Proximal advantage in knowledge diffusion : the time dimension

Wang, Jue; Zhang, Liwei

2018

Wang, J., & Zhang, L. (2018). Proximal advantage in knowledge diffusion : the time dimension. *Journal of Informetrics*, 12(3), 858-867. doi:10.1016/j.joi.2018.07.006

<https://hdl.handle.net/10356/139614>

<https://doi.org/10.1016/j.joi.2018.07.006>

© 2018 Elsevier Ltd. All rights reserved. This paper was published in *Journal of Informetrics* and is made available with permission of Elsevier Ltd.

Downloaded on 26 Apr 2023 21:11:23 SGT

Proximal advantage in knowledge diffusion: the time dimension

Jue Wang^{1*} and Liwei Zhang¹²

¹ Public Policy and Global Affairs Division, Nanyang Technological University, 14 Nanyang Drive, Singapore 637332; Email: wangjue@ntu.edu.sg; ORCID 0000-0002-3401-713X

² WISE Lab, School of Humanities and Social Science, Dalian University of Technology, Dalian 116024 China

* Corresponding author

Abstract

This paper intends to explore the impact of geographic proximity on the diffusion of knowledge in the form of publication citations, and argues that codified knowledge is transmitted faster in proximity and is subject to similar geographic constraints as tacit knowledge. The geographic proximity advantage would be particularly relevant in the early stage of dissemination. We collected three sets of research articles published in 1990, 2000 and 2010 and compared the longitudinal citations they received domestically and from abroad. The study found that domestic citations accumulate faster and reach their peak much earlier than foreign citations, and the difference is most evident in the early years after publication. The result shows that geographic proximity does play a role in the speed of knowledge diffusion and points to the network effect for citations. Those located closer to the knowledge origin would be exposed and react to publications faster due to the additional opportunities of research exchange and network.

Keywords: speed of diffusion; publication citation; proximal advantage; network effect; time dimension

1. Introduction

The impact of geographic proximity on knowledge creation and diffusion has been discussed extensively in the literature on the geography of economics (Amin & Wilkinson, 1999; Audretsch & Feldman, 1996; Boschma, 2005; Jaffe, Trajtenberg, & Henderson, 1993), and the literature has been continuously growing over the years (Bouba-Olga, Carrincazeaux, Coris, & Ferru, 2015). The discussion can be traced back to the 1950s, when Polanyi (1958) suggested distance as a factor in the transmission of knowledge depending on the property of knowledge – tacit or codifiable. Tacit knowledge refers to the knowledge pieces that are accumulated through experience and often are hard to formalize and express. It is often disseminated in interpersonal networks from one researcher to another researcher, through channels such as training, observation, discussion, and conferences etc. The transmission is typically localized in both physical and social space as it often requires face-to-face interaction (Stefano Breschi & Lissoni, 2005). Therefore, the dissemination of tacit knowledge is bounded by spatial constraints and more likely to occur among those in close proximity, and gradually reaches more distant parties (Sorenson & Fleming, 2004). By contrast, codified knowledge refers to knowledge that can be converted into symbols and is transmittable in formal and systematic language. It is less bounded in space and can be broadcast over long distances and across organizational boundaries (Polanyi, 1958), in particular with the globalization and the development of information and communication technologies (ICTs). The trend is evidenced in studies found a growing number of internationally co-authored publications (Coccia & Wang, 2016; Caroline S. Wagner & Leydesdorff, 2005; Caroline S. Wagner, Whetsell, & Leydesdorff, 2017), and studies observed a diminishing role of distance (Choi, 2012) and territorial borders (Hoekman, Frenken, & Tijssen, 2010).

However, some other empirical analysis of the spatial dimension of codified knowledge presented different evidence. Following Jaffe et al's (1993) seminal work, a series of studies have examined the spatial dimension of the knowledge diffusion via the channel of patent citations (Audretsch & Feldman, 1996; Jaffe, 1989; Jaffe & Trajtenberg, 1996, 1999; Jaffe et al., 1993; LeSage, Fischer, & Scherngell, 2007), and found that as a form of codified knowledge, patents are more likely to receive citations from proximal agents. Other studies have examined spatial spillover by estimating knowledge production function (Anselin, Varga, & Ac, 1997; Autant-Bernard, 2001; Bottazzi & Peri, 2003; Jaffe, 1989) and from the perspectives of mobility and collaborative network (Autant-Bernard, Mairesse, & Massard, 2007; Stefano Breschi, Lenzi, Lissoni, & Vezzulli, 2010; Frenken et al., 2009; Glänzel & Schubert, 2005; Katz, 1994; Ponds, Van Oort, & Frenken, 2007; Ponds, van Oort, & Frenken, 2010; Ter Wal, 2014), which have all demonstrated the existence of geographic influence.

This study intends to join the debate and further explore the role of geography in knowledge diffusion using a less explored approach: publication citations. Scientific publications serve as a carrier of codified knowledge and are an important channel of knowledge diffusion (Sorenson & Fleming, 2004), with the embedded knowledge transmitted between studies, research fields, and geographic locations (Lewison, Rippon, & Wooding, 2005). Publication citation is one of the popular instruments to trace knowledge accumulation as the later publication builds on the cited publication (Garfield, 1979; Leydesdorff, 1998; D.J. de Solla Price, 1965). It better reflects the accumulation process than patent citations (Wuestman, Hoekman, & Frenken, 2018) as the latter implies similarity instead of contribution and can be added by the examiners (Meyer, 2000). Publication citation networks have been used to capture knowledge flow at several levels including those of disciplines (Jian Wang, Veugelers, &

Stephane, 2017; Yan, 2014), journals (Bollen, Rodriguez, & Van De Sompel, 2006), institutions (Börner, Penumarthy, Meiss, & Ke, 2006; Yan & Sugimoto, 2011) and countries (Chen & Hicks, 2004; Lewison et al., 2005; Liu & Rousseau, 2010). However, only a small number of studies have explicitly used publication citations to examine the influence geographic proximity on knowledge diffusion. Matthiessen et al. (2002) analyzed the citation networks of publications produced in top 40 cities and found “national links dominate over international links”. Glänzel & Schubert (2005) did a similar study for the top 36 countries and found the domesticity of citations vary across countries. At a regional level, both Börner et al (2006) and Pan (2012) found space matters for knowledge diffusion, and citations are more likely to come from institutions or cities nearby. Wuestman (2018) pointed out that geographic proximity enhances citations but also noted that the effect is dependent on cognitive proximity. Building on this line of literature, our study intends to further examine the spatial boundary in the dissemination of publications, particularly its influence on the speed of the dissemination. Having the time dimension helps to reveal the dynamics of knowledge diffusion and explain why geographical proximity matters.

The outline of this paper is as follows. In Section 2, we start with a literature review and propose our hypothesis. In Section 3, we discuss features of our dataset and apply the model. The estimation results are presented in section 4. In Section 5, we conclude with remarks.

2. Geographical proximity and speed of diffusion

The study examines the effect of geography on the speed of knowledge dissemination and argue that space and time interactively play a role in the diffusion process. Following Jaffe and Trajtenberg (1996; 1999), we use country as the geographic unit in this study as country borders

typically demarcate zones with different languages, history, culture, and political institutions (Bornmann, Wagner, & Leydesdorff, 2018; Girvan & Newman, 2002), all of which would influence knowledge spillover. In the context of this study, speed of diffusion refers to how fast the publication receives citations domestically or from abroad. Citation practices are not constant in time (Wallace, Larivière, & Gingras, 2009). When research is newly published, it may not receive citations immediately and often takes time for the publication to be recognized in the scientific community. The delayed recognition could be a result of clogged information channels due to information explosion (Garfield, 1980), low author visibility (Barber, 1962), or most importantly, poor quality of paper (Cole, 1970) or paper content being pre-mature (Garfield, 1980) or too novel (Jian Wang et al., 2017). Geographic proximity eases the exchange of ideas and the transmission of knowledge (Feldman 1999), and we expect it to affect the speed of diffusion by alleviating the above mentioned problems.

Working in an “invisible college”, scientists communicate with each other in various ways to exchange ideas and these communication patterns affect the development of knowledge (Diana Crane, 1972). Scientists may form social groups based on shared research interest and maintain high level informal communication (D.J. de Solla Price, 1963). They could exchange research in progress, research findings and pre-prints with one another (Diana Crane, 1969). Intellectual and interpersonal ties are important in citations (White, Wellman, & Nazer, 2004). A study by Milard (2014) shows that only 25% of references are cited by the researcher without knowing any of the authors. When faced with several references, the personal relationship – knowing the author – could be the criteria for the choice (Milard, 2014). Levels of acquaintanceship with the authors cited vary from as close as colleagues and friends to those met occasionally in conferences or only knowing by name (Milard, 2014). Being embedded in the

same network enhances a rapid diffusion of knowledge among partners (Victor Gilsing, 2005; Victor Gilsing, Nooteboom, Vanhaverbeke, Duysters, & Oord, 2008). Social network is usually locally concentrated as agents co-located in a region are more likely to be connected with social ties (Stefano Breschi & Lissoni, 2005). Scholars residing in the same country have more opportunities to know each other by attending national conferences, joining national academic associations and competing for funding from national funding agencies. These social network and exchange activities contribute to the awareness of the research and its publication (Cohen, Nelson, & Walsh, 2002; Goldfinch, Dale, & Derouen, 2003). Similar to the early view benefit brought by posting pre-prints on ArXiv (Kurtz et al., 2005; Moed, 2007), early disclosure of the on-going or recent completed research within the network could shorten the lag between the publication and awareness.

In addition, the relationship between social network and quick accumulation of domestic citations could be reinforced by possible similar research interest in the country. As noted by Tang et al (2015), the domestic scientific community is likely to be influenced by national policies or funding programs favoring some priority areas and tends to have similar research agenda. When there is a new finding published on those topics, it could quickly capture attention in the field and gets cited by peers in the country. The international community usually does not respond as fast as the domestic community since communication channels are relatively fewer.

Lastly, physical proximity is particularly important for knowledge with a strong tacit component (Polanyi, 1966). Knowledge codified in a publication alone is not sufficient for fellow scientists to fully judge its meaning and reliability (Collins, 1985). To do this it may often require “complementary know-how” (Dasgupta & David, 1994). Craft skills transferred through demonstration and personal instructions, which are not communicated in a publication, could be

essential for others to acquire the knowledge (Dasgupta & David, 1994) or replicate the study (Collins, 1985). Geographic proximity provides more opportunities for face-to-face interaction between scientists. They would have more chances to learn technical details that are not conveyed in publications from informal conversation or formal presentations, and thus have a better understanding of the research. It helps to transfer the tacit component of the knowledge. As mentioned above, scientists are mostly keen in introducing and promoting their new findings in various occasions, mostly inside the country. The domestic research community, as the audience, can benefit from the conversation and have more knowledge about the new publications.

To summarize, social network increases the visibility of the scholars and their research, and face-to-face interaction allows for better understanding of the research and its implications, both of which are facilitated by geographic proximity. As the scientific exchange occurs most frequently when scholars have new research findings and publications, the domestic community gets to learn the research and its details ahead of foreign peers. As a consequence, we expect to see the knowledge diffuses faster in the domestic community and generates more immediate citations – those received in the first two years (Wallace et al., 2009) – domestically. And here comes the hypothesis: *Domestic citations come earlier than foreign citations.*

3. Methodology

3.1 Data

The data used in this study was retrieved from the Science Citation Index (SCI) of Thomson Reuters Web of Science (WoS). The search was done in February and March 2016 using search term "WC=Physics, Multidisciplinary". We selected the subject category of

multidisciplinary physics as it is one of the basic research fields that has a long history and is a subject universally studied. We then retrieved all research articles in multidisciplinary physics published in the period of Dec 16th – 31th for three years – 1990, 2000 and 2010. We set three time periods spaced ten years apart to see whether and how the knowledge diffusion speed would vary in different decade, and intentionally restricted the dates to the last two weeks of the year (Dec 16th – 31th) for convenience purpose and also to reduce the volume. We used *seed articles* to refer to this set of articles. In the next step we manually collected bibliometric information for all citations by year 2015 for each seed article, and this set of articles is referred as *citing articles*. While the seed articles are only restricted to research articles, the citing articles include all types of publications such as review articles, proceeding papers and letters etc, as the focus here is more on the spread of knowledge. We then extracted author affiliation information for both seed articles and citing articles, manually cleaned and standardized the field of affiliation country.

As a result, we identified 804 seed articles (225 for 1990, 237 for 2000 and 342 for 2010), and 11,200 citing articles for seed articles in 1990, 9970 citing articles for seed articles in 2000, and 10,005 citing articles for seed articles in 2010 (See table 1). The total number of articles in our dataset is 32,024. The citing articles for the set of articles in 1990 accumulated over 26 years from 1990 to 2015, while the citing articles for the latter two sets spanned 16 and 6 years respectively. They constitute an unbalanced panel of time series. An increasingly internationalized trend is observed from the authorship of the publications, which is consistent with previous studies (Caroline S. Wagner et al., 2017). The number of countries contributed to the three sets of seeding articles increased from 37 in 1990 to 43 in 2000 and 58 in 2010. The average citation per article is the highest with the 1990 set of seed articles, and the lowest with

the most recent set of articles, which is expected as papers published earlier have wider time window for generating attention and tend to have more citations. However, after controlling for the time window, the more recent seed articles are found with more citations as indicated by the average citation per article per year, which can be attributed to the “exponential growth of scientific output” in recent years (Bornmann & Mutz, 2015).

Table 1. Seed articles and citing articles

Year	Seed Articles	Citing Articles	Citation per article	Citation per article per year
1990	225	11,200	49.78	1.92
2000	237	9,970	42.07	2.63
2010	342	10,005	29.25	4.88

3.2 Description of domestic and foreign citation

Citing articles are divided into two groups based on the country location of the authors. If any of the authors of the citing article is from a country associated with the cited seeding article, we label the citing article as a domestic citation. Otherwise, if there is no overlapping between the countries of the citing article and those of the seeding article, the citing article is categorized as a foreign citation. For example, if the seed article was a result of international collaboration between the US and Japan, a citing article with an author based in the US or Japan is counted as a domestic citation, while a citing article with no authors from the US or Japan is treated as a foreign citation. Around 4.5% of the 804 seed articles do not have citations at all, and 13.6% of them have no foreign citations. Articles published in recent years are more likely to have cross boarder citations. Only 8% of seed articles published in 2010 have no foreign citations, while the figure for those published in 1990 was 19% (Table 2).

Table 2. Share of seed articles without foreign citations

Year	Seed articles_1990	Seed articles_2000	Seed articles_2010
1990	99.6%	-	-
1991	68%	-	-
1992	52%	-	-
1993	54%	-	-
1994	54%	-	-
1995	55%	-	-
1996	59%	-	-
1997	62%	-	-
1998	63%	-	-
1999	64%	-	-
2000	62%	99%	-
2001	63%	68%	-
2002	66%	46%	-
2003	70%	51%	-
2004	69%	55%	-
2005	71%	54%	-
2006	71%	56%	-
2007	75%	57%	-
2008	72%	57%	-
2009	74%	57%	-
2010	73%	57%	99%
2011	72%	60%	47%
2012	74%	61%	35%
2013	75%	62%	32%
2014	80%	65%	31%
2015	74%	67%	32%
Overall share	19%	12%	8%
# Seed_articles	225	237	342

As mentioned above, knowledge diffusion is bounded by geography and faster in proximity. We expect the domestic citations to accumulate earlier than foreign citations. Table 3 shows the distribution of domestic and foreign citations over the years. The total counts of foreign citations are higher than the domestic citations, mostly because the size of the

international scientific community is much larger than that of the domestic community. As expected, domestic citations are accumulated faster than foreign citations for all three sets of seed articles. In the first two or three years of publication, articles receive more citations from the domestic research community (See shaded cells in Table 3). Starting from the third or fourth year, citations from foreign peers exceed domestic peers.

Table 3. Domestic citations and foreign citations

	Seed articles 1990		Seed articles 2000		Seed articles 2010	
	Domestic citations	Foreign citations	Domestic citations	Foreign citations	Domestic citations	Foreign citations
1990	4	1	-	-	-	-
1991	210	162	-	-	-	-
1992	363	387	-	-	-	-
1993	290	428	-	-	-	-
1994	269	444	-	-	-	-
1995	258	386	-	-	-	-
1996	185	393	-	-	-	-
1997	183	367	-	-	-	-
1998	134	335	-	-	-	-
1999	134	314	-	-	-	-
2000	126	335	8	3	-	-
2001	143	319	294	182	-	-
2002	114	343	489	465	-	-
2003	128	317	445	515	-	-
2004	115	283	397	469	-	-
2005	108	340	330	461	-	-
2006	98	319	311	458	-	-
2007	82	265	248	430	-	-
2008	90	281	217	419	-	-
2009	69	263	217	407	-	-
2010	62	222	187	383	10	4
2011	72	216	204	371	682	535
2012	71	243	203	355	1001	1037
2013	73	202	173	329	969	1210
2014	87	231	161	304	900	1264
2015	87	215	166	320	815	1219
2016	10	24	17	32	149	207
Sum	3,565	7,635	4,067	5,903	4,529	5,476

3.3 Variables

The descriptive statistics show that domestic citations not only accumulate faster than foreign citations but also reach its peak earlier. We would like to further investigate if the phenomena exists after eliminating influences from other factors such as the author country and journal impact factor etc. Given seed articles have citations over the years, we constructed a panel data set with each seed article i being a unit of analysis, and the citation of the seed article i in year t being an observation. The total number of seed articles is 808 and the total number of observations is 11,694.

Dependent variables. Four variables are created to measure the diffusion of knowledge: the count of total citations for seed article i in a given year t (variable TC_{it}), the count of domestic citations in year t (variable DC_{it}), the count of foreign citations in year t (variable FC_{it}), and the difference between foreign citations and domestic citations in that year (variable $DIFF_FC_DC_{it}$). The difference variable $DIFF_FC_DC_{it}$ intends to give a more straightforward comparison for citations in any given year. A positive value shows that the publication receives more citations from foreign scholars than from domestic scholars in that year, and a negative value shows the opposite.

Independent variables. The key independent variable is the time effect as we are interested in the temporal heterogeneity in knowledge diffusion. It is measured by the number of years between citations in the given year t and the year of publication for the seed article i (variable $Time_{it}$). As displayed in Figure 1, most citation curves show that citations increase in the first several years of the publication and then decrease (Garfield 1989). Hence, a squared term of *time* (variable $Time^2_{it}$) is included for testing the curvilinear relationship between citations and the passage of time.

Control variables. The number of author countries for the seed article i (variable $Country_num_i$) is used to control the “bonus effect” of international collaboration as internationally co-authored papers tend to receive more citations (Kato & Ando, 2003; Narin, Stevens, & Whitlow, 1991). It is also used to control the size of domestic research community as there are multiple home countries in the paper and the chances for citations to occur from these countries are higher (Schmoch & Schubert, 2008). Other control variables include the journal impact factor (in 2014) for the seed article i (variable IF_i), the year of publication (variables $PY1990_i$ and $PY2000_i$) and whether author countries include the US (variable US_i). The impact factor indicates a journal’s popularity, and varies across fields depending on disciplinary citing propensity (Zitt & Small, 2008). Publications in high-profile journals are perceived to be of higher quality (Didegah & Thelwall, 2013), and hence are more likely to be cited (Bornmann & Leydesdorff, 2017; Onodera & Yoshikane, 2015; Perneger, 2010; Jue Wang & Shapira, 2015), but the relationship is weakening in the digital age when papers are individually available (Lozano, Larivière, & Gingras, 2012). Two dummies of publication year – 1990 and 2000 – are added to control for the possible variation in citations due to the difference in the type or topic of research conducted in different decades, while year 2010 is used as a benchmark. Lastly, an author country dummy – the US – is included as the US is the leading country in science and may have a higher influence in the knowledge diffusion process than other countries (Bornmann et al., 2018; Gingras & Khelifaoui, 2018). The description and summary statistics of these variables are presented in Table 4 and Table 5.

Table 4. Variable description

VARIABLES	Type	Description
TC_{it}	Count	Count of total citations for the seed article i in year t

DC_{it}	Count	Count of domestic citations for the seed article i in year t
FC_{it}	Count	Count of foreign citations for the seed article i in year t
$DIFF_FC_DC_{it}$	Count	Difference between foreign and domestic citations for the seed article i in year t
$Time_{it}$	Interval	Number of years between citation and publications for the seed article i in year t
$Time^2_{it}$	Interval	Squared term of $Time_{it}$
$Country_num_i$	Count	Number of countries in author affiliations of the seed article i
IF_i	Interval	Journal impact factor in 2014 for the seed article i
US_i	Dummy	Whether the US is in the author country of the seed article i
$PY1990_i$	Dummy	Publication year for the seed article is 1999
$PY2000_i$	Dummy	Publication year for the seed article is 2000

Table 5. Summary statistics of variables

VARIABLES	(1) N	(2) mean	(3) sd	(4) min	(5) max
TC_{it}	11,694	2.628	6.864	0	129
DC_{it}	11,694	1.023	2.835	0	51
FC_{it}	11,694	1.605	4.721	0	102
$DIFF_FC_DC_{it}$	11,694	0.581	3.678	-25	81
$Time_{it}$	11,694	9.124	7.057	0	25
$Time^2_{it}$	11,694	133.0	165.5	0	625
$Country_num_i$	11,694	1.328	1.064	1	37
IF_i	11,694	4.572	3.138	1.359	17.49
US_i	11,694	0.302	0.459	0	1
$PY1990_i$	11,694	0.500	0.500	0	1
$PY2000_i$	11,694	0.324	0.468	0	1
Num of ID	804	804	804	804	804

3.4 Empirical estimation

To test the hypotheses proposed above, we use four models. The first model examines the temporal effect of citations while Model 2 and Model 3 separate domestic citations from foreign citations. Using $DIFF_FC_DC_{it}$ as the dependent variable, Model 4 looks at the difference between the citations from abroad and those from author countries, and tests for the temporal influence on the difference. The models are estimated using random effect panel regression.

4. Results

The first model provides an overall picture of knowledge diffusion by looking at the trend of citation over time (Table 6). Both the linear term *Time* and the squared term *Time*² are significant, implying a curvilinear relationship between time and citation. According to the coefficients (0.189 and -0.009), citations increase steadily over time at decreasing rate, reaching a peak at 10.5 years after publications, and then decline at an increasing rate – a typical citation curve with an inverse U shape. It shows research findings in the field of physics could generate interest for ten years and become less popular afterwards. While the number of author countries and having US-based authors are not contributing to the citations in a statistically significant way, the journal impact factor appears to be highly relevant. Articles published in high impact journals tend to generate more attention and thus are more likely to receive citations. In particular, an increase of impact factor by 1 would lead to the increase of total citations by 0.667. Both PY1990 and PY2000 are negatively affecting citations. Compared with articles published in 1990 and 2000, those published in 2010 tend to have more citations (1.57 and 1.351 respectively) in each equivalent time window. The popularization of the Internet and ICTs and open access options since the 2000s facilitates knowledge transmit to a wider audience and leads to more reads and citations (Lawrence, 2001). At the same time the rapidly expanded knowledge production in recent years also contributes to the upsurge of citing articles.

Table 6. Regression output

VARIABLES	(1) TC	(2) DC	(3) FC	(4) DIFF FC DC
Time	0.189*** (0.019)	0.014* (0.009)	0.174*** (0.014)	0.029*** (0.004)
Time2	-0.009*** (0.001)	-0.002*** (0.000)	-0.008*** (0.001)	
Country_num	0.130	0.202***	-0.072	-0.275***

	(0.135)	(0.058)	(0.090)	(0.069)
IF	0.677***	0.271***	0.406***	0.136***
	(0.062)	(0.027)	(0.041)	(0.031)
US	0.731	0.713***	0.019	-0.688***
	(0.455)	(0.195)	(0.302)	(0.229)
PY1990	-1.570***	-0.766***	-0.804**	0.104
	(0.496)	(0.213)	(0.330)	(0.251)
PY2000	-1.351***	-0.552***	-0.799**	0.017
	(0.481)	(0.207)	(0.320)	(0.242)
Constant	-0.094	-0.016	-0.076	0.207
	(0.499)	(0.215)	(0.332)	(0.253)
Num of Obs	11,694	11,694	11,694	11,694
Num of ID	804	804	804	804
Wald chi ² (7)	354.29***	373.3***	309.45***	309.45***

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

The next two models (Model 2 and 3) test for the influence of time on domestic and foreign citations separately. The variable *Time* and its curvilinear term *Time*² are associated with both domestic citations and foreign citations in the same direction. Consistent with the result in Model 1, citations increase with time initially and then decline. However, the peak comes at a different time. Domestic citations reach its peak at 3.5 years after publication (based on coefficients of 0.014 and -0.002), while foreign citations reach the peak at 11 years after publication (based on coefficients of 0.174 and -0.008), with other variables being controlled (Figure 1). It shows knowledge diffuses quickly inside the country but also fades fast, while in the distant scientific community, it is slow in generating interest but the interest tends to last longer. Among the control variables, the number of author countries has positive impact on domestic citations, but is negatively yet insignificantly associated with foreign citations. Having one more country in the authorship could increase domestic citations by 0.2, which is expected as the more countries that the publication is involved, the bigger the pool of domestic scholars and hence more domestic citations. Consistent with the findings from Model 1, journal impact

factor significantly improves domestic citation counts and even more for foreign citations. The years of publication for seed articles also appear to have similar influence on citations, domestic and abroad. Articles published in 1990 received fewer citations in the same time window than those published in 2000, and even fewer than 2010. While having US-based authors does not matter for foreign citations, it is particularly effective for domestic citations. The finding is similar to results in Glänzel & Schubert (2005) and Bornmann et al (2018) showing that US scholars are more likely to cite one another compared with scholars in other countries.

Figure 1 Accumulation of domestic citations and foreign citations

Note: plot based on regression output in Table 6

The last model tests how the difference between foreign citations and domestic citations changes over time. The variable *Time* is found to have a positive and significant impact on the difference in the citations. There will be more foreign citations than domestic citations with the passage of time. With each additional year, the difference between foreign and domestic citations increases by 0.029. After the promotion period with better exposure domestically, the publication is gradually discovered by the rest of the scientific community, and the proximal advantage diminishes. The geography may still play a role, but the effect is likely to be overshadowed by

the attention from the international community, which is much larger in size. The number of countries in author affiliations is negatively associated with the difference. The more countries the publication involves, the smaller the difference is between foreign citations and domestic citations as there are more chances for the publication to be cited by scholars in the same countries. The journal impact factor is positively contributing to the gap between foreign and domestic citations. Articles that appear in top journals with high impact factors are receiving more attention from scholars around the world, which leads to higher foreign citations. By contrast, publications authored by American scholars tend to have less difference between foreign and domestic citations. It is most likely because the US produces the largest number of publications in the field and the citing articles published by researchers in the US are no less than those by other countries. The other two dummies variables controlling the year of publications – PY1990 and PY2000 – do not seem to matter for the difference in citations. It shows the wider access and early view benefit from open access options in the recent years (Kurtz et al., 2005; Moed, 2007) is shared by both domestic and foreign communities.

5. Discussion and conclusions

The paper intends to investigate the proximal advantage in knowledge diffusion in the form of publication citations and empirically test whether knowledge diffuses differently in vicinity versus distance along the timeline. Using a set of research articles in one of the basic research fields – physics, the study compares the number of citations from scholars in the domestic research communities with the citations from abroad in each subsequent year of publication. The distribution of citations shows domestic citations appear faster, which supports the hypothesis. Subsequently, using random effect panel regressions, we test the effect of time on

the accumulation of domestic and foreign citations and found the peak of domestic citations comes earlier, which again confirms the hypothesis. Knowledge diffusion is indeed affected by spatial boundaries, and the influence is stronger in the early years after publication. The finding reveals the importance of networking and exchange in knowledge dissemination. While knowledge in the form of publication can be retrieved with little access barrier, it needs time to be discovered and cannot grasp people's attention immediately. Presenting the research formally in conferences or informally in social occasions will shorten the time and generate more awareness. Interaction, networking and interpersonal ties are still important for knowledge diffusion even in the era with advanced ICTs.

The study is among the first to introduce the time dimension in examining the spatial effect of knowledge diffusion. The prevailing understanding in the current literature is that codified knowledge is less bounded by spatial constraints in its transmission (Bernard, Bureth, & Cohendet, 2000). As a form of dissemination, publication releases knowledge from the interpersonal network to broadcast, which should exhibit little geographic effect on the diffusion process. However, this expectation is not supported by some empirical studies, based on data from patent citations and collaboration networks. To complement previous research, the study provides additional empirical evidence to the spatial dimension of knowledge distribution with publication citation data. In addition, the study reveals the temporal effect in spatial distribution by comparing the different dissemination speed in and outside the country along the timeline. As a result, the study confirms the advantage of physical proximity, but adds that the spatial advantage is not constant. It is the social component in knowledge diffusion that makes geography matters, and the social component is most evident in new publications when authors are active and keen in promoting their research.

There are a few caveats with the study. Domestic citations may be inflated by self-citations (Aksnes, 2003). Authors often cite their previous publications to reflect the source of knowledge and to boost the impact of former work (Lawani, 1982). Self-citations could appear rather early as authors may simultaneously produce a series of papers that are inter-related based on their research projects and cite each other even before the official publication, which might lead to the over-estimation of the temporal effect of domestic citations. However, as pointed in Lewison (2005), self-citation still implies some knowledge transfer as it is rare for a citing paper to have identical authorship as the cited paper. Hence we still keep them in the analysis.

Secondly, the findings cannot be over generalized as this study only focuses on the field of physics. As physics is among the most internationalized fields (Girvan & Newman, 2002; Caroline S. Wagner et al., 2017), a stronger effect of national boundary might exist in other disciplines. For instance in fields of humanities and social sciences where research subjects are highly contextualized, the diffusion inside the country will be even more discernable, but this needs to be confirmed in future studies. In addition, given that the analysis is based on publication data derived from the WoS, the effect of national boundaries of knowledge spillover is likely to be under-estimated. The WoS database is biased against non-English publications (Van Leeuwen, Moed, Tijssen, Visser, & Van Raan, 2001), and research activities outside the developed world are less visible and accessible (Duque et al., 2005; Tang & Hu, 2013; Caroline S. Wagner & Wong, 2012). Most scholars in non-English-speaking countries still publish in domestic journals that are not indexed by WoS or not available digitally in full text, see for example studies of the BRIC countries (Caroline S. Wagner & Wong, 2012). The influence of the scientific work in domestic scientific community would be much larger if we were able to use publication data from all possible journals. As such, the use of WoS in fact suppressed the

national boundary effect of knowledge diffusion. In view of this limitation, our findings on the localization of knowledge spillover is not only valid but could also potentially be more significant. Lastly, this study does not take into account of country variance. Country size is found to be relevant for collaboration pattern (Mattsson, Laget, Nilsson, & Sundberg, 2008) and the domesticity of citations (Glänzel & Schubert, 2005). The effect of country size on the speed of knowledge diffusion could be a worthwhile subject for further study.

References

- Aksnes, D. W. (2003). A macro study of self-citation. *Scientometrics*, 56(2), 235–246.
- Amin, A., & Wilkinson, F. (1999). Learning, proximity and industrial performance: an introduction. *Cambridge Journal of Economics*, 23(2), 121–125.
- Anselin, L., Varga, A., & Ac, Z. (1997). Local Geographic Spillovers between University Research and High Technology Innovations. *Journal of urban economics*, 42, 422–448.
- Audretsch, D. B., & Feldman, M. P. (1996). R&D spillovers and the geography of innovation and production. *American Economic Review*, 86(3), 641–652.
- Autant-Bernard, C. (2001). Science and knowledge flows: evidence from the French case. *Research Policy*, 30(7), 1069–1078.
- Autant-Bernard, C., Mairesse, J., & Massard, N. (2007). Spatial knowledge diffusion through collaborative networks. *Papers in Regional Science*, 86(3), 341–350.
- Barber, B. (1962). Resistance by Scientists to Scientific Discovery. In B. B. a. W. Hirsch (Ed.), *The Sociology of Science*. New York Free Press.
- Bernard, A., Bureth, A., & Cohendet, P. (2000). The economics of knowledge: the debate about codification and tacit knowledge. *Industrial and corporate change*, 9(2), 255–287.
- Bollen, J., Rodriguez, M. A., & Van De Sompel, H. (2006). Journal status. *Scientometrics*, 69(3), 669–687.
- Börner, K., Penumarthy, S., Meiss, M., & Ke, W. (2006). Mapping the diffusion of scholarly knowledge among major U.S. research institutions. *Scientometrics*, 68(3), 415–426.
- Bornmann, L., & Leydesdorff, L. (2017). Skewness of citation impact data and covariates of citation distributions: A large-scale empirical analysis based on Web of Science data. *Journal of Informetrics*, 11(1), 164–175.
- Bornmann, L., & Mutz, R. (2015). Growth rates of modern science: A bibliometric analysis based on the number of publications and cited references. *Journal of the Association for Information Science and Technology*, 66(11), 2215–2222.
- Bornmann, L., Wagner, C., & Leydesdorff, L. (2018). The geography of references in elite articles: What countries contribute to the archives of knowledge. *PloS one*, 13(3), e0194805.
- Boschma, R. (2005). Proximity and innovation: A critical assessment. *Regional Studies*, 39(1), 61–74.
- Bottazzi, L., & Peri, G. (2003). Innovation and spillovers in regions: Evidence from European patent data. *European Economic Review*, 47(4), 687–710.
- Bouba-Olga, O., Carrincazeaux, C., Coris, M., & Ferru, M. (2015). Editorial: Proximity Dynamics, Social Networks and Innovation. *Regional Studies*, 49(6), 901–906.
- Breschi, S., Lenzi, C., Lissoni, F., & Vezzulli, A. (2010). The geography of knowledge spillovers: The role of inventors' mobility across firms and in space. In R. Boschma & R. Martin (Eds.), *The handbook of evolutionary economic geography* (pp. 353–369). Mass, USA: Edward Elgar Publishing.
- Breschi, S., & Lissoni, F. (2005). "Cross-Firm" Inventors and Social Networks: Localized Knowledge Spillovers Revisited *Annales d'Économie et de Statistique*, 79/80, 189–209.
- Chen, C., & Hicks, D. (2004). Tracing knowledge diffusion. *Scientometrics*, 59(2), 199–211.
- Choi, S. (2012). Core-periphery, new clusters, or rising stars?: international scientific collaboration among 'advanced' countries in the era of globalization. *Scientometrics*, 90(1), 25–41.
- Coccia, M., & Wang, L. (2016). Evolution and convergence of the patterns of international scientific collaboration. *PNAS*, 113(8), 2057–2061.
- Cohen, W. M., Nelson, R. R., & Walsh, J. P. (2002). Links and Impacts: The Influence of Public Research on Industrial R&D. *Management Science*, 48(1), 1–23.
- Cole, S. (1970). Professional Standing and the Reception of Scientific Discoveries. *American Journal of Sociology*, 76(2), 286–306.

- Collins, H. (1985). *Changing Order: Replication and Induction in Scientific Practice*. Chicago: University of Chicago Press.
- Crane, D. (1969). Social Structure in a Group of Scientists: A Test of the "Invisible College" Hypothesis. *American Sociological Review*, 34(3), 335-352.
- Crane, D. (1972). *Invisible Colleges: Diffusion of Knowledge in Scientific Communities*. Chicago, IL: University of Chicago Press.
- Dasgupta, P., & David, P. A. (1994). Toward a new economics of science. *Research Policy*, 23(5), 487–521.
- Didegah, F., & Thelwall, M. (2013). Determinants of Research Citation Impact in Nanoscience and Nanotechnology. *Journal of the American Society for Information Science and Technology*, 64(5), 1055–1064.
- Duque, R. B., Ynalvez, M., Sooryamoorthy, R., Mbatia, P., Dzorgbo, D. B. S., & Shrum, W. (2005). Collaboration paradox: Scientific productivity, the Internet, and problems of research in developing areas. *Social Studies of Science*, 35(5), 755–785.
- Frenken, K., Hoekman, J., Kok, S., Ponds, R., Oort, F. v., & Vliet, J. v. (2009). Death of Distance in Science? A Gravity Approach to Research Collaboration. In A. Pyka & A. Scharnhorst (Eds.), *Innovation Networks: New Approaches in Modeling and Analyzing*: Springer-Verlag Berlin Heidelberg.
- Garfield, E. (1979). *Citation indexing: Its theory and application in science, technology, and humanities*. New York: John Wiley & Sons.
- Garfield, E. (1980). Premature discovery or delayed recognition—Why. *Current Contents*, 21, 5–10.
- Gilsing, V. (2005). *The Dynamics of Innovation and Interfirm Networks: Exploration, Exploitation and Co-evolution*. Cheltenham: Edward Elgar.
- Gilsing, V., Nooteboom, B., Vanhaverbeke, W., Duysters, G., & Oord, A. v. d. (2008). Network embeddedness and the exploration of novel technologies: Technological distance, betweenness centrality and density. *Research Policy*, 37(10), 1717–1731.
- Gingras, Y., & Khelifaoui, M. (2018). Assessing the effect of the United States’ “citation advantage” on other countries’ scientific impact as measured in the Web of Science (WoS) database *Scientometrics*, 114(2), 517–532.
- Girvan, M., & Newman, M. E. (2002). Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, 99(12), 7821–7826.
- Glänzel, W., & Schubert, A. (2005). Domesticity and internationality in co-authorship, references and citations. *Scientometrics*, 65(3), 323-342.
- Goldfinch, S., Dale, T., & Derouen, K. J. (2003). Science from the periphery: Collaboration, networks and Periphery Effects in the citation of New Zealand Crown Research Institutes articles, 1995-2000. *Scientometrics*, 57(3), 321-337.
- Hoekman, J., Frenken, K., & Tijssen, R. J. W. (2010). Research collaboration at a distance: Changing spatial patterns of scientific collaboration within Europe. *Research Policy*, 39(5), 662–673.
- Jaffe, A. B. (1989). Real effects of academic research. *American Economic Review*, 79(5), 957–970.
- Jaffe, A. B., & Trajtenberg, M. (1996). Flows of knowledge from universities and federal laboratories: Modeling the flow of patent citations over time and across institutional and geographic boundaries. *Proceedings of the National Academy of Sciences*, 93(23), 12671–12677.
- Jaffe, A. B., & Trajtenberg, M. (1999). International knowledge flows: evidence from patent citations. *Economics of Innovation and New Technology*, 8(1-2), 105-136.
- Jaffe, A. B., Trajtenberg, M., & Henderson, R. (1993). Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations. *The Quarterly Journal of Economics*, 108(3), 577-598.
- Kato, M., & Ando, A. (2003). The relationship between research performance and international collaboration in chemistry. *Scientometrics*, 97(3), 535–553.

- Katz, J. S. (1994). Geographical proximity and scientific collaboration. *Scientometrics*, 31(1), 31-43.
doi:10.1007/bf02018100
- Kurtz, M. J., Eichhorn, G., Accomazzi, A., Grant, C., Demleitner, M., Henneken, E., & S.Murray, S. (2005). The effect of use and access on citations. *Information Processing & Management*, 41(6), 1395-1402.
- Lawani, S. M. (1982). On the heterogeneity and classification of author self-citations. *Journal of the American Society for Information Science and Technology*, 33(5), 281-284.
- Lawrence, S. (2001). Free online availability substantially increases a paper's impact. *Nature*, 411, 521.
- LeSage, J. P., Fischer, M. M., & Scherngell, T. (2007). Knowledge spillovers across Europe: Evidence from a Poisson spatial interaction model with spatial effects. *Papers in Regional Science*, 86(3), 393-422.
- Lewison, G., Rippon, I., & Wooding, S. (2005). Tracking knowledge diffusion through citations. *Research Evaluation*, 14(1), 5-14.
- Leydesdorff, L. (1998). Theories of citations? *Scientometrics*, 43(1), 5-25.
- Liu, Y., & Rousseau, R. (2010). Knowledge diffusion through publications and citations: A case study using ESI - fields as unit of diffusion. *Journal of the American Society for Information Science and Technology*, 61(2), 340 - 351
- Lozano, G. A., Larivière, V., & Gingras, Y. (2012). The weakening relationship between the impact factor and papers' citations in the digital age. *Journal of the Association for Information Science and Technology*, 63(11), 2140-2145.
- Matthiessen, C. W., Schwarz, A. W., & Find, S. (2002). The Top-level Global Research System, 1997-99: centres, networks and nodality. An Analysis Based on Bibliometric Indicators. *Urban Studies*, 39(5-6), 903-927.
- Mattsson, P., Laget, P., Nilsson, A., & Sundberg, C.-J. (2008). Intra-EU vs. extra-EU scientific co-publication patterns in EU. *Scientometrics*, 75(3), 555-574.
- Meyer, M. (2000). What is special about patent citations? Differences between scientific and patent citations. *Scientometrics*, 49(1), 93-123.
- Milard, B. (2014). The social circles behind scientific references: Relationships between citing and cited authors in chemistry publications. *Journal of the Association for Information Science and Technology*, 65(12), 2459-2468.
- Moed, H. F. (2007). The Effect of "Open Access" on Citation Impact: An Analysis of ArXiv's Condensed Matter Section. *Journal of the American Society for Information Science and Technology*, 58(13), 2047-2054.
- Narin, F., Stevens, K., & Whitlow, E. S. (1991). Scientific co-operation in Europe and the citation of multinationally authored papers. *Scientometrics*, 21(3), 313-323.
- Onodera, N., & Yoshikane, F. (2015). Factors Affecting Citation Rates of Research Articles. *Journal of the Association for Information Science and Technology*, 66(4), 739-764.
- Pan, R. K., Kaski, K., & Fortunato, S. (2012). World citation and collaboration networks: uncovering the role of geography in science. *Scientific Reports*, 2(902), doi:10.1038/srep00902.
- Perneger, T. V. (2010). Citation analysis of identical consensus statements revealed journal-related bias. *Journal of Clinical Epidemiology*, 63(6), 660-664.
- Polanyi, M. (1958). *Personal Knowledge: Towards a Post-Critical Philosophy*. Chicago: University of Chicago Press.
- Polanyi, M. (1966). *The Tacit Dimension*. Chicago, IL University Of Chicago Press.
- Ponds, R., Van Oort, F., & Frenken, K. (2007). The geographical and institutional proximity of research collaboration. *Papers in Regional Science*, 86(3), 423-443.

- Ponds, R., van Oort, F., & Frenken, K. (2010). Innovation, spillovers and university–industry collaboration: an extended knowledge production function approach. *Journal of Economic Geography*, 10(2), 231–255.
- Price, D. J. d. S. (1963). *Little Science, Big Science*. New York: Columbia University Press.
- Price, D. J. d. S. (1965). Networks of scientific papers. *Science*, 149(3683), 510–515.
- Schmoch, U., & Schubert, T. (2008). Are international co-publications an indicator for quality of scientific research? *Scientometrics*, 74(3), 361–377.
- Sorenson, O., & Fleming, L. (2004). Science and the diffusion of knowledge. *Research Policy*, 33(10), 1615–1634.
- Tang, L., & Hu, G. (2013). Tracing the Footprint of Knowledge Spillover: Evidence From U.S.–China Collaboration in Nanotechnology. *Journal of the American Society for Information Science and Technology*, 64(9), 1791–1801.
- Tang, L., Shapira, P., & Youtie, J. (2015). Is There a Clubbing Effect Underlying Chinese Research Citation Increases? *Journal of the Association for Information Science and Technology*, 66(9), 1923–1932.
- Ter Wal, A. L. J. (2014). The dynamics of the inventor network in German biotechnology: geographic proximity versus triadic closure. *Journal of Economic Geography*, 14, 589–620.
- Van Leeuwen, T. N., Moed, H. F., Tijssen, R. J. W., Visser, M. S., & Van Raan, A. F. J. (2001). Language biases in the coverage of the Science Citation Index and its consequences for international comparisons of national research performance. *Scientometrics*, 51(1), 335–346.
- Wagner, C. S., & Leydesdorff, L. (2005). Mapping the network of global science comparing international co-authorships from 1990 to 2000. *International Journal of Technology and Globalization*, 1(2), 185–208.
- Wagner, C. S., Whetsell, T. A., & Leydesdorff, L. (2017). Growth of international collaboration in science: revisiting six specialties. *Scientometrics*, 110(3), 1633–1652.
- Wagner, C. S., & Wong, S. K. (2012). Unseen science? Representation of BRICs in global science. *Scientometrics*, 90(3), 1001–1013.
- Wallace, M. L., Larivière, V., & Gingras, Y. (2009). Modeling a century of citation distributions. *Journal of Informetrics*, 3(4), 296–303.
- Wang, J., & Shapira, P. (2015). Is There a Relationship between Research Sponsorship and Publication Impact? An Analysis of Funding Acknowledgments in Nanotechnology Papers. *PloS one*, 10(2), e0117727. DOI: 0117710.0111371/journal.pone.0117727.
- Wang, J., Veugelers, R., & Stephane, P. (2017). Bias against novelty in science: A cautionary tale for users of bibliometric indicators. *Research Policy*, 46(8), 1416–1436.
- White, H. D., Wellman, B., & Nazer, N. (2004). Does Citation Reflect Social Structure? Longitudinal Evidence From the “Globenet” Interdisciplinary Research Group. *Journal of the American Society for Information Science and Technology*, 55(2), 111–126.
- Wuestman, M., Hoekman, J., & Frenken, K. (2018). *The Complex Geography of Scientific Knowledge Accumulation*. Paper presented at the DRUID Academy Conference, Odense, Denmark.
- Yan, E. (2014). Finding knowledge paths among scientific disciplines. *Journal of the Association for Information Science and Technology*, 65(11), 2331–2347.
- Yan, E., & Sugimoto, C. R. (2011). Institutional interactions: Exploring social, cognitive, and geographic relationships between institutions as demonstrated through citation networks. *Journal of the American Society for Information Science and Technology*, 62(8), 1498–1514.
- Zitt, M., & Small, H. (2008). Modifying the Journal Impact Factor by Fractional Citation Weighting: The Audience Factor. *Journal of the American Society for Information Science and Technology*, 59(11), 1856–1860.