

Patent Citations and the Geography of Knowledge Spillovers: Evidence from Inventor- and Examiner-Added Citations

Peter Thompson

Florida International University

March 2004

(This revision: June 2005)

I report new evidence for localized knowledge spillovers identified by within-patent variations in the geographic matching rates of citations added by inventors and citations added by examiners. Evaluated at the mean citation lag, inventor citations are 20 percent more likely than examiner citations to match the country of origin of their citing patent, while US inventor citations are 25 percent more likely to match the state or metropolitan area of their citing patent. The localization of intranational knowledge spillovers declines with the passage of time, but international borders present a persistent barrier to spillovers. (*JEL* O310, O340)

Department of Economics, Florida International University, Miami, FL 33199. This research was supported by the National Science Foundation under Grant No. SBR-0296192. This paper has benefited from especially helpful comments by two referees and the editor, Daron Acemoglu.

I. Introduction

Well over a century has passed since Marshall's (1890) observation that agglomeration of specialized industries is due in part to the localization of knowledge spillovers. The notion continues to resonate among economists [e.g. Jacobs (1969), Feldman (1994a, 1994b), Glaeser et al. (1992), Manski (2000)], and is implicit in most theories of economic growth [e.g. Romer (1990), Grossman and Helpman ((1991)]. Nonetheless, about the only *direct* evidence we have for localized knowledge spillovers is based on Jaffe, Trajtenberg and Henderson's (1993 – hereafter JTH) pioneering analysis of patent citations using a case-control matching method.¹

Thompson and Fox Kean (2005) recently argued that at least part of JTH's evidence for strong localization effects is driven by imperfect matching that generates the appearance of localization effects even when none exist. Although they refine JTH's matching method, they also contend that one cannot really trust evidence about localization effects obtained after selecting control patents by technology classification. This conclusion is particularly disheartening because patent citations remain the only counterexample to Krugman's (1991:53) observation that “knowledge flows . . . leave no paper trail by which they may be measured and tracked.”

This note reports results from an alternative identification scheme that continues to follow the paper trail left by patent citations. Since January 2001, the USPTO has indicated whether each citation in a patent was added by the inventor or by the examiner. The analysis exploits this new information to examine within-patent and within-examiner variations in the citing-cited geographic matching rates of citations added by inventors and those added by examiners.

Using a sample of over 27,000 citing-cited patent pairs, the estimations produce consistent evidence of localization effects at all geographic levels. Inventor citations are 20 percent more likely to match the country of origin of the citing patent than are examiner citations, while for domestic patents they are 25 percent more likely to match the citing patent's location within the United States. The localization of in-

¹ JTH's methodology has been applied in more specialized settings [e.g. Agrawal, Cockburn, and McHale (2003), Almeida (1996), Frost (2001), Hicks et al. (2001)], with similar results.

transnational knowledge spillovers declines with the passage of time, but international borders present a persistent barrier to spillovers.

Section II of this note describes the data. Section III discusses the identification strategy. Section IV provides the results, and Section V briefly concludes.

II. Data

The citing patents in the sample consist of all US patents granted during the first week of January 2003 and for which there is an institutional assignee. The analysis of citations is restricted to patents granted after January 1, 1976, the contents of which are available in machine-readable form. The numbers of every such patent referenced by at least one of the citing patents were collected. Programs in *perl* extracted the following details for each citing and cited patent: assignee name and location,² inventor names and locations, date of issue, date of application, US classification codes, international classification codes, examiner's field of search codes, and name of primary examiner. Additional coding noted whether a cited patent had an institutional assignee, and whether any citing-cited patent pair was a self-citation (i.e. the two patents had an assignee in common). Finally, the patent image files were then checked manually to detect which patents were added by the examiner, and to make numerous corrections.

The location assigned to each patent was determined by the residence of its first-named inventor.³ Patents report the town and state or country of residence of each inventor. Towns in the US were matched to counties, and, where relevant, to one of some 300 metropolitan statistical areas and to one of 17 consolidated metropolitan statistical areas (CMSAs) as defined by the US Census Bureau in 1990. This

² In the relatively rare cases with multiple institutional assignees, the name and location of the first-named assignee was used. Patents with private individuals and an institution as co-assignees were assigned to the institution.

³ Using only the first-named inventor has been a common approach because of programming ease. Thompson and Fox Kean (2004) discuss the merits of this approach compared with alternative assignment rules.

matching was initially done using correlation files provided by the Office of Social and Economic Data Analysis (OSEDA) of the University of Missouri. Some 1,000 place names not in the OSEDA files, usually neighborhoods in major metropolitan areas, were identified using digital maps available from <http://maps.yahoo.com>.

To check that the technology coverage of the sample is not unusual, the three-digit primary classes of the citing patents in the sample were compared with the universe of patents granted between 1997 and 1999, taken from the NBER patent data file. The distribution of the sample across technology classes is highly correlated with the corresponding distribution in the NBER data file, but there are five outliers.⁴ No doubt, part of this difference is attributable to changes in the patent population during the five years or so that separate the samples. Nonetheless, the results to be presented in Section IV were checked after eliminating these five technology classes. Doing so did not change the results reported.

Table 1 provides some basic information about the sample. After the elimination of seven cited patents⁵, the sample contains 31,377 citations generated by 2,670 citing patents, an average of just under twelve citations made by each patent. Examiners are a very important source of citations, accounting for over 41 percent of the sample. Moreover, they accounted for *all* the citations made by 38 percent of the citing patents, compared with only 8.5 percent for inventors. Surprisingly, the self-citation rate is only a little higher among inventor citations (12.5 percent against 10.9 percent), which suggests at the least a certain casualness with which inventors prepare their applications.

⁴ Class 257 (transistors; 3.26% of the sample, 1.61% of the NBER data file), 438 (semiconductors; 3.33% sample, 1.79% NBER) and 365 (information storage; 1.87% sample, 1.07% NBER) are overrepresented in the sample, while classes 435 (molecular biology, 1.8% sample, 2.65% NBER) and 514 (drug compounds, 2.29% sample, 2.82% NBER) are underrepresented.

⁵ One was eliminated because the inventor location could not be traced, six because the applications were filed prior to 1970.

TABLE 1. *Summary Statistics*

CITING PATENTS			
NUMBER OF OBSERVATIONS			2,670
CITATIONS PER PATENT			11.8
FRACTION OF CITING PATENTS WITH ALL CITATIONS ADDED BY EXAMINER			.380
FRACTION OF CITING PATENTS WITH ALL CITATIONS ADDED BY INVENTOR			.085
CITED PATENTS, ISSUED AFTER JAN 1, 1976.			
	ALL CITATIONS	ADDED BY INVENTOR	ADDED BY EXAMINER
NUMBER OF OBSERVATIONS	31,377	18,413	12,964
MEAN FILING DATE	May, 1992	May, 1991	May, 1993
MEDIAN FILING DATE	Apr, 1994	Feb, 1993	Nov, 1995
MEAN ISSUE DATE	Jun, 1994	May, 1993	Nov, 1995
MEDIAN ISSUE DATE	Mar, 1996	Dec, 1994	Mar, 1998
FRACTION SELF-CITATIONS	.118	.125	.109
FRACTION INSTITUTIONAL ASSIGNEE	.906	.904	.910

III. Identification Strategy

The identification strategy rests on two assumptions. The first is that examiners, who work in a single campus located in Alexandria, VA and most commonly enter the USPTO directly from college, cannot be learning about prior art because of geographic proximity to related technological activities.⁶ The second is that an inventor citation is more likely to represent a true knowledge flow than is an examiner cita-

⁶ The patent examination procedure is explained in detail in Cockburn, Kortum and Stern (2003), the contents of which need not be replicated here.

tion. The clearest support for the second assumption comes from the NBER/Case Western Reserve survey of patentees [Jaffe, Trajtenberg, and Fogarty (2002)], from which it was found that examiner citations are more likely to reflect ignorance on the part of the inventor.

It is not necessary that all knowledge flows be captured by inventor citations. The NBER/Case Western Reserve survey documents several sources of noise in the citations data. Inventors may cite prior art by conducting a search (or having their lawyers do so) after completing the invention, thereby adding citations that do not reflect a knowledge flow.⁷ Inventors may also fail to cite prior art that they do know about, and these are eventually added by the examiner. These sources of noise reduce the power of tests of geographic differences between inventor and examiner citations, and lead to a systematic underestimate of the magnitude of differences in geographic matching rates. However, the tests in this paper, as in previous work, turn on the statistical significance of any differences found. Power lost by noise can always be recovered by increasing sample size.

But this identification scheme will fail, and produce spurious evidence that knowledge flows are constrained by geography, if those examiner citations that do represent knowledge flows are less likely to produce geographic matches than inventor citations. One mechanism generating this bias is as follows⁸ Suppose that all citations capture a knowledge flow with equal probability. Knowledge flows by word of mouth, so that knowledge from more distant sources has passed through more agents before reaching an inventor. Because of the longer chain, the inventor does not learn of the source of ideas with more distant sources. If the inventor also does not undertake a patent search of these ideas, he or she is more likely to cite local patents, while the examiner fills in the gap.

This alternative story appears observationally equivalent to the identification scheme, so it is difficult to discount directly. However, if the story is correct, one would expect that examiners, who are citing sources that have passed through a longer chain, will on average cite older patents. The data show the opposite. Figure 1

⁷ See Breschi and Lissoni (2004) for a review of the evidence.

⁸ This insight is due to Daron Acemoglu.

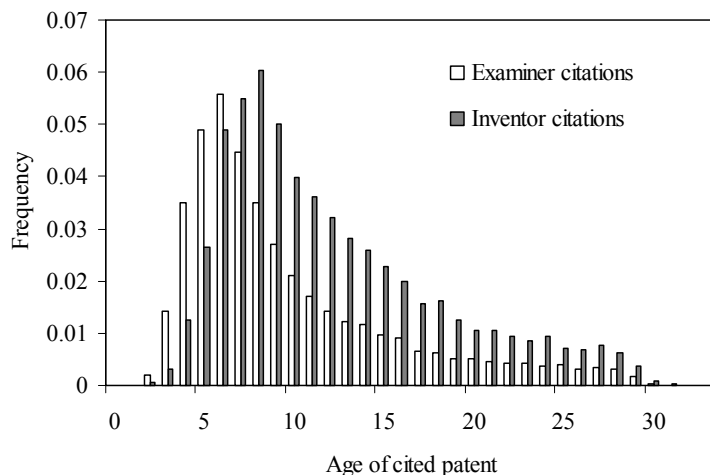


FIGURE 1. Age distribution of cited patents, by citation source.

plots the distribution of cited patent ages by citation source. Not only is the mean age of examiner citations lower, it is readily verified that the age of examiner citations first-order stochastically dominates the age of inventor citations. Even if we remove from the sample cited patents younger than five years old (because examiners are particularly likely to cite them), stochastic dominance holds.

IV. Results

All results in this section were obtained after eliminating self-citations from the sample. International localization effects are assessed by comparing the country of residence of the first-named inventor in the citing and cited patents. The variable `MATCH COUNTRY` was set to one if both inventors resided in the same country, and zero otherwise. Intranational localization effects are reported at three levels. First, if the citing patent has a US origin, does the cited patent have the same state origin as the originating patent (`MATCH STATE`)? Second, if the origin of the citing patent is in one of the 300 or so MSAs, does the cited patent share the same MSA (`MATCH MSA`)? Third, if the origin of the citing patent is in one of the 17 CMSAs, does the cited patent share the same CMSA (`MATCH CMSA`)? The first column of Table 2 indicates the number of observations available at each level of analysis. After elimi-

nating self-citations, 27,665 observations are in the sample. Of these, 18,737 citing patents have a US origin, 17,826 have an origin in an MSA, and 9,721 have an origin in a CMSA.

TABLE 2. *Crude Geographic Matching Rates*

	N	ALL OBSERVATIONS	CITATIONS ADDED BY . .	
			INVENTOR	EXAMINER
MATCH COUNTRY	27,665	.539	.584	.476
MATCH STATE ^a	18,737	.111	.115	.101
MATCH CMSA ^b	9,721	.104	.111	.090
MATCH MSA ^c	17,826	.093	.098	.082

Self-citations excluded. ^a Conditional on the citing patent having a US inventor. ^b Conditional on the citing patent coming from a CMSA. ^c Conditional on the citing patent coming from an MSA.

The remaining columns of Table 2 summarize the crude geographic matching rates. The table presents a uniform picture at all geographic levels, with matching rates for inventor citations exceeding the rates for examiner citations by between 7 and 23 percent. Of course, the localization effects suggested by these crude matching rates may be confounded with important composition effects, whereby examiner citations just happen to figure more prominently in patents that exhibit lower geographic matching rates. To eliminate such potential confounding, I turn now to conditional logit estimation.

Panel A of Table 3 reports odds ratio estimates from logits with fixed effects for each cited patent. Two general comments are in order. First, because the number of observations per citing patent is modest, the bias induced by the well-known incidental parameter problem may be quantitatively important.⁹ Estimation is therefore car-

⁹ Katz (2001) has simulated fixed-effect unconditional logit models for finite samples and found that the bias is modest when the number of observations per group, T , reaches about 10, and is essentially zero when T exceeds 20. In the present sample, although T averages 11.8, it varies from 1 to over 150. Little is yet known about the extent of bias in the uncondi-

ried out using Chamberlain's (1980) conditional logit model. Doing so means, first, that we cannot obtain consistent estimates of the patent fixed effects, numbers that might be of considerable interest in their own right. Second, when the outcome is identical for all observations within a group (either no citations in a single patent have a match, or they all do), the fixed effect alone is a sufficient statistic for the estimated matching rate. Such observations do not contribute useful information in estimating the parameters of interest. Table 3 consequently reports sample sizes reflecting the number of observations that make a positive contribution to the likelihood.¹⁰

Each model contains three regressors. First, a dichotomous variable is set equal to one if the citation was added by the inventor, zero otherwise. An odds ratio in excess of one provides *prima facie* evidence of localization effects. Second, a dichotomous variable is set equal to one if the cited patent has no institutional assignee. Because the geographic distribution of non-institutional patents differs markedly from institutional patents (they are more likely to be American, and more likely to be from outside metropolitan areas), it was expected that the type of assignee would influence the matching rate. Third, many theories of technological diffusion

tional logit estimator when T varies by group to this extent. My colleague, Jonathan Hill, kindly ran some simulations on samples with the degree of variation in group size exhibited by the sample, and found that bias appears to be a significant problem in the unconditional logit.

¹⁰ A significant number of observations are lost, especially in the finer geographic classes, because of the lack of within-group variation in the dependent variables. One way to reduce the loss of effective observations is to estimate a model with fewer fixed effects, which reduces the number of groups with no variation in the dependent variable. One candidate for an alternative set of fixed effects is the primary examiner on the citing patents, who have distinct specialties [Cockburn, Kortum and McHale (2003)] and therefore may serve as an adequate, albeit somewhat cruder, control. With 975 primary examiners in the sample, the mean number of observations per examiner is about three times the mean number of observations per citing patent. The results are not modified by this alternative set of controls, and so are not reported here

suggest that matching rates may be lower the older the cited patent, so a linear trend for the age of the cited patent is included.¹¹

TABLE 3. *Odds Ratios for Geographic Matching Rates*

DEPENDENT VARIABLE	<i>N</i> ^a	INVENTOR CITATION	NON-INSTITUTIONAL	CITED PATENT AGE	AGE X INVENTOR CITATION
PANEL A					
MATCH COUNTRY	22,198	1.207 (3.76)	1.28 (4.32)	0.990 (-3.33)	—
MATCH STATE	11,864	1.31 (3.17)	1.072 (0.78)	0.966 (-6.51)	—
MATCH CMSA	6,036	1.292 (2.10)	0.832 (-1.37)	0.966 (-4.35)	—
MATCH MSA	10,470	1.300 (2.76)	0.805 (-2.07)	0.954 (-7.64)	—
PANEL B					
MATCH COUNTRY	22,198	1.151 (1.73)	1.281 (4.33)	0.988 (-2.66)	1.00 (0.76)
MATCH STATE	11,864	1.782 (4.07)	1.072 (0.78)	0.988 (-1.25)	0.969 (-2.74)
MATCH CMSA	6,036	1.666 (2.50)	0.831 (-1.38)	0.987 (-0.87)	0.972 (-1.57)
MATCH MSA	10,470	1.732 (3.40)	0.804 (-2.09)	0.976 (-2.09)	0.970 (-2.22)

Z-scores in parentheses. ^a Number of observations contributing to the likelihood function.

The estimated odds ratios for INVENTOR CITATION are in general a little higher than the crude geographic matching rates reported in Table 2. Inventor citations are 20 percent more likely to show a country match than are examiner citations, which is similar to the crude rates; but they are 30 percent more likely to show a state, CMSA or MSA match, which are about 1.5 times the corresponding crude matching rates.

¹¹ Age is measured in years since the filing date of the citing patent.

The two additional controls behave in just the way anticipated. First, the matching rate for non-institutional patents is higher at the country level, because the majority of the sample consists of domestic patents and most non-institutional patents are domestic. In contrast, the matching rate is lower at the MSA and CMSA levels, because non-institutional patents are less likely to be found in metropolitan areas. The intermediate level of the state shows no difference. Second, there is consistent evidence at all geographic levels that matching rates decline with the age of the cited patent. Quantitatively, the effect is quite marked. The odds ratio falls by between 1 and 5 percent per year, so that, say, a ten year old cited patent is between 10 and 50 percent less likely to generate a match than the most recent cited patent.¹² However, the estimated rate of decline is markedly lower for the country matching rate.

The decline in matching rates for older cited patents does not constitute direct evidence that knowledge spillovers become less localized with the passage of time. In principle, it is possible to distinguish between two confounding effects – the diffusion of industrial activity over wider geographic areas and changes in the localization of spillovers – by adding to the regressions an interaction term between the indicator variable for inventor citations and the cited patent age. Panel B of Table 3 reports the results. The odds ratios for inventor citations, which have risen markedly at the intranational levels, now measure the effect for the youngest cited patents. The results continue to show that increases in the age of the cited patent reduce matching rates, reflecting diffusion of technological activity both intranationally and nationally.

The odds ratios for the interaction term tell an interesting story. It is essentially unity at the international level, while all estimates of the odds ratio at intranational levels are less than one. Put another way, knowledge spillovers appear to become less localized over time within the US, but not between countries. However, at the sample mean age, 11.4 years, all four levels of analysis show very similar degrees of localization; the estimated odds ratios at this age are 1.21 (country), 1.25 (state), 1.23 (CMSA), and 1.25 (MSA), each of which is significantly greater than unity at the customary five percent level. To aid interpretation of these results, Figure 2 plots the odds ratio for inventor citations by age of the cited patent for country and MSA

¹² Replacing the linear trend with age dummies does not alter these results.

matching.¹³ There is a modest, but persistent, localization of knowledge spillovers at the international level. At the MSA level, there are strong localization effects for the most recent cited patents, but this effect decays with time so that, by age 14, the effect is no longer statistically significant.

This decay of localization effects within the US, but not across countries is just what one would expect if the decay of localization is caused primarily by inventor relocation, which is much more likely to occur between regions of the US than between countries. While 5.6 percent of the 2001 US population had changed their county of residence during the prior year and 2.8 percent had changed state, only 0.6 percent had arrived from abroad [US Census (2004)].

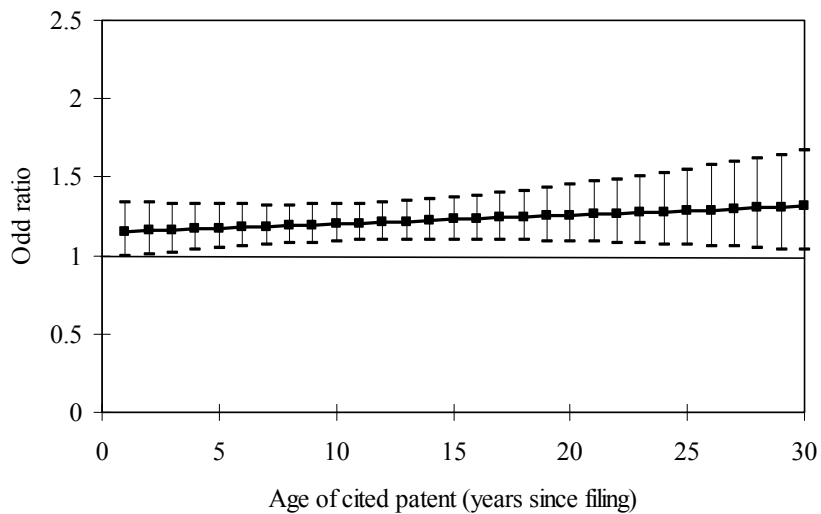
Table 4 reports some further analyses. In Panel A, a distinction is made between patent pairs for which the citation crosses technology classes, and those pairs for which both patents share the same primary three-digit class. One might expect the more homogeneous a network, the more readily it transcends geography.¹⁴ One would therefore expect geography to matter more when technology classes differ. This is exactly what the data show. Intranationally, geographic matches are more likely when patent pairs have the same primary technology class (surprisingly the same does not appear to be true at the country level). After controlling for technology class level effects, localization effects turn out to be markedly stronger at every level when patent pairs do not share the same technology class.

Almeida and Kogut (1999), among others, have shown that localization effects are stronger in certain high-technology regions, such as Silicon Valley, the Route 126 corridor and Austin, TX, than in other regions. The evidence is consistent with the widespread perception that in these regions ideas are stimulated by local technologi-

¹³ The figures for CMSA and state matching are very similar to the figure for MSA matching.

¹⁴ There is a simple analogy in our own profession. The economists we know in our own specialty are likely to be more dispersed than economists we know in other specialties, and they in turn are likely to be more dispersed than the non-economist academics that we know. In a similar vein, Agrawal, Kapur and McHale (2004) show that co-ethnicity substitutes for co-location.

PANEL A. Odds Ratio for Country Match



PANEL B. Odds Ratio for MSA Match

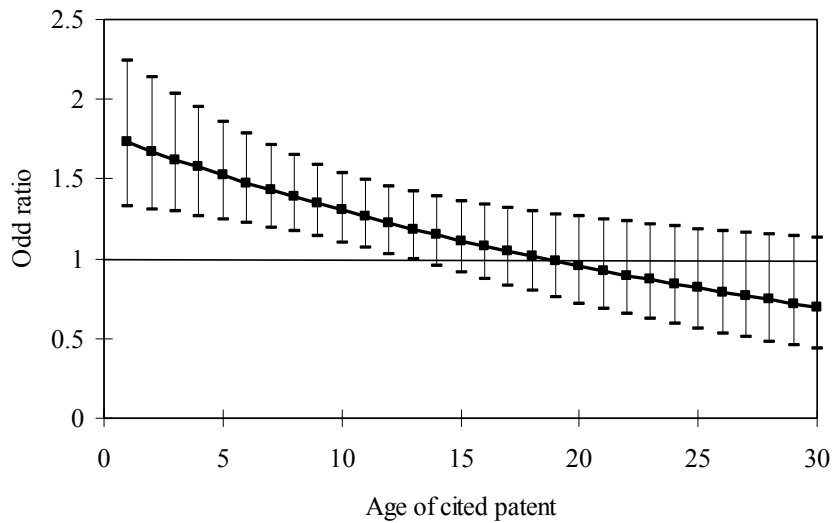


FIGURE 2. Odds ratios for inventor added citations, by age of cited patent.

TABLE 4. *Odds Ratios for Geographic Matching Rates*
Conditioning on Technology Classification Match or Assignee Location

	A. TECHNOLOGY CLASSIFICATION MATCH			B. ASSIGNEE LOCATION	
	INVENTOR CITATION		TECHNOLOGY MATCH	INVENTOR CITATION	
	MATCH	NO MATCH		ASSIGNEE NOT IN CA, TX, MA	ASSIGNEE IN CA, TX, OR MA
MATCH COUNTRY	1.137 (2.04)	1.273 (3.91)	1.047 (0.77)	1.189 (2.95)	1.236 (2.19)
MATCH STATE	1.216 (1.85)	1.427 (3.26)	1.147 (1.24)	1.220 (1.74)	1.450 (2.87)
MATCH CMSA	1.232 (1.44)	1.447 (2.28)	1.317 (1.76)	1.285 (1.49)	1.337 (1.63)
MATCH MSA	1.162 (1.29)	1.499 (3.24)	1.269 (1.87)	1.163 (1.25)	1.600 (2.99)

Z-scores in parentheses. Odds ratios for NON-INSTITUTIONAL and CITED PATENT AGE are similar to previous results and hence are not reported here.

cal developments to a greater extent than elsewhere. Panel B of Table 4 reports separate regressions after excluding California, Texas and Massachusetts, and for these states alone.¹⁵ Consistent with prior evidence, localization effects are stronger in these states, but geography does matter elsewhere.

V. Conclusions

This paper combines Jaffe, Trajtenberg and Henderson’s (1993) innovative use of patent citations to study knowledge flows with a new identification strategy based on differences between geographic matching rates for inventor-added and examiner-added citations. The paper has produced prima facie evidence that knowledge spillovers are geographically localized both internationally and intranationally. It was also found that only intranational localization effects become weaker with the passage of time. These are not surprising results. In particular, the finding that intranational but

¹⁵ These states account for about half the observations in the sample.

not international localization effects decay with time is consistent with the conventional wisdom that geography matters because tacit knowledge is embodied in individual researchers, who relocate frequently within in the United States but only infrequently across international borders.

Of course, these results should be interpreted with caution. Two issues merit particular attention. First, the elimination of self-citations – in this and in all prior work – remains far from satisfactory, in ways that may well generate false localization effects. Although I have manually checked the sample for cases where company names are sufficiently similar to identify self-citations between parents and their subsidiaries, partners, and joint ventures, this effort can only get us so far. One could presumably advance the process using directories of company ownership [e.g. Dun & Bradstreet (1998)]. But, daunting as that task would be, one must then decide when a citation is a self-citation and when it is a spillover. Presumably, the judgment depends on the degree of interaction taking place between related firms. This does not seem to be a criterion that lends itself to measurement.

Second, the analysis fails to distinguish adequately between the effects of geography on knowledge spillovers and the effects of industry boundaries. For example, patent examiners may be more likely than inventors to cite related technologies in different industries, and in so doing we confound the effects of industry with the effects of geography. A first look at the data raises the hope that this is not too great a concern, because for various technological criteria it turns out that examiners are *less* likely to cite across technology classes. Inventor citations match the US primary class 41 percent of the time, compared with 54 percent for examiner citations. Examiner citations also match the US sub-class, the international classification code, and the field of search more often. However, these numbers are only suggestive. Because of the way that examiners undertake searches for prior art, it is possible that they are more likely to cite prior art within the same technology class, while at the same time unobserved heterogeneity within classes implies that they are less likely to cite prior art in the same industry. It is not obvious to me how one might answer this question from information contained in patent data, but it should encourage caution: we cannot be sure whether geography or industry boundaries present the real barrier to knowledge spillovers.

Dealing with these, and possibly other, caveats, will no doubt require more work on both data collection and experimental design. But, before undertaking that considerable effort, there is perhaps a more immediate task: to figure out exactly what it means from the perspective of policy design and welfare to have, say, a 20 percent greater matching rate for inventor citations. The most appropriate line of inquiry here looks to be introducing such differences in spillover rates to an appropriate calibrated growth model. This seems to me to be the next task at hand.

References

- Agrawal, Ajay, Iain M. Cockburn, and John McHale (2003): "Gone but not forgotten: labor flows, knowledge spillovers, and enduring social capital." NBER working paper 9950.
- Agrawal, Ajay., Devesh Kapur, and John McHale (2004): "Defying distance: examining the influence of the diaspora on scientific knowledge flows." University of Toronto: Working paper.
- Almeida, Paul. (1996): "Knowledge sourcing by foreign multinationals: patent citation analysis in the US semiconductor industry." *Strategic Management Journal*, **17**(Winter Special Issue):101-23.
- Audretsch, David B., and Maryann P. Feldman (2003): "Knowledge spillovers and the geography of innovation." In V. Henderson and J.F. Thisse, eds, *Handbook of Urban and Regional Economics*, volume 4, forthcoming.
- Breschi, Stefano, and Francesco Lissoni (2004): "Knowledge networks from patent data: methodological issues and research targets." CESPRI, Milan: Working paper no. 150.
- Chamberlain, Gary (1980): "Analysis of covariance with qualitative data." *Review of Economic Studies*, **47**:225-238.
- Cockburn, Iain M., Samuel Kortum, and Scott Stern (2003): "Are all patent examiners equal? The impact of characteristics on patent statistics and litigation outcomes." NBER working paper 8980.
- Dun & Bradstreet (1998): *Who Owns Whom. North & South America*. -- High Wycombe, UK: Dun & Bradstreet Ltd.
- Feldman, Maryann P. (1994a): *The Geography of Innovation*. Boston: Kluwer Academic Publishers.

- Feldman, Maryann P. (1994b): "Knowledge complementarity and innovation." *Small Business Economics*, **6**(3):363-372.
- Frost, Tony S. (2001): "The geographic sources of foreign subsidiaries' innovations." *Strategic Management Journal*, **22**(2):101-23.
- Glaeser, Edward, H. Kallal, Jose Scheinkman, and Andre Schleifer. (1992): "Growth of cities." *Journal of Political Economy*, **100**(6):1126-1152.
- Grossman, Gene, and Elhanan Helpman (1991): *Innovation and Growth in the Global Economy*. Cambridge, MA: MIT Press.
- Hicks, Diane, Tony Breitzman, Dominic Olivastro, and Kimberly Hamilton, (2001): "The changing composition of innovative activity in the US – a portrait based on patent analysis." *Research Policy*, **30**(4):681-703.
- Jacobs, Jane (1969): *The Economy of Cities*. New York: Random House.
- Jaffe, Adam B., Manuel Trajtenberg, and Michael S. Fogarty (2002): "The meaning of patent citations: report on the NBER/Case Western Reserve survey of patentees." In A.B. Jaffe and M. Trajtenberg, *Patents, Citations, and Innovations: A Window on the Knowledge Economy*. Cambridge: MIT Press, pp. 379-401.
- Jaffe, Adam B., Manuel Trajtenberg, and Rebecca Henderson, (1993): "Geographic knowledge spillovers as evidenced by patent citations." *Quarterly Journal of Economics*, **108**(3):577-98.
- Katz, Ethan (2001): "Bias in conditional and unconditional fixed effects logit." *Political Analysis*, **9**(4):379-384.
- Krugman, Paul R. (1991) *Geography and Trade*. Cambridge, MA: MIT Press, 1991.
- Manski, Charles F. (2000): "Economic analysis of social interaction." *Journal of Economic Perspectives*, **14**:115-136.
- Marshall, Alfred. (1890): *Principles of Economics*, London: Macmillan.
- Romer, Paul M. (1990): "Endogenous technological change." *Journal of Political Economy*, **98**(5, pt. 2):S71-S102.
- Thompson, Peter, and Melanie Fox-Kean (2005): "Patent citations and the geography of knowledge spillovers: a reassessment." *American Economic Review*, **95**(1):450-460.
- US Census (2004): "Annual geographic mobility rates, by type of movement: 1947-2001. www.census.gov/population/socdemo/migration/tab-a-1.txt. Accessed March 15, 2004.