WORKING PAPER

Sure, But Who Has the Energy? The Importance of Location for Knowledge Transfer in the Energy Sector

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

Colorado College Working Paper 2011-14
November, 2011

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

Sure, But Who Has the Energy?

The Importance of Location for Knowledge Transfer in the Energy Sector

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

Abstract

Abstract: Using over 250,000 U.S. patent citations, we test whether knowledge transfers in the energy 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.

Keywords: energy, patent, citation, spillover, distance

JEL codes: L6 --- Industry Studies: Manufacturing
N9 --- Regional and Urban History
O3 --- Technological Change
R1 --- Urban, Rural and Regional Economics

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@ColoradoCollege.edu

Daniel K.N. Johnson is the Gerald L. Schlessman Professor of Economics and Director of the Innovative Minds Program 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.

<u>Sure, But Who Has the Energy?</u>

<u>The Importance of Location for Knowledge Transfer in the Energy Sector</u>

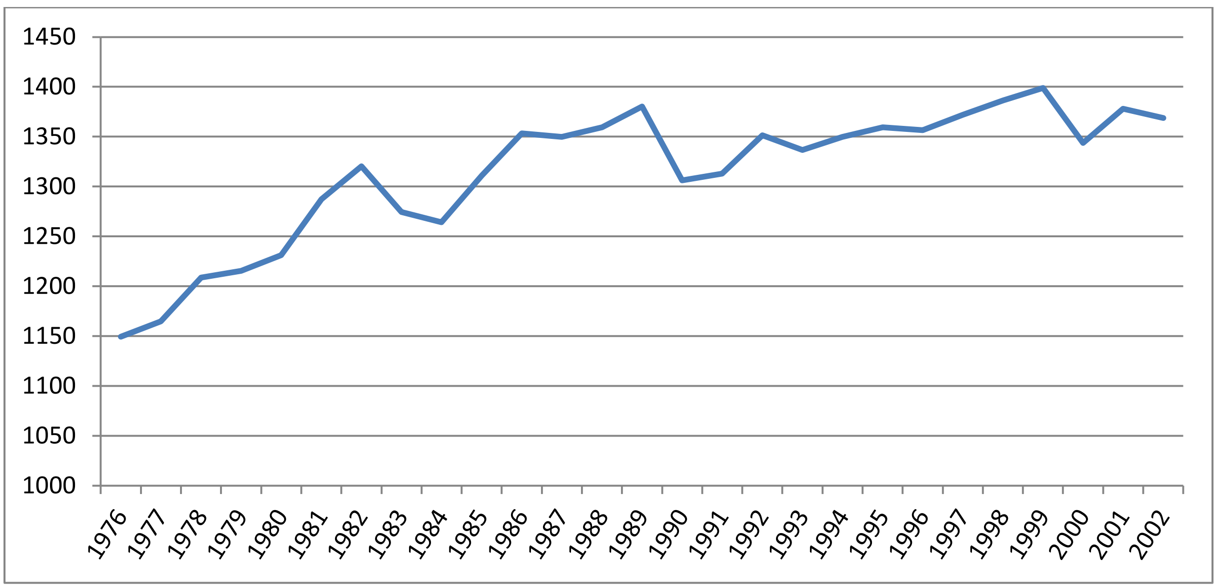
It has long been noted that firms within an industry often cluster geographically, for a variety of reasons. Localization economies, or Marshall-Arrow-Romer externalities that reduce the cost of inputs to firms in the local industry, have been studied in a variety of contexts (see for example Henderson (1986) or Smith and Florida (1994)). For some industries, it is the nature of the knowledge itself, its speed of progress, and the degree to which it is tacit that encourages firms to locate near other firms in the sector. Both of these aspects of knowledge -- diffusion speed and tacitness -- have been identified in the literature (see Caballero and Jaffe (1993) for diffusion speed and Von Hippel (1994) for tacitness); but to our knowledge no analysis has tested this effect within the energy sector.

This paper examines knowledge flows within the energy industry, confirming the traditional evidence that inter-firm knowledge transfers decrease with distance, while adding the important caveat that the impact of physical distance has been diminishing over time. Thus, historically there has been a knowledge spillover-based reason for energy-based firms to cluster geographically, although that reason is less powerful now than ever before.

Using all energy-related patents granted in the U.S. between 1976 and 2002, we show that there is a marked tendency to cite patents from nearby areas. Thus, this paper follows the "paper trail" literature (Feldman, 1999) on inter-firm knowledge transfers or "spillovers" in documenting the creation of knowledge via patent citations. We test whether this pattern could naturally arise from a tendency to cite other patents listing the same inventor, the same firm assignee, the same patent examiner, or the same technology class. We conclude that the geographic clustering of citations holds over and above the effects of these other factors, suggesting that there is a local nature to knowledge spillovers (at least insofar as patent citations reflect knowledge flows).

Cursory investigation suggests that the importance of distance has perhaps weakened over time. As Figure 1 shows, the average distance between a citing patent and its bibliographic references has grown by over 200 miles, or twenty percent, over the observation period between 1976 and 2002. While the paper's analysis controls for other factors that have changed over time, the fundamental pattern remains.

Figure 1: Average citation distance (in kilometers)



In section 2 of the paper, we briefly review the relevant literature on energy technology clustering and 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. The result is a geographic clustering of citations.

Empirical evidence stresses the important role of geography in the spillover of knowledge from one member of an innovation network to another (see, for example, a review by Gelsing, 1992). It also emphasizes the importance of frequent personal contact and research collaboration. In particular, Zucker et al. (1998) show that firms using biotechnology are strongly influenced by the location of superstars in academic research institutions.

Lundvall (1992) points out that the importance of geography should differ predictably by technology type. While geography has little impact on stationary technologies (facing constant needs and opportunities), its importance grows quickly for technologies undergoing incremental innovation and radical innovation. During technological revolutions, there is a dramatic effect on the geographic pattern of subsequent innovations. Since certain subsectors within the energy industry have enjoyed aggressive growth, we expect some large geographic impacts on knowledge flows.

Geographic proximity has already been used to explain the location of R&D-intensive activities (for Boston's high-technology district, see Dorfman (1988); for France see Carrincazeaux et al., (2001)), due to evidence of localized spillovers within an industry. However, the location of *firms* is not always a good predictor of the location of *innovation*. Feldman (1994) tests whether U.S. state patent levels can be predicted simply by the presence of particular industries in that state: in fact there is only a 0.42 correlation between innovation measures and value-added in each state by industry. This result is confirmed for the 1975-1994 period in the U.S. (Johnson and Brown, 2004) in an exploration of why the northeastern states lost a dramatic share of the national patenting total. It occurred not only due to the location of industries, but also due to industries that were unable to maintain the patenting rates enjoyed by the same sector in other regions.

Localization of patent citations has been firmly established by the leading paper on the topic (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 small but statistically significant, and are more intense where knowledge becomes obsolescent rapidly, such as electronics, optics, and nuclear technology (Jaffe and Trajtenberg, 1996). The result has also been confirmed for semiconductors (Almeida and Kogut, 1997).

Other researchers have demonstrated a geographic pattern to European patent citations. In a limited sample of Swedish patent applications, international trade flows, rather than physical distance, was the only variable that robustly explained international references (Sjoholm, 1996). In a larger study of over 100,000 patent citations between European regions, there is strong evidence of geographic clustering (Maurseth and Verspagen, 1999). Regressions show that distance between regions is an important driving factor, along with technological similarity between regions.

On the other hand, there are strong voices in the literature who argue that either distance has never mattered as much as was thought (e.g. 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 (e.g. Vasileiadou and Vliegenthart, 2009; Graham, 2001).

3. Data

3.1. Measurement issues

This paper relies exclusively on patent citations from energy-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 of materials that were consulted during the invention process 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. However, within the U.S. on a state-by-state level, patents have a high correlation with other measures of innovative activity. For example, there is a 0.88 correlation between patents and R&D expenditures, 0.99 between patents and research employment records, and 0.93 between patents and a census of innovation citations in scientific and trade journals conducted by the Small Business Administration (Feldman, 1994).

Citations themselves 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) relates survey evidence showing that only a quarter of all patent citations correspond to a clear spillover of knowledge, another quarter have some possibility of a spillover, and the remaining half do not reflect knowledge transfers. However, their statistical tests indicate that overall citations can be interpreted as a signal of spillovers, albeit a noisy signal.

As a final definitional challenge, "energy sector" definitions differ between nations and over time. Therefore, we follow World Intellectual Property Organization's definition for alternative energy, namely the sectors of solar, wind/geothermal and fuel cells. These definitions include portions of 14 separate International Patent Classification (IPC) category codes on the 4-digit level.[1]

Accordingly, we appended our dataset with the set of patents cited by energy 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, as the task of geo-coding all patent applications globally was too Herculean a feat for our team at this time. However, there is evidence in the literature that international citations are increasing in frequency across a host of technologies (e.g. Johnson and Sneed, 2009), evidence which is at least sympathetic to the hypothesis here that citation distances have been increasing over time (Johnson and Lybecker, 2011).

Our focus is thus exclusively upon energy-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.

---

[1] See http://www.wipo.int/patentscope/en/technology_focus/pdf/landscape_alternative_energy.pdf for alternative energy definitions. Specifically, we used the IPC categories given for solar, wind power and fuel cells. Traditional energy was coded as all IPCs starting with C10.

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, citing them more frequently than others, which would give a biased impression of the importance of geography. 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. While inventor and assignee self-citation may drive a pattern of geographic clustering, they confuse the issue of "local knowledge spillovers," which is the primary focus of our analysis. On the other hand, we do not wish simply to ignore self-citations as being obviously local. If an assignee firm is located in several different locations, high familiarity with other inventions by the same assignee may actually work against a geographic clustering of citations. The same may be true of an inventor who moves during his or her career. 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 energy-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 277,850 U.S.-based energy patent documents that include a total of 916,009 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 or not patents *k* and *K* have the same inventor (hereafter, SI),
- whether or not patents *k* and *K* have the same assignee (SA),
- whether or not 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 energy (E),
- 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, and
- year T of citing patent *K*, to account for citation inflation (Y).

We traced all self-citations by inventors, allowing for some flexibility in name spellings (since the USPTO does not standardize name format). These include not only first inventors, but all inventors listed for each patent. We found that self-citation accounted for almost five percent of all citations from energy patents, suggesting that while some self-citation is present, there are very strong inter-inventor knowledge spillovers. This is a much lower self-citation rate than has been documented in other sectors like biotechnology (twelve percent, according to Johnson and Lybecker, 2011), indicating that knowledge transfers between individuals are more common in energy. Unlike academic citations, there is very little reason here to self-cite as a means of advertising, so we can be fairly sure that self-citations are indicators of useful capital or legal protection. Self-citation was coded as a binary variable (SI) for each citation.

On the other hand, using the same energy patents, thirty-two percent of all citations were to the same assignee firm or person, again checking rigorously for variations of firm names. This intra-firm pattern is obviously an important component of the flow of knowledge, but they still imply that 2 out of 3 citations are to other firms. Nonetheless, this rate is higher than has been measured in biotechnology (twenty-four percent, according to Johnson and Lybecker, 2011). Self-citation by assignees coded as a binary variable (SA) for each citation.

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 at the 4-digit level. 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). There are 634 clusters at the 4-digit level, so an indicator that the patents share a class is a powerful signal of technological similarity, and a strong indicator that they were both processed by patent examiners with very similar scientific training. Including this variable serves to ensure that the spatial pattern observed among citations is true, even controlling for spatial patterns in the underlying geography of innovation itself. If, for example, energy technology has gradually clustered in two distant locations, controlling for technology type will permit the analysis to abstract from that pattern in discerning whether geography matters as much as it has in the past.

The fourth variable, SC or the technological correlation between citing and cited states, 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. Again, controlling for technological similarity will defuse the power of the data to show an importance of geography that may be superficially the result of two distant (or proximate) regions sharing the same technological portfolio and hence attracting citation flows. While it would have been interesting to calculate this correlation on a city-to-city basis, it was simply not practical as many locations did not have sufficient activity to populate many technology classes.

The analysis also includes an indicator of whether the cited patent is classified as energy (E). Obviously, all citing patents have been defined as such, and there should be a higher probability for them to cite other energy patents than to cite a random other technology group.

This variable is significantly different than the ‘same IPC’ variable, as the energy technology definition fourteen separate IPCs. Furthermore, the patents were categorized using indicator variables for what sub-category of energy they fell into, whether it be solar (ES), wind/geothermal (EW), fuel cell (EF), traditional (ET) or transportation technology (TR).

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 277,850 citations in the sample dataset are presented in Table 1. Notice first that the average citation is around 1350 kilometers long, but the distribution has a very wide variance and has a high number of zeroes (self-citations) versus some extremely long citation distances (e.g. from Hawaii to Florida). 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. As Figure 1 showed, the average has varied from a low of around 1150 kilometers to a peak of around 1400 kilometers, so it is reasonable to ask whether that change has been due to other correlated factors.

Table 1: Summary statistics for citation dataset

Variable	Mean	Std. Dev.	Min	Max
Citation distance (Km)	1343.27	1284.25	0	12565.20
SC= correlation in technology portfolios between cited and citing states	0.85	0.14	0.01	1
A= age of citation (years)	8.22	5.80	0	27
	Number of zeros			
E_S= citing patent categorized as solar energy	270575 (97%)		0	1
E_T= citing patent categorized as traditional energy	219434 (79%)		0	1
E_F= citing patent categorized as fuel cell technology	272171 (97%)		0	1
E_{TR}= citing patent categorized as transportation technology	74692 (26%)		0	1
E_W= citing patent categorized as wind energy	274528 (99%)		0	1
SI = same inventor in cited and citing patents	265017 (95.3%)		0	1
SA = same assignee in cited and citing patents	188356 (67.8%)		0	1
ST = same IPC in cited and citing patents	129334 (46.5%)		0	1
E = cited patent is also energy	92299 (33.2%)		0	1
G = government assignee	274030 (98.6%)		0	1
U = university assignee	274811 (98.9%)		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 citation as the unit of analysis. While Petersen and Rajan studied the expanding distance between commercial lenders and their customers, the same method is clearly applicable here. The model is a recognition that the distance between a cited patent k granted in year t and a subsequent citing patent, K, granted in year T, can be explained at least in part 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 non-geographic attributes of patents k and K that affect the probability of citation, and ε is a randomly distributed error term. Specifically, we postulate the reduced functional form, using the log of distance (or technically 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 due to the loglinear nature of the data's underlying relationship:

$$\begin{aligned} Distance = {} & \alpha_0 + \alpha_{EnergySolar}ES + \alpha_{EnergyTraditonal}ET + \\ & \alpha_{EnergyFuelcell}EF + \alpha_{EnergyTransport}TR + \alpha_{SameAssignee}SA + \\ & \alpha_{SameInventor}SI + \alpha_{SameTech}ST + \alpha_{Energy}E + \alpha_{Govt}G + \alpha_{University}U + \\ & \alpha_{Age}A + \alpha_{AgeSquared}A^2 + \textstyle\sum_{1976}^{2002} \alpha_{Year=i}Y_i + \varepsilon_K + u \end{aligned} \qquad \textbf{(2)}$$

where the distance of each observed citation is explained by the attributes of the citing and cited patents. Notice that we use a fixed effect specific to the citing patent (ε_K), since there are presumably immeasurable factors specific to the citing patent 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 miles), with White-corrected errors to accommodate the presence of heterskedasticity in the sample, using fixed effects at the level of the citing patent where each individual citations is the unit of analysis. For simplicity, we estimate using only a time trend (and nonlinear versions of it) as an explanatory variable. Considering all citations, the average distance unambiguously increases with time, with strong evidence of a non-linear pattern. When considering only inter-city citations with distances greater than 100 kilometers, the evidence is still very strong that a nonlinear pattern exists, one with distance rising with time.

Table 2: Tobit weighted regressions on log(distance+1), time trend only

	All citations				Only citations with distance > 100 km			
	Time trend		Nonlinear time trend		Time trend		Nonlinear time trend	
	coefficient	t-statistic	coefficient	t-stat	coefficient	t-stat	coefficient	t-stat
Trend	2.719×10^{-2}	(35.87)***	5.168×10^{-2}	(13.04)***	-4.95×10^{-4}	(-1.54)	6.28×10^{-3}	(3.77)***
Trend^2	---	---	-7.22×10^{-4}	(6.52) ***	---	---	-1.984×10^{-4}	(4.16) ***
Constant	5.555	(349.41)***	5.379	(159.94) ***	7.110	(1083.8) ***	7.059	(517.70) ***
F-stat Obs		1286.44*** 277850		643.35*** 277850		2.36 218768		9.98*** 218768

Notes: *** indicates 99% confidence, ** 95% confidence, * 90% confidence.

To permit maximum flexibility to these nonlinearities, and potential nuances in particular years, Table 3 offers the same analysis, using separate year indicator variables. Notice that while increasing, the annual indicator variables do not uniformly increase over time (e.g. 1982-83, 1986-87).

Table 4 presents the primary results, confirmed by the ancillary results in Table 5 which use time-based indicator variables instead of a time trend. Both sets of results differ only minimally from the results of a model which uses the citing patent as the unit of analysis (not presented here), where each citation is weighted appropriately according to the number of citations referenced by the citing patent in question.

Table 3: Tobit weighted regressions on log(distance+1), separate year time dummies

	All citations		**Only citations with distance> 100 km**	
Variable	**coefficient**	**t-statistic**	**coefficient**	**t-statistic**
citingyear77	0.085	(0.54)	7.20×10^{-3}	(0.11)
citingyear78	2.00×10^{-4}	(0.00)	0.069	(1.11)
citingyear79	0.044	(0.30)	0.097	(1.56)
citingyear80	0.255	(1.77)*	0.059	(0.97)
citingyear81	0.300	(2.11)**	0.124	(2.08)**
citingyear82	0.357	(2.53)**	0.129	(2.19)**
citingyear83	0.313	(2.23)**	0.121	(2.06)**
citingyear84	0.393	(2.82)***	0.114	(1.94)*
citingyear85	0.490	(3.53)***	0.110	(1.89)*
citingyear86	0.589	(4.26)***	0.183	(3.14)***
citingyear87	0.582	(4.23)***	0.134	(2.31)**
citingyear88	0.549	(3.99)***	0.166	(2.87)***
citingyear89	0.622	(4.56)***	0.135	(2.35)**
citingyear90	0.512	(3.75)***	0.108	(1.87)*
citingyear91	0.612	(4.49)***	0.085	(1.48)
citingyear92	0.643	(4.71)***	0.119	(2.07)**
citingyear93	0.626	(4.60)***	0.114	(1.98)**
citingyear94	0.704	(5.18)***	0.102	(1.77)*
citingyear95	0.719	(5.29)***	0.117	(2.04)**
citingyear96	0.772	(5.69)***	0.094	(1.64)
citingyear97	0.780	(5.75)***	0.127	(2.22)**
citingyear98	0.801	(5.92)***	0.133	(2.33)**
citingyear99	0.842	(6.22)***	0.122	(2.13)**
citingyear00	0.742	(5.44)***	0.109	(1.90)*
citingyear01	0.876	(6.45)***	0.109	(1.89)*
citingyear02	0.865	(6.40)***	0.094	(1.65)*
Constant	5.388	(40.03)***	6.98	123.05
F-statistic		52.74***		4.85***
Observations		277850		218768

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

Table 4: Tobit weighted regressions on log(distance+1), time trend

	All citations		distance>100km	
Variable	**coefficient**	**t-statistic**	**coefficient**	**t-statistic**
solar energy (E_S)	-0.123	(2.07)**	0.103	(3.67)***
traditional energy (E_T)	-0.503	(9.88)***	-0.111	(4.89)***
fuel cell energy (E_F)	-0.618	(9.61)***	-0.308	(9.32)***
transportation energy (E_{TR})	-0.315	(6.33)***	-0.193	(8.78)***
same assignee (SA)	-0.692	(55.35)***	0.128	(19.90)***
same inventor (SI)	-3.323	(83.22)***	-0.527	(13.92)***
same ipc (ST)	-0.249	(17.13)***	-0.048	(6.39)***
citing-cited state correlation (SC)	-6.657	(145.03)***	-1.429	(65.11)***
cited energy (E)	0.140	(9.15)***	0.059	(7.30)***
assignee govt (G)	0.066	(1.62)	0.130	(5.83)***
assignee university (U)	0.289	(6.26)***	0.189	(7.63)***
citation age (A)	6.44×10^{-2}	(19.97)***	6.40×10^{-3}	(3.76)***
square age of citation (A^2)	-1.80×10^{-3}	(13.41)***	-2.217×10^{-4}	(3.06)***
trend	5.52×10^{-3}	(1.29)	9.944×10^{-3}	(4.44)***
trend^2	-2.141×10^{-4}	(1.70)*	-2.836×10^{-4}	(4.32)***
constant	12.119	(184.21)	8.272	(258.51)***
F-statistic Observations		2943.3*** 277850		375.81*** 218768

Regressions using a time trend display statistically significant and positive impacts of the passage of time on average distance. Using annual indicators in Table 5, each annual coefficient differs from the comparison year (1976) at a statistically insignificant level, but all time periods show a coefficient value higher than that of the implicit comparison year. The test statistic for the null hypothesis that annual coefficients do not change over time is F(24, 277811) = 1.05, which is insignificant at any level. We suspect that each year is insignificant due to the random nature of what materials happen to be patented in any given year, but we are open to considering other possible explanations.

Table 5: Tobit weighted regressions on log(distance+1), separate year time dummies

	All citations		distance>100 km	
Variable	**coefficient**	**t-statistic**	**coefficient**	**t-statistic**
solar energy (E_S)	-0.126	(2.12)**	0.104	(3.69)***
traditional energy (E_T)	-0.504	(9.91)***	-0.110	(4.85)***
fuel cell energy (E_F)	-0.622	(9.65)***	-0.308	(9.29)***
transportation energy (E_{TR})	-0.320	(6.41)***	-0.192	(8.70)***
same assignee (SA)	-0.693	(55.38)***	0.128	(19.92)***
same inventor (SI)	-3.324	(83.34)***	-0.528	(13.96)***
same ipc (ST)	-0.249	(17.12)***	-0.048	(6.39)***
citing-cited state correlation (SC)	-6.657	(145.1)***	-1.429	(65.12)***
cited energy (E)	0.141	(9.19)***	0.060	(7.28)***
assignee govt (G)	0.064	(1.58)	0.130	(5.83)***
assignee university (U)	0.290	(6.30)***	0.189	(7.62)***
citingyear77	-0.293	(2.41)**	-0.060	(0.92)
citingyear78	-0.307	(2.65)***	-9.010×10^{-3}	(0.14)
citingyear79	-0.321	(2.72)***	1.191×10^{-3}	(0.02)
citingyear80	-0.277	(2.44)**	-9.456×10^{-3}	(0.15)
citingyear81	-0.221	(1.98)**	4.509×10^{-3}	(0.75)
citingyear82	-0.247	(2.21)**	3.792×10^{-3}	(0.64)
citingyear83	-0.328	(2.96)***	1.771×10^{-3}	(0.30)
citingyear84	-0.301	(2.73)***	5.381×10^{-3}	(0.91)
citingyear85	-0.222	(2.03)**	2.534×10^{-3}	(0.43)
citingyear86	-0.261	(2.38)**	5.582×10^{-3}	(0.95)
citingyear87	-0.253	(2.33)**	4.461×10^{-3}	(0.76)
citingyear88	-0.260	(2.40)**	6.654×10^{-3}	(1.14)
citingyear89	-0.222	(2.07)**	5.655×10^{-3}	(0.98)
citingyear90	-0.257	(2.39)**	3.869×10^{-3}	(0.67)
citingyear91	-0.276	(2.58)***	2.343×10^{-3}	(0.41)
citingyear92	-0.236	(2.21)**	3.198×10^{-3}	(0.55)
citingyear93	-0.228	(2.14)**	5.883×10^{-3}	(1.02)
citingyear94	-0.240	(2.26)**	6.148×10^{-3}	(1.07)
citingyear95	-0.284	(2.67)***	4.257×10^{-3}	(0.74)
citingyear96	-0.261	(2.46)**	4.017×10^{-3}	(0.70)
citingyear97	-0.291	(2.74)***	3.626×10^{-3}	(0.63)
citingyear98	-0.281	(2.66)***	7.397×10^{-3}	(1.30)
citingyear99	-0.272	(2.58)***	5.548×10^{-3}	(0.97)
citingyear00	-0.295	(2.76)***	3.968×10^{-3}	(0.69)
citingyear01	-0.268	(2.52)**	3.849×10^{-3}	(0.67)
citingyear02	-0.303	(2.88)***	1.422×10^{-3}	(0.25)

citation age (A)	6.492×10^{-2}	$(20.04)^{***}$	6.130×10^{-3}	$(3.59)^{***}$
square age of citation (A^2)	-1.825×10^{-3}	$(13.53)^{***}$	-2.097×10^{-4}	$(2.89)^{***}$
Constant	12.408	$(103.86)^{***}$	8.305	$(132.76)^{***}$
F-stat Observations		1136.89^{***} 277850		145.83^{***} 218768

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

Sensitivity tests find very similar results if we restrict our consideration to citations of more than 10 kilometers, of more than 50 kilometers, or of more than 100 kilometers. Results for citations spanning more than 100 kilometers, that is, excluding short and intra-city citations, are presented in Table 4 and Table 5, and tell a very similar story. Alternatively, results omitting citations from the states with the most citations (CA, NJ and MI) again show the same pattern, with all coefficients of a size and sign similar to those presented here.

Moving to other elements of the regression results, notice that fuel cell and traditional energy patent citations are likely to transit shorter distances than other forms of energy knowledge. In contrast, solar energy appears to be the most mobile or easy to transmit over long distances.

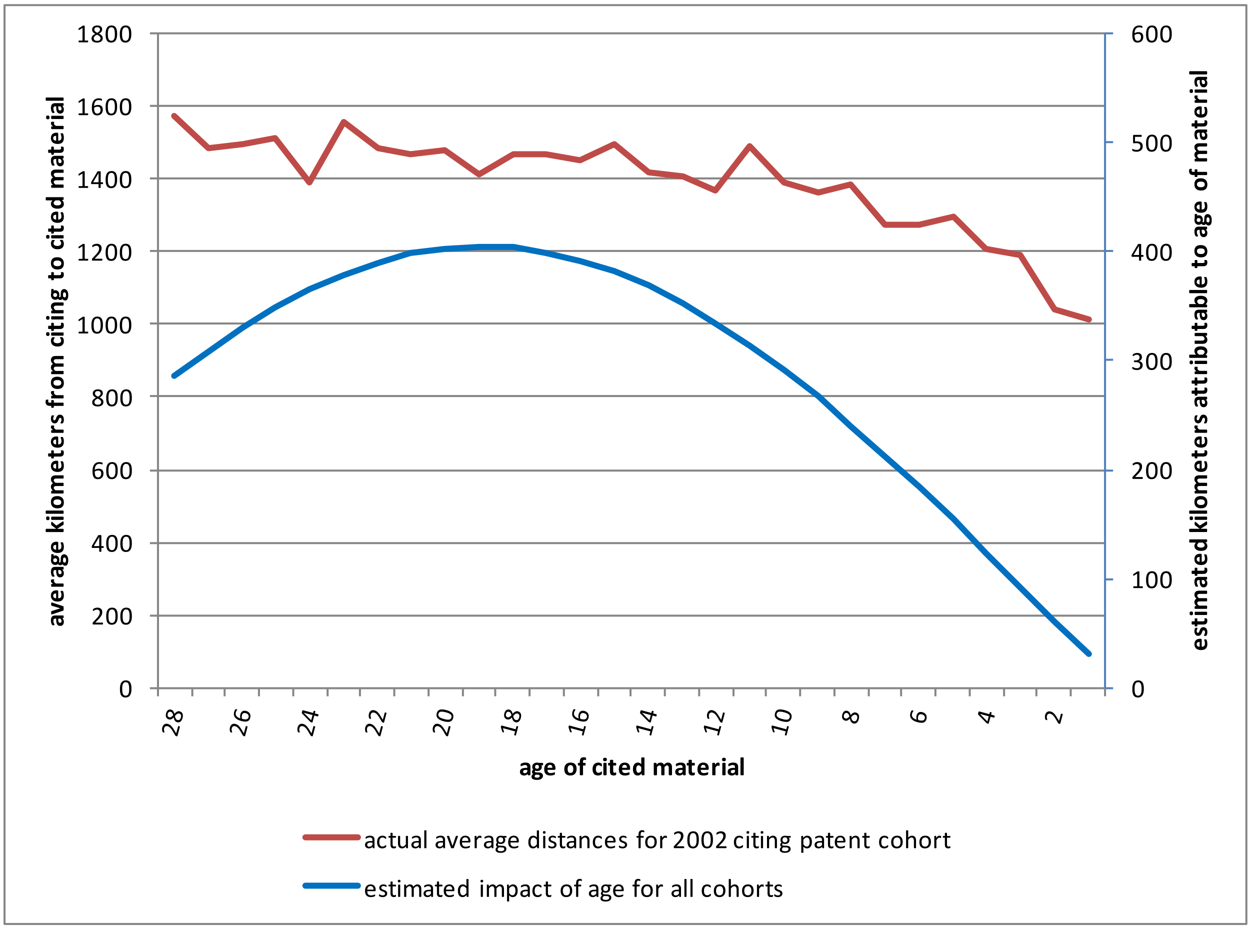
Unsurprisingly, citations with the same assignee or same inventor are more likely to be proximate than are other citations. The effect is especially strong and significant for inventors, suggesting that at least within energy, inventors are not likely to move locations between self-citations. Citations within the same specific technology class appear to reference citing and cited patents that are closer to each other than more dissimilar patents (the ST coefficient is negative), and states that have similar technology sets in their innovative portfolios tend to be close together, a fact captured by the negative coefficient on that variable (SC).

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

Citations from government-assigned patents tend to travel longer transmission distances for the knowledge they cite, a result that is much more pronounced when we consider only long-distance (>100 km) citations. Academic patents tend to be longer than their peers from the business sector as well, but that effect is less pronounced among long-distance citations (>100 km).

The age of the cited patent matters as well: older citations travel longer distances, an effect which other studies (e.g. Johnson and Popp, 2002) have confirmed for an array of technologies. Figure 2 shows the average citation distances of an actual cohort of patents, those cited by patents granted in 2002. Notice that the average distance appears to lengthen as the age of the cited material rises (presumably because it takes time for knowledge to diffuse longer distances). Presented on the same graph is the estimated impact of the age of cited material on citation distance, as estimated from the nonlinear model coefficients. When considering all cohorts, it appears that material eighteen to twenty years old diffuses the farthest, while concurrent knowledge clearly diffuses the shortest distance.

Figure 2: Change in distance according to age of cited material



While all citation distances have statistically lengthened over time, that effect is not economically significant for all consecutive years (due, presumably, to the idiosyncratic nature of the specific technologies being patented in any given year). Citation length increases by an average of 0.7 percent, or 8.33 kilometers per year over the period in question, but that average hides great variation over time. On average, we can conclude that distances increased roughly 0.7 percent per year over the period, a much slower rate than the documented rise in distances for biotechnology of 2.4 percent per year (Johnson and Lybecker, 2011). While the rise has not been monotonic, it has been steady, with no clear oddities or eccentric patterns to explain.

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 in the era of computerization, Internet, teleconferencing, and cellular communication. After all, this is in line with the literature (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 energy technology knowledge. Controlling for other factors, knowledge flows used to diminish 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.

Alternatively, perhaps it is the nature of the energy sector itself that changed over this period in time. In the late 1970s, radically different energy technologies were patented, with nascent solar energy technology competing with traditional technologies but few wind or fuel cell patents. By the end of our analysis period, solar/wind/fuel cell technologies comprised 8.09 percent of all patents in the energy sector, up from 3.71 percent at the beginning of the sample period. The majority of patents fall into the energy transportation sector, as this sector accounted for over 80 percent of the total energy patent citations throughout the analysis period.

Long-distance knowledge transfers are increasingly the norm in energy technology. As the importance of physical distance has diminished over time, 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 energy 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.

Federal Communications Commission (FCC) (2001), "Statistics of the Long Distance Telecommunications Industry". Industry Analysis Division, Common Carrier Bureau, Federal Communications, Washington, DC.

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

Feldman, M.P. (1999) "The New Economics of Innovation, Spillovers and Agglomeration: A Review of Empirical Studies". Economics of Innovation and New Technology, Vol 8, p.5-25.

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.

Greene, W.H. (1993) Econometric Analysis, Macmillan Publishing Company, New York.

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 R.E.Evenson (2000) "How Far Away Is Africa? Technological Spillovers to Agriculture and Productivity". American Journal of Agricultural Economics. Vol 82, p.743-749.

Johnson, D.K.N. and K.M. Lybecker (2011) "Does distance matter less now? The changing role of geography in biotechnological innovation". Forthcoming *Review of Industrial Organization,* Fall 2011.

Johnson, D.K.N. and V.Santaniello (2000) "Biotechnology Inventions: What Can We Learn from Patents?". Agriculture and Intellectual Property Rights: Economic, Institutional, and Implementation Issues in Biotechnology. V. Santaniello et al., editors. CABI Publishing, New York.

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.

Knoke, D. and J.H.Kuklinski (1982) Network Analysis. Sage Publications, London.

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.

Maurseth, P.B. and B.Verspagen (1999) "Europe: one or several systems of innovation? An analysis based on patent citations". Economic Challenge to Europe: Adapting to Innovation-Based Growth. Edward Elgar Publishing, 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

Sjoholm, F. (1996) "International Transfer of Knowledge: The Role of International Trade and Geographic Proximity". Weltwirtschaftliches Archiv, Vol 132, p. 97-115.

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.

USPTO Technology Assessment and Forecast Program (1998) "Technology Profile Report: Patent Examining Technology Center Groups 1630-1650, Biotechnology". United States Patent and Trademark Office report.

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

Valente, T.W. (1995) Network Models of the Diffusion of Innovations. Hampton Press, Inc. Cresskill, New Jersey.

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.

Winkler, A.E., S.G. Levin and P.E. Stephan (2009) "The Diffusion of IT in Higher Education: Publishing Productivity of Academic Life Scientists". Andrew Young School of Policy Studies Research Paper Series No. 09-01.