié, E. P. (2009) 'Is Geographical Proximity Necessary for Knowledge Spillovers Within a Cooperative Technological Network? The Case of the French Biotechnology Sector', *Regional Studies*, 43/1: 33–42.

Gambardella, A., Harhoff, D., and Verspagen, B. (2008) '*The Value of European Patents*', Working Paper.

Gertler, M. S., and Levitte, Y. M. (2005) 'Local Nodes in Global Networks: The Geography of Knowledge Flows in Biotechnology Innovation', *Industry and Innovation*, 12: 487–507.

Gluckler, J. (2007) 'Economic Geography and the Evolution of Networks', *Journal of Economic Geography*, 7: 619–34.

Hall, B. H., Jaffe, A., and Trajtenberg, M. (2005) 'Market Value and Patent Citations', *RANT Journal of Economics*, 36/1: 16–38.

Harhoff, D., Narin, F., Scherer, F. M., and Vopel, K. (1999) 'Citation Frequency and the Value of Patented Inventions', *The Review of Economics and Statistics*, 81/3: 511–3.

Hess, M. (2004) 'Spatial" Relationships? Towards Reconceptualization of Embeddedness', *Progress in Human Geography*, 28/2: 165–86.

Jaffe, A., and Trajtenberg, M. (2002) *Patents, Citations and Innovations*. Cambridge, MA: Cambridge University Press.

——, ——, and Henderson, R. (1993) 'Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations', *Quarterly Journal of Economics*, 108: 577–98.

Jenkins, M., and Tallman, S. (2010) 'The Shifting Geography of Competitive Advantage: Clusters, Networks and Firms', *Journal of Economic Geography*, 10: 599–618.

Li, G. C., Lai, R., D'Amour, A., Doolin, D. M. et al. (2014) 'Disambiguation and Co-authorship Networks of the U.S. Patent Inventor Database (1975-2010)', *Research Policy*, 43/6: 941–55.

Lissoni, F. (2001) 'Knowledge Codification and the Geography of Innovation: The Case of Brescia Mechanical Cluster', *Research Policy*, 30: 1479–500.

Malmberg, A., and Maskell, P. (2002) 'The Elusive Concept of Localization Economies: Towards a Knowledge Based Theory of Spatial Clustering', *Environment and Planning A*, 34: 429–49.

——, and —— (2006) 'Localized Learning Revisited', *Growth and Change*, 37: 1–19.

Markusen, A. (2003) 'Fuzzy concepts, scanty evidence, policy distance: The case for rigor and policy relevance in critical regional studies', *Regional Studies*, 33: 869–84.

Marx, M., Singh, J., and Fleming, L. (2014) 'Regional Disadvantage? Employee Non-Compete Agreements and Brain Drain', *Research Policy*, 44/2: 394–404.

Massard, N., and Autant-Bernard, C. (2015) 'Editorial: Geography of Innovation: New Trends and Implications for Public Policy Renewal', *Regional Studies*, 49/11: 1767–71.

Miguelez, E., and Moreno, R. (2014) 'What Attracts Knowledge Workers? The Role of Space and Social Networks', *Journal of Regional Science*, 54/1: 33–60.

Mowery, D. C., and Ziedonis, A. A. (2004) 'The geographic reach of market and non-market channels of technology transfer: comparing citations and licenses of University Patents'. in J Cantwell (ed.) *Globalization and the Location of Firms*. Northampton, MA: Edward Elgar.

Nagaoka, S., Motohashi, K., and Goto, A. (2010) 'Patent Statistics as an Innovation Indicator', in BH Hall and N Rosenberg (eds) *Handbook of the Economics of Innovation*, pp. 1083–127. London: North-Holland.

Owen-Smith, J., and Powell, W. W. (2004) 'Knowledge Networks as Channels and Conduits: The Effects of Spillovers in the Boston Biotechnology Community', *Organization Science*, 15/1: 5–21.

Powell, W. W., and Sandholtz, K. (2012) 'Chance, necessiate, et naivete: Ingredients to Create a New Organizational Form', in JF Padgett and WW Powell (eds) *The Emergence of Organizations and Markets*, pp. 379–433. Princeton: Princeton University Press.

Peri, G. (2005) 'Determinants of Knowledge Flows and Their Effect on Innovation', *The Review of Economics and Statistics*, 87/2: 308–22.

Petruzzelli, A. M., Rotolo, D., and Albino, V. (2015) 'Determinants of Patent Citations in Biotechnology: An Analysis of Patent Influence Across the Industrial and Organizational Boundaries', *Technological Forecasting and Social Change*, 91: 208–21.

Singh, J. H. (2005) 'Collaborative networks as determinants of knowledge diffusion patterns', *Management Science*, 51: 756–70.

——, and Marx, M. (2013) 'Geographic Constrains on Knowledge Spillovers: Political Borders vs. Spatial Proximity', *Management Science*, 59/9: 2056–78.

Sonn, J. W., and Storper, M. (2008) 'The Increasing importance of geographical proximity in knowledge production: An analysis of U.S. patent citations, 1975-1997', *Environment and Planning a*, 40/5: 1020–39.

Downloaded from https://academic.oup.com/spp/article-abstract/doi/10.1093/scipol/scx001/2964537/Inventor-mobility-and-the-geography-of-knowledge
by lib-electronic@uic.edu user
on 02 October 2017

Steiner, M., and Hartmann, C. (2006) 'Organizational Learning in Clusters: A Case Study on Material and Immaterial Dimensions of Cooperation', *Regional Studies*, 40/5: 493–506.

Ter Wal, A., and Boschma, R. A. (2009) 'Applying Social Network Analysis in Economic Geography', *Annals of Regional Science*, 43: 739–56.

Thompson, P., and Fox-Kean, M. (2005) 'Patent Citations and the Geography of Knowledge Spillovers: A Reassessment', *American Economic Review*, 95: 450–60.

Von Wartburg, I., Teichert, T., and Rost, K. (2005) 'Inventive Progress Measured by Multi-stage Patent Citation Analysis', *Research Policy*, 34: 1591–607.

Appendix

z - test is defined as:

$$z = \frac{P_1 - P_2}{SE}$$

$$SE = \sqrt{p* \ (1-p)* \left[\left(\frac{1}{n_1}\right) + \left(\frac{1}{n_2}\right) \right]}$$

$$p = \frac{(P_1 * n_1) + (P_2 * n_2)}{(n_1 + n_2)}$$

Where: $p1$ and $p2$ are the sample proportion estimate for the citing and the control-cited patents that match geographically. n_1 and n_2 are the size of the citing-cited and control-cited samples (in this research: $n_1 = n_2$).

The odds of co-location for citing-cited patent pairs is defined as:

$$O_1 = \frac{P_1}{1 - P_1}$$

Let O_2 be the corresponding odds of co-location for control cited patent pairs. The odds ratio is defined as $OR = \frac{O_1}{O_2}$.

Downloaded from https://academic.oup.com/spp/article-abstract/doi/10.1093/scipol/scx001/2964537/Inventor-mobility-and-the-geography-of-knowledge
by lib-electronic@uic.edu user
on 02 October 2017