WORKING PAPER

Planes, Trains and Automobiles: The Importance of Location for Knowledge Transfer in the Transportation Sector

by
Jeffrey Moore, Daniel K.N Johnson & Kristina M. Lybecker

Colorado College Working Paper 2012-10
June, 2012

Department of Economics and Business
Colorado College
Colorado Springs, Colorado 80903-3298
www.coloradocollege.edu/dept/EC

Planes, Trains and Automobiles: The Importance of Location for Knowledge Transfer in the Transportation Sector

Jeffrey A. Moore, Daniel K.N. Johnson* and Kristina M. Lybecker

Jeffrey Moore is the Technical and Statistical Coordinator of Colorado College's Department of Economics and Business, 14 East Cache La Poudre Street, Colorado Springs, CO 80903, tel: (719) 389-6447, fax: (719) 389-6927, email: jeffrey.moore@Colorado College.edu

Daniel K.N. Johnson is the Gerald L. Schlessman Professor of Economics at Colorado College, 14 East Cache La Poudre Street, Colorado Springs, CO 80903, tel: (719) 389-6654, fax: (719) 389-6927, email: djohnson@ColoradoCollege.edu. *Please direct correspondence to this co-author.

Kristina M. Lybecker is an Assistant Professor of Economics at Colorado College, 14 East Cache La Poudre Street, Colorado Springs, CO 80903, tel: (719) 389-6445, fax: (719) 389-6927, email: Kristina.Lybecker@ColoradoCollege.edu.

Special thanks to a team led by Nicole Gurley for their patience while geo-coding, to Matt Gottfried for his continued patience with our GIS-based questions, and generous support from Colorado College in the forms of the Schlessman Professorship, a Faculty Grant, a Mrachek Research Fellowship, a Mellon Research Block and assistance from Colorado College's Chapman Fund.

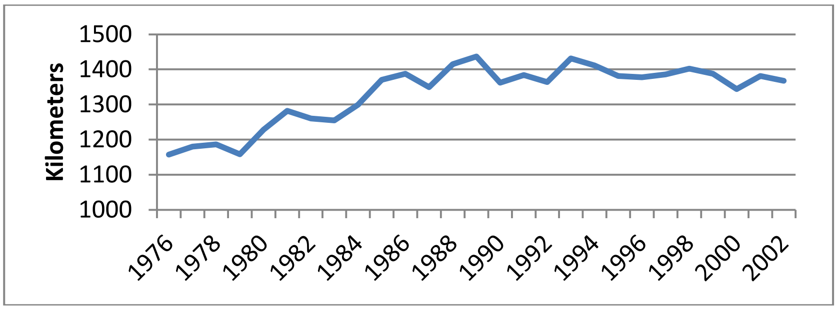
Abstract: Using over 200,000 U.S. patent citations, we test whether knowledge transfers in the transportation sector are sensitive to distance, and whether that sensitivity has changed over time. Controlling for self-citation by inventor, assignee and examiner, multivariate regression analysis shows that physical distance is becoming less important for spillovers with time, albeit in a nonlinear fashion.

Keywords: transportation, patent, citation, spillover, distance

Historically, firms within an industry have often clustered geographically, due to localization economies (Henderson (1986) or Smith and Florida (1994)) such as the speed (see Caballero and Jaffe (1993) and tacitness of the transfer of knowledge (see Von Hippel (1994)). To our knowledge, however, this phenomenon has not been measured within the transportation sector, nor has its importance been compared over time for this sector.

This paper examines knowledge flows within the transportation industry, both confirming traditional evidence that inter-firm knowledge transfers typically decreased with distance, and documenting their increase over time (albeit at a decreasing rate). Figure 1 portrays the pattern without any statistical correction for other potentially correlating factors.

Figure 1: Average citation distance (in kilometers)



In section 2 of the paper, we briefly review the relevant literature on the geographic nature of knowledge spillovers. Section 3 describes our data set, designed for compatibility with the literature, and Section 4 presents multivariate regression analysis that controls for non-geographic effects in presenting the declining role of distance. Section 5 concludes with implications for policy and further research.

2. Literature review

Most of the literature suggests that knowledge spillovers cluster geographically, with higher spillovers (shown by more patent citations) within a short distance. The underlying supposition is that inventors are more aware of (or find more use for) inventions located close to them, and therefore build more heavily on local inventions. Alternatively, it might be a function of where innovations actually occur, smaller distances if the industry clusters around a central hub. The result is a geographic clustering of citations.

Empirical evidence stresses the role of geography in the transfer of knowledge within an innovation network (see, for example, a review by Gelsing, 1992), and emphasizes the importance of frequent personal contact and research collaboration. Lundvall (1992) points out that the importance of geography should differ predictably by technology type. While geography has little impact on stationary technologies (which face constant needs and opportunities), its importance grows quickly for technologies undergoing incremental, and particularly, radical innovation.

Geographic proximity has already been used to explain the location choices of R&D-intensive activities (e.g. Dorfman (1988)). However, the location of *firms* is not always a good predictor of the location of *innovation* (Feldman (1994), Johnson and Brown, (2004)).

Localization of patent citations has been firmly established (Jaffe et al., 1993), with a random sample of patents clearly more likely to cite local patents than patents by parties that are located farther away, at every geographically aggregated level. The effects are more intense where knowledge becomes obsolescent rapidly, such as electronics, optics, and nuclear technology (Jaffe and Trajtenberg, 1996). Alternatively, strong voices in the literature argue that either distance has never mattered much (Thompson and Fox-Kean, 2005), or that the impact of communication technology on productivity or on knowledge transmission across distance will not be that great (Vasileiadou and Vliegenthart, 2009; Graham, 2001).

3. Data

3.1. Measurement issues

This paper relies exclusively on patent citations from transportation-sector patents as a geographic measure of knowledge spillovers in the sector. When a patent application is submitted for approval, it is accompanied by a list of citations to other patents and literature which have been instrumental in the creation of this technology, or which delineate the legal limits of this application. The intention is twofold: to build a convincing case that this application is novel and unobvious to someone trained in the field, and to provide a legal record in order to protect patent rights in the future. To this list of citations, a patent examiner may suggest his or her own list of citations for the applicant to include. The result is a paper trail of knowledge creation.

Of course, patents records do not perfectly reflect the creation of technology, as some innovations are never patented and patents vary greatly in size and importance (Feldman (1994)). Likewise, citations do not perfectly reflect the transfer of knowledge, as they may be inserted for a variety of other reasons including legal protection or examiner privilege (Jaffe et al. (2000)). However, their statistical tests indicate that overall citations can be interpreted as a signal of spillovers, albeit a noisy signal.

In order to define the scope of the study, we follow the World Intellectual Property Organization's definition for transportation, definitions which include portions of 5 separate International Patent Classification (IPC) category codes on the 4-digit level: B60 through B64 inclusive. We appended our dataset with the set of patents cited by transportation patents, at least those that were themselves granted between 1976 and 2002. While the citing patents date exclusively from this period, patents cited by our observation sample may predate 1976, but were truncated from consideration simply due to data availability. Furthermore, citing and cited patents from all non-U.S. inventors have been excluded, for reasons of feasibility. However, Johnson and Sneed (2009) present evidence in the literature that international citations are increasing in frequency across a host of technologies, evidence which is sympathetic to the hypothesis here that citation distances have been increasing over time (Johnson and Lybecker, 2011).

Our focus is thus exclusively upon transportation-sector knowledge flows within the United States, within a banded period of time, inviting subsequent scholars to continue to work of expanding the dataset's coverage.

3.2. Clustering of knowledge citations

Patent citations may cluster for non-geographic reasons, coincidentally causing a pattern that appears geographic merely through correlation with other phenomena. For example, inventors may be more familiar with their own patents, and the same may be true of assignees (the legal holders of the intellectual property rights) if employees of a firm are familiar with other patents held by the same firm. Therefore we include self-citations in the analysis but control for them separately.

Using U.S. patent data from a combination of sources (NBER website as described in Hall et al., 2001, in addition to raw data collected by the independent firm MicroPatent), this paper relies on citations collected from all transportation-sector patents granted between 1976 and 2002. Each citation's endpoints (citing and cited) were then geo-coded for the primary location of each U.S.-based patent at the geographic center of the city listed (as specific addresses are available for less than ten percent of all patent documents).

The result is a dataset of 203,158 U.S.-based transportation patent documents that include a total of 263,796 citations to other U.S.-based patent documents. Previous literature (e.g. Johnson and Lybecker, 2011) indicates that each of the following factors may play some role in the distance of a citation, so this research measured each for every observed citation:

- whether patents *k* and *K* have the same inventor (SI),
- whether patents *k* and *K* have the same assignee (SA),
- whether patents *k* and *K* are in the same technology cluster (ST),
- how similar the citing and cited states are in technology types (SC),
- whether the cited patent is also classified as transportation (T),
- whether the assignee is a government agency (G),
- whether the assignee is an educational institution (U),

- how old the citation is, in years between citing and cited patent (A), along with its squared term to account for the potentially nonlinear effects of age (A^2), and
- year of citing patent *K*, to account for citation inflation (Y).

Self-citation by inventors accounted for almost three percent of all citations from transportation patents, a much lower self-citation rate than has been documented in other sectors like biotechnology (twelve percent, according to Johnson and Lybecker, 2011). On the other hand, thirty-four percent of all citations were to the same assignee, higher than has been measured in biotechnology (twenty-four percent, according to Johnson and Lybecker, 2011).

It is also possible that patents closer in technology space may have longer or shorter citation distances than more diverse cited patents. The data are coded so that a binary variable, ST, indicates whether the International Patent Classification (IPC) system places both citing and cited patents in the same technology cluster, out of 634 mutually exclusive and exhaustive clusters in the system. This system, in global use since 1975, is the standard by which all patents are organized (and thus assigned to examiners for processing, or searched by inventors and lawyers to establish claims).

The fourth variable, the technological correlation between citing and cited states (SC), is included for a similar reason. Each state's technological profile was calculated as the share of patent activity assigned to each of the 634 IPC technology classes. Pair-wise correlations between state vectors then provide a measure of technological similarity between locations. The analysis also includes an indicator of whether the cited patent is classified as transportation.

Because government (G) and university (U) patents may cite knowledge differently than do private sector patents, we include those indicators as controls as well. Linear and squared age terms are included to accommodate nonlinear effects for older knowledge. Finally, since the goal of the analysis is to test whether distance changes over time, it is necessary to include dummy variables for each time period (year Y).

Summary statistics of the 203,158 citations in the sample are presented in Table 1. The average citation is around 1375 kilometers long, but the distribution has a very wide variance. For this reason, tests are performed not only using distance as a dependent variable, but on alternative specifications using a dependent variable of the logarithm of distance.

Table 1: Summary statistics for citation dataset

Variable	Mean	Std. Dev.	Min	Max
Citation distance (km)	1375.48	1272.63	0	12565.15
SC= correlation in technology portfolios between cited and citing states	0.84	0.15	0.01	1
A= age of citation (years)	8.38	5.91	0	27
	Number of zeros			
SI = same inventor in cited and citing patents	196846 (97.0%)		0	1
SA = same assignee in cited and citing patents	133889 (66.0%)		0	1
ST = same IPC in cited and citing patents	98900 (48.7%)		0	1
T = cited patent is also Transportation	63470 (31.2%)		0	1
G = government assignee	200,595 (98.7%)		0	1
U = university assignee	202,030 (99.4%)		0	1

4. Model and results

Our regression analysis follows the literature (Johnson and Lybecker, 2011) in using a simple model by Petersen and Rajan (2002) with the patent citation as the unit of analysis. The model expresses the distance between a cited patent k granted in year y and a subsequent citing patent, K, granted in year Y, as a function of the attributes of patents k and K:

$$\delta_{k,K} = \alpha(k,K) + \varepsilon \qquad \textbf{(1)}$$

where and $\delta_{k,K}$ represents the distance between patents k and K, $\alpha(k,K)$ is a vector of the other attributes of patents k and K that affect the probability of citation, and ε is a randomly distributed error term. We postulate a reduced functional form, using the log of [distance plus one] in order to avoid taking the log of a zero distance) because the fit of the equation is better:

$$\mathrm{Distance} = \alpha_0 + \alpha_{\mathrm{SameAssignee}}\mathrm{SA} + \alpha_{\mathrm{SameInventor}}\mathrm{SI} + \alpha_{\mathrm{SameTech}}\mathrm{ST} + \alpha_{\mathrm{Transportation}}\mathrm{T} + \alpha_{\mathrm{Govt}}\mathrm{G} + \alpha_{\mathrm{University}}\mathrm{U} + \alpha_{\mathrm{Age}}\mathrm{A} + \alpha_{\mathrm{AgeSquared}}\mathrm{A}^2 + \sum_{1976}^{2002}\alpha_{\mathrm{Year=i}}\mathrm{Y_i} + \varepsilon_{\mathrm{K}} + \mathrm{u} \qquad \textbf{(2)}$$

We include a fixed effect specific to the citing patent (ε_K), since there are presumably immeasurable factors which might dictate a longer or shorter average citation distance.

Table 2 presents multivariate regression Tobit estimates (left-censored for intra-city citations with a distance of 0 kilometers), with White-corrected errors, using fixed effects. Considering all citations, the average distance unambiguously increases with time, with strong evidence of a non-linear pattern. Sensitivity tests find very similar results if we restrict our consideration to citations of more than 10 kilometers, more than 50 kilometers, or more than 100 kilometers. Results for citations spanning more than 100 kilometers (i.e., excluding short and intra-city citations) are presented in the second panel of Table 2, and tell a similar story.

To permit maximum flexibility to these nonlinearities, and potential nuances in particular years, we performed the same analyses using separate year indicator variables instead of a time trend variable.

Unsurprisingly, citations with the same assignee or same inventor are more likely to be proximate than are other citations, an effect is especially strong and significant for inventors. Citations within the same technology class are closer to each other than more dissimilar patents (the ST coefficient is negative), and states that have similar technology portfolios tend to be close together, a fact captured by the negative coefficient on that variable (SC).

Table 2: Tobit weighted regressions on log(distance+1)

	All citations		Only citations with distance>100km	
variable	**coefficient**	**t-statistic**	**coefficient**	**t-statistic**
SA	-0.41	(-30.08)***	0.18	(26.51)***
SI	-3.57	(-66.99)***	-0.59	(-12.69)***
ST	-0.29	(-18.26)***	-0.06	(-6.61)***
SC	-6.36	(-130.57)***	-1.33	(-56.03)***
T	0.14	(8.11)***	0.06	(6.5)***
G	0.16	(3.41)***	0.23	(8.53)***
U	0.34	(5.04)***	0.22	(5.63)***
A	6.96×10^{-2}	(19.46)***	8.60×10^{-3}	(4.54)***
A^2	-1.98×10^{-3}	(-13.42)***	-2.86×10^{-4}	(-3.56)***
Y	1.23×10^{-2}	2(.49)**	1.58×10^{-2}	(6.14)***
Y^2	-3.50×10^{-4}	(-2.42)**	-4.35×10^{-4}	(-5.83)***
constant	11.38	(212.22)***	7.92	(275.75)***
F-stat		(2583.47)***		(418.31)***
Observations		203158		165647

Notes: *** indicates 99% confidence, ** 95% confidence, * 90% confidence. Implicit impacts are calculated at the sample mean for the group in question.

Citations that cite other transportation patents average a slightly longer distance than their peers. Apparently distance matters less for the transfer of purely transportation-related knowledge than it matters for the transfer of non-transportation innovations into the transportation sector.

Government-assigned knowledge tends to travel longer transmission distances for the knowledge they cite, a result that is more pronounced when we consider only long-distance (>100 km) citations. Academic patents tend to be longer than their peers as well, but that effect is less pronounced among long-distance citations (>100 km). Finally, age matters; older citations travel longer distances, an effect which other studies (e.g. Johnson and Popp, 2003) have confirmed.

The results point to the fact that physical distance has become less of a constraint with the passage of time. Perhaps the trend is due to the nature of the knowledge being created, but we suspect that it is more due to advances in communication, which allows easier transmission of information across great distances (e.g. Kim et al., 2006). In short, the principles underlying the inter-firm transfer of knowledge are changing in a striking fashion, making spillovers easier and longer than ever before.

5. Conclusion and Policy Implications

We are left with a striking picture of the inter-firm transfer of transportation technology knowledge. Controlling for other factors, knowledge flows historically diminished with physical distance, but the importance of distance has been receding with time. That is, knowledge is more likely to transfer over longer distances now than it was twenty years ago. Long-distance knowledge transfers are increasingly the norm in transportation technology. Innovation has become possible at a wider array of locations, potentially drawing on a wider range of raw materials and ideas. This might imply a potential for the deliberate fostering of non-traditional locations for transportation technology, with a prerequisite of vibrant communication with the research community elsewhere.

6. References

Caballero, R.J. and A.B. Jaffe (1993), “How High are the Giants’ Shoulders: An Empirical Assessment of Knowledge Spillovers and Creative Destruction in a Model of Economic Growth,” in Olivier J. Blanchard and Stanley Fischer, eds., NBER Macroeconomics Annual 1993, MIT Press, Cambridge, MA.

Dorfman, N.S. (1988) “Route 128: The Development of a Regional High Technology Economy”. The Massachusetts Miracle: High Technology and Economic Revitalization. D. Lampe, editor. MIT Press, Cambridge.

Feldman, M.P. (1994) The Geography of Innovation. Kluwer Academic Publishers, Dordrecht, Netherlands.

Gelsing, L. (1992) "Innovation and Development of Industrial Networks". National Systems of Innovation: Towards a Theory of Innovation and Interactive Learning. B. Lundvall, editor. Pinter Publishers, London.

Graham, A. (2001) "The Assessment: Economics of the Internet". Oxford Review of Economic Policy, Vol. 17(2), p. 145-158.

Hall, B.H., A.B. Jaffe and M. Trajtenberg (2001) "The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools". NBER Working Paper #8498.

Henderson, J.V. (1986) "Efficiency of Resource Usage and City Size". Journal of Urban Economics, Vol 19, p.47-70.

Jaffe, A.B., M. Trajtenberg and R. Henderson (1993) "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations". Quarterly Journal of Economics, Vol 108, p.577-598.

Jaffe, A.B. and M. Trajtenberg (1996) "Flows of Knowledge From Universities and Federal Labs: Modeling the Flow of Patent Citations Over Time and Across Institutional and Geographic Boundaries". NBER Working Paper #5712.

Jaffe, A.B., M. Trajtenberg and M.S. Fogarty (2000) "Knowledge Spillovers and Patent Citations: Evidence from a Survey of Inventors". American Economic Review, Vol 90, p.215-218.

Johnson, D.K.N. and A. Brown (2004) "How the West Has Won: Regional and Industrial Inversion in U.S. Patent Activity". Economic Geography, Vol 80, No 3, p.241-260.

Johnson, D.K.N. and D.C. Popp (2003) "Forced Out of the Closet: The Impact of the American Inventors Protection Act on the Timing of Patent Disclosure". Rand Journal of Economics 34(1), p. 96-112.

Johnson, D.K.N. and K.M. Lybecker (2011) "Does distance matter less now? The changing role of geography in biotechnological innovation". Review of Industrial Organization, Vol 40, No 1, Fall 2011, p.21-35.

Johnson, D.K.N. and K.A. Sneed (2009) "Are Many Heads Better Than Two? Recent Changes in International Technological Collaboration". Journal of Law, Ethics and Intellectual Property, Vol 3, No 1, June 2009, p.1-9.

Kim, E.H., A. Morse and L. Zingales (2006) "Are Elite Universities Losing their Competitive Edge?" CEPR Discussion Paper 5700, June.

Lundvall, B.A. (1992) "User-Producer Relationships, National Systems of Innovation and Internationalisation". National Systems of Innovation: Towards a Theory of Innovation and Interactive Learning. B. Lundvall, editor. Pinter Publishers, London.

Petersen, M.A. and R.G. Rajan (2002) "Does Distance Still Matter? The Information Revolution in Small Business Lending". The Journal of Finance, Vol. 57, No. 6, p. 2533-2570

Smith, D. and R. Florida (1994) "Agglomeration and Industry Location: An Econometric Analysis of Japanese-Affiliated Manufacturing Establishments in Automotive Related Industries". Journal of Urban Economics, Vol 36, p. 23-41.

Thompson, P. and M. Fox-Kean, "Patent Citations and the Geography of Knowledge Spillovers: A reassessment". American Economic Review, Vol 95, p. 450-460.

Vasileiadou, E. and R. Vliegenthart (2009) "Research Productivity in the Era of the Internet Revisited". http://ssrn.com/abstract=1431457

Von Hippel, E. (1994) "Sticky Information and the Locus of Problem-Solving". Management Science, Vol 40, p. 429-439.