Discussion Paper Series
No. 1004
Jan 2010

Inventor Mobility and Knowledge Transmission in Nanotechnology

Jinyoung Kim, Sangjoon John Lee, and Gerald Marschke

The Institute of Economic Research - Korea University
Anam-dong, Sungbuk-ku, Seoul, 136-701, South Korea, Tel: (82-2) 3290-1632, Fax: (82-2) 928-4948

Inventor Mobility and Knowledge Transmission in Nanotechnology*

Jinyoung Kim†, Sangjoon John Lee††, and Gerald Marschke†††

†Korea University
††Alfred University
†††University at Albany-SUNY, NBER and IZA

PRELIMINARY—PLEASE DO NOT QUOTE WITHOUT PERMISSION

April, 2008

Abstract

Using U.S. patent records in nanotechnology, we study the relationship between inventor mobility among firms and knowledge diffusion. We find evidence consistent with a story that, in one important nanotechnology subfield, when inventors move among firms they spread knowledge. In particular, we find that if we consider any two patents in the "Chemicals, misc." subclass, A and B, where A and B are assigned to different firms and where A is granted after B, patent A is more likely to cite patent B if the patent A firm employs an inventor who earlier worked for the patent B firm.

JEL Classification: J62, O31, O33
Key words: Nanotechnology; Patents; Innovation; Knowledge spillovers;

---

*Prepared for "The Nanobank Research Conference," May 3, 2008, Cambridge, MA. All errors are exclusively the responsibility of the authors. Comments welcome at Jinyoung Kim, Department of Economics, Korea University, 5-1 Anam-Dong, Sungbook-Ku, Seoul, Korea, jinykim@korea.ac.kr; and Gerald Marschke, Department of Economics, SUNY Albany, BA-110, Albany, NY 12222, marschke@albany.edu.

I. Introduction

Nanotechnology is one of the fastest-advancing fields in science and technology and is anticipated to make great contributions to and possibly fundamentally transform many large and economically important industries such as information technology and medicine. Nanotechnology thus has the potential to have a major impact on technological progress and economic growth. The speed with which the benefits of nanotechnology are realized and who ultimately enjoys these benefits depend on how nanotechnology diffuses and is transmitted throughout the economy and across economies. Economists and other scholars who study science and technology have long suspected that the inter-firm mobility of scientists transmits technological know-how across firms (Arrow, 1962; Stephan, 1996), but evidence is often anecdotal and econometric evidence is scarce. In this paper we study the link between the mobility of innovators and technological diffusion. When innovators patent, they leave a paper trail. We use this trail to follow inventors from firm to firm and measure whether an employer's innovation reflects the employment history of its inventors.

Knowledge diffuses across organizations via a number of potential paths, for example, through published patents, papers, and textbooks, at conferences where research is presented and where industry and academic research personnel comingle, and via informal social networks.[1] But a firm also learns about other firms' research after employing or collaborating with innovators who work or have worked in other firms' laboratories. In fact, social scientists who study innovation suspect that certain kinds of important knowledge—*tacit* knowledge—become available to a firm only with sustained,

[1] See Cohen, Nelson and Walsh (2002) on the various means by which innovating firms access know-how developed externally. See Agrawal, Cockburn, and McHale (2003) for evidence of the importance of social networks in promoting diffusion.

close interaction with researchers who possess this knowledge as through an employment or collaborative research arrangement.

We propose to examine the role of research personnel as a pathway for the diffusion of ideas in nanotechnology, utilizing information contained in U.S. patent records. The inventors behind the patented invention are listed on each patent, as are the firms, government organizations, and universities to which the patents are assigned. Using a procedure proposed in Kim, Lee, and Marschke (2008), we match names on patents to construct a panel data set of inventors that contains the patents in each year of the inventors' careers. Because the assignee is usually the inventor's employer, the inventor reveals her employer when she invents. On the occasions that the assignee is not the inventor's employer, often the inventor is working under an inter-organizational research agreement of some kind. Thus we interpret the appearance of multiple assignees among an inventor's patent portfolio as evidence of either job mobility or collaborations across organizational boundaries. Hiring an inventor formerly employed elsewhere and collaborating with another firm in an R&D project are both means by which tacit knowledge can be transmitted. For inventors that invent frequently the U.S. patent data yield the employment and collaboration history of inventors.

At the same time, one can use patent citations to infer the extent to which a patent in one firm accesses the knowledge generated by inventors in other firms. Patent applicants are legally obligated to disclose any knowledge they have of previous relevant inventions. The patent examiner may add to the application relevant citations omitted by the applicant. Thus, through the patent citations each patent documents the "prior art" upon which the new innovation builds. These citations provide an additional window on the

pathways of knowledge (for evidence that citations proxy for knowledge flows, see Jaffe, Fogarty, and Banks, 1998; and Duguet and MacGarvie, 2005).

Our principle finding is consistent with a story that, in one important nanotechnology subfield, when inventors move among firms they spread knowledge. In particular, we find if we consider any two patents in the "Chemicals, misc." subclass, A and B, where A and B are assigned to different firms and where A is granted after B, patent A is more likely to cite patent B if the patent A firm employs an inventor who earlier worked for the patent B firm.

The paper is organized as follows. The next section summarizes the literature on technology spillovers, scientist collaboration and mobility, and the use of patent citations to trace technological diffusion. Section III describes our data, focusing on the construction of the inventor panel. Section IV describes our empirical strategy and Section V describes our results. Section VI offers concluding remarks.

II. Literature Review

Nanotechnology is a multidisciplinary area of applied science that includes such disciplines as physics, chemistry, and engineering. According to a definition proposed by the U.S. National Nanotechnology Initiative, nanotechnology involves the control and understanding of matter at the approximate scale of 1 to 100 nanometers (nm)—a nanometer is one billionth of a meter—and the assembly of useful devices in the nanoscale atom by atom or molecule by molecule. An attraction of the nanoscale environment is that the physical and chemical properties of matter change as matter's scale gets very small. A material may become stronger gain transparency or conductivity

in the nanoscale. For example, a nanoscale tube of carbon, approximately 1/100,000 the diameter of a human hair, is very strong. Materials composed of carbon nanotubes hold the potential to replace steel in cars, for example, vastly increasing cars' gas mileage. Carbon nanotubes possess superior heat and electricity conductivity making them ideal materials for electronic devices. The applications of nanotechnology so far, however, have been fairly mundane. The main commercial applications of nanotechnology are in suntan lotions, cosmetics, and surface coatings.

This paper focuses on how technologies—nanotechnologies in particular—diffuse through the economy. Understanding how knowledge spillovers within and across economies work is of interest because of the role spillovers may play in economic growth and because of its implications for science and technology policy. Studies in both the economics and sociology of innovation literatures argue that new technologies are frequently "tacit" and difficult to transmit to the uninitiated via spoken or written communication (Polyani, 1958, 1966). The most efficient means of transmission across organizational boundaries for tacit knowledge may be via person-to-person contact involving a transfer or exchange of personnel. Recent findings that technological diffusion appears to be geographically limited (e.g., Jaffe, 1989; Jaffe, Trajtenberg, and Henderson, 1993; Audretsch and Feldman, 1996; Zucker, Darby, and Brewer, 1998; and Mowery and Ziedonis, 2001) are often interpreted as evidence of the tacitness of knowledge (e.g., Feldman, 1994). Some survey evidence exists that person-to-person interaction is important for the diffusion of technology. Cohen, Nelson, and Walsh (2002) surveyed R&D managers on the means by which they gather and assimilate new technologies and find that firms access externally-located technology partly through the

hiring of and collaboration with researchers from the outside. Moreover, they find that hiring/collaboration with outside researchers is complementary to other means of accessing externally produced knowledge, such as through informal communications with outsiders and more formal (such as consulting) relationships with outsiders.

Much of the literature that examines the mobility of scientists and innovators as a source of knowledge transmission focuses on the movement of academic scientists from academe to industry. Certainly universities and academic ideas are important to the high-technology sector. A number of recent papers offer evidence for geographically localized spillovers occurring in areas around major universities (Jaffe, 1986, 1989; Audretsch and Feldman, 1996; Henderson et al, 1998), suggesting both that academe is an important source of commercially-important ideas and that such ideas are not easily transmitted from the university labs in which they originate to the firms where they can be turned into commercial products. Work by Jensen and Thursby (2001), Agrawal and Henderson (2002), and Thursby and Thursby (2002) find that the best predictor that an academic idea leads to successful product roll-out is the participation of the inventing scientist. Thus, the hands-on involvement of academic scientists may in fact be necessary for an academic idea to take root in industry. In the biotechnology sector, Darby, Zucker and co-authors have examined the importance of working relationships between firms' bench scientists and top academic, or "star", scientists. They find that firms in the U.S. and Japan are more likely to enter the biotechnology industry in regions where star academics publish (Zucker, Darby, and Armstrong, 1998, 2001; Zucker, Darby, and Brewer, 1998; and Zucker and Darby, 2001). They also find that university influence on nearby firm

R&D productivity exists almost exclusively in firms whose bench scientists have working relationships with star academic scientists.

Darby, Zucker and co-authors also examine these issues with respect to the nanotech industry. Zucker and Darby (2007) and Darby and Zucker (2003) find evidence of regional agglomeration in both science and commercial applications of nanoscience, with firm entry in nanotechnology occurring in the vicinity of the major research universities publishing in scientific papers. Zucker, Darby, Furner, Liu, and Ma (2007) find counts of patents and articles in nanotechnology are correlated with the degree of collaboration observed between industry and universities (as measured by co-authorship on scientific papers) suggesting that knowledge transfer takes place via collaboration and this is conducive to nanotechnology progress.

The evidence suggests that the interaction between universities and firms, and possibly the movement of academic scientists to industry are important in the development of high technology, generally, and the nanotechnology industry, in particular. In this paper, we look at the mobility from firm to firm, within industry, of non-academic innovators as a source for knowledge transmission. Rates of mobility of computer scientists, engineers and scientists in industry are very high in many regions of the country. Job-hopping is such a part of the landscape in places like Silicon Valley, that engineers relate, presumably jokingly, that opportunities are so plentiful one can change jobs without changing carpools and that they consequently switch jobs so often, that they don't even bother to tell their spouses.

But econometric evidence that job to job mobility facilitates transmission of knowledge is indirect and circumstantial. Almeida and Kogut (1999) find that firms are

more likely to cite patents of other firms in their region if inventor mobility rates are high, offering circumstantial evidence that ideas in the semiconductor industry are spread by the movement of key engineers among firms, especially within a geographical area.

Kim and Marschke (2005) develop and test a model of the patenting and R&D decisions of an innovating firm whose scientist-employees sometime quit to join or start a rival. In their model, the innovating firm patents to protect itself from its employees. Kim and Marschke find that firms are more likely to patent in environments where scientists are likely to switch employers, suggesting firms perceive a threat that their workers will pass along proprietary know-how.

If such technical knowledge acquired by the researcher in an employer's lab can be transmitted to future employers, then such knowledge is a form of general human capital. Researchers would be willing to accept lower wages to acquire technical knowledge that they can exploit with multiple employers. Moen (2005) finds some evidence of this: he shows that technical workers in R&D intensive firms in Norway accept lower wages early in their career in exchange for higher wages later, evidence that workers trade on the intellectual property they acquire in an employer's lab.

Stolpe (2002) examines knowledge diffusion in the field of liquid crystal display. He employs an empirical strategy very similar to the one employed in this paper. He uses a firm's patent citations to evaluate one firm's access of the knowledge located at another firm. He finds that sharing inventors does not significantly influence two firms' likelihood of citing one another and he interprets this finding as evidence that the mobility of innovators is not an important means by which knowledge diffuses in the LCD industry.

III. Data

The focus of this paper is how firm-to-firm mobility of innovators transmit knowledge in the nanotechnology industry. We use patent citations to trace knowledge flows between firms. We use a unique panel data set on inventors that allows us to identify the firms for which inventors have invented. Using these data we can test whether a firm's likelihood of citing the patents of other firms reflect the patenting history of its inventors. In this way, we hope to learn whether inventors' mobility influences the diffusion of new knowledge.

The data used in this paper are a part of the inventor-firm panel database that we created (see Kim, Lee and Marschke, 2008, for the detailed description of the construction procedure). The data for this paper are derived from three sources: (1) Patent Bibliographic data (Patents BIB) released by the U.S. Patent and Trademark Office (USPTO) that contain bibliographic information for U.S. utility patents issued from 1975 to 2002; (2) the NBER Patent-Citations data collected by Hall, Jaffe and Trajtenberg (2001) which contain all citations made by patents granted in 1975-1999; and (3) the Nanobank database collected by Zucker and Darby (2007) that identifies patents in nanotechnology. To create our data from these sources, we match inventor names in the Patents BIB database, and add information from the citation data. We then select only those patents in nanotechnology based on the Nanobank database. The following describes the name matching method.

Inventor name matching

Since the 1960's the information contained in patent data have been extensively used to investigate various issues such as technology spillovers and R&D productivity at the industry or firm level. The information on inventors contained in patent data, however, has not been fully utilized possibly because of the difficulty in identifying whether two names in the inventor name field from two patents belong to the same inventor. Using inventor's name (last, first, and middle), address, city, state, zip (often missing), and country at the time of grant of the patent, we attempt in this paper to match inventor names and produce each inventor's life-cycle productivity in patenting.

To start, we treat each entry that appears in the inventor name field of every patent in the Patents BIB data as a unique inventor. Given N number of names in this name pool, we pair each name with all other names, which generates N(N-1)/2 number of unique pairs. The 5.1 million names in the Patents BIB data (2.05 inventors per patent) thus produce 13 trillion unique pairs. For each pair, we consider the two names as belonging to the same inventor if the SOUNDEX codes of their last names and their full first names are the same, and at least one of the following three conditions is met: (1) the full addresses for the pair of names are the same; (2) one name from the pair is an inventor of a patent that is cited by another patent whose inventors include the other name from the pair; or (3) the two names from the pair share the same co-inventor. In implementing the second and third conditions, we make comparisons based on whether the first and last names are spelled identically. After our name matching procedure is completed, we go back and check that these conditions are still valid based on the inventor identifier constructed by the matching procedure. If not, we repeat the name matching process to create a new inventor identifier.

SOUNDEX is a coded index for last names based on the way a last name sounds in English rather than the way it is spelled. Last names that sound the same, but are spelled differently, like SMITH and SMYTH, have the same SOUNDEX code. We use the SOUNDEX coding method to expand the list of similar last names to overcome the potential for misspellings and inconsistent foreign name translations to English; misspellings are common in the USPTO data as are names of non-Western European origin (see Appendix A for the detailed SOUNDEX coding method).

We also consider a pair of names as a match if two have the same full last and first names as spelled in the Patent BIB data, and at least one of the following conditions is met: (1) the two have the same zip code; (2) they have the same full middle name; or (3) they reside in the same MSA area. Given all pairs of names that are considered as matches by the preceding procedures, we impose an additional matching criterion that a pair of names is not treated as a match if their middle name initials are different. We then impose transitivity in the following sense: If name A is matched to name B and name B is matched to name C, name A is then matched to name C. We iterate this process until all possible transitivity matches are completed. At this point we assign the same inventor ID number for all the names matched. Using this method, we identified 1.72 million unique inventors (34%) out of 5.1 million names in the entire patent data.

After name matching, we add information on all citations from the NBER Patent-Citations data collected by Hall, Jaffe and Trajtenberg (2001) where each citing patent that was granted between 1975 and 1999 is matched to all patents cited by the patent. As the final step, we select only those patents in nanotechnology that are identified in the Nanobank database (see Zucker and Darby, 2007, for a description of these data).

IV. Empirical Method

Nanoscience and nanotechnology can be found across a variety of disciplines and applications. Figures 1 and 2 show the patent counts by major technological category by year (we use the aggregated technological classification system of Hall, Jaffe and Trajtenberg, 2001). Figure 1 shows all U.S. patents and Figure 2 shows only those nanotechnology patents. Figure 2 shows that the chemistry field dominates nanotechnology patents early on and remain dominant throughout the period. Note that over 25% of chemical patents are classified as nanotechnology. Patents in the Electrical and Electronic and Drugs and Medical fields begin to take off in the mid-1980s.

We are interested in estimating the determinants of the decision to cite a patent. Our independent variable of interest is employment mobility. In particular, we wish to discover whether firm A's patent is more likely to cite firm B's patent, if an inventor at firm A once patented for firm B. We examine only those nanotechnology patents in patent technological category "Chemicals, miscellaneous" (USPTO subcategory 19). This is the largest subcategory in the nanotechnology patent population. By isolating a single technology we reduce at least some unobserved heterogeneity. We wish to estimate the determinants of a patent's citing another patent. One strategy for estimating the citation probability is to estimate for each feasible patent pair combination the probability that the newer patent cites the older one. Because for each citing patent there are thousands of potential or cite-able patents we choose an empirical strategy similar to strategies adopted elsewhere in the patenting literature (Singh, 2006; Stolpe, and Jaffe et al, 1993). We estimate a weighted logit (Manksi and Lerman, 1977), where the

probability modeled is the probability that a nanotech patent in the subcategory 19 cites another patent in the subcategory 19.

The sample used in the analysis consists of all nanotechnology patents in subcategory 19 that cites any patent in subcategory 19. To produce the comparison group we draw a random subsample of nanotech patents from subcategory 19 that do not cite another patent in subcategory 19. Each citing patent in the random sample is then matched randomly to one cite-able patent from the subcategory (pairs of patents from the same assignee are dropped). Thus the comparison group consists of a random subset of citing and cite-able patent pairs, in which the citing patent does not cite the cite-able patent. We set our sampling rates to yield a 2 to 1 ratio of non-citing to citing patents. In the estimation of the logit, the randomly sampled patent pairs in the comparison group are weighted up by the inverse of their probability of being sampled so that their contribution to the likelihood is proportional to their numbers in the population.

The focus of the analysis is the effect of sharing an inventor on the probability that a patent by one firm cites a patent by another. Our measure of mobility, *MOBILITY*, is an indicator variable which is coded 1 if an inventor on the citing patent was also an inventor on a patent assigned to the assignee of the cited or cite-able patent filed sometime between five years prior to the cited/cite-able patent's filing date and the citing patent's filing date. Suppose for example, patent B is filed on June 1, 1980 by assignee b and A is filed on June 1, 1990 by assignee a. If an inventor on A was an inventor on any patent assigned to b between February 1, 1980 and June 1, 1990 then the dummy is coded 1, else it is coded zero.

Table 1 reports the means of variables used in our analysis.[2] In estimating the determinants of a citation, we include, in addition to MOBILITY, CITELAG (and its square), the time in years that have elapsed between the cited/cite-able patent's application date and the citing patent's application date. We hypothesize that the more recent the patent, the greater its likelihood of being cited. We also include LCLAIMS_A and LCLAIMS_B, respectively the log of the number of claims made by the citing and cited/cite-able patents. The number of claims represents the number of pieces or "building blocks" to the underlying innovation. Hence, the number of claims may be informative of the innovative territory covered by the patent. We expect that patents that cover more ground are more likely to be cited. The number of citations received in the first five years following the patent's grant date, CRECEIVE_A for the citing patent and CRECEIVE_B for the cited-cite-able patent, is included as a measure of patent importance. We expect that a patent is more likely to cite another patent if it is important (has high CRECEIVE_B).

We also include Hall, Jaffe, and Trajtenberg's measures of patent originality and generality. Let s_{ij} be the share of citations received by patent that belong to patent class *j*, out of n_i patent classes. Their measure of generality is $Generality_i = 1 - \sum_{1}^{n_i} s_{ij}^2$. Note that this is one minus the Herfindahl index: the more diffuse the contribution (the greater the number of classes the citations cover) the larger this index. Their measure of originality is defined the same way except it is the citations *made* that are used in the calculation. Thus, if the patent cites patents in only a few technologies, the originality measure will be low,

[2] We do not weight the data in the production of the descriptive statistics. Thus patents in the subclass 19 that cite other patents in the same subclass are overrepresented, and patents that do not cite are underrepresented.

whereas if it cites patents across many fields, then the measure will be high. We anticipate that patent A is more likely to cite patent B if B is more general and more original.

V. Results

Table 2 presents the results from the estimation of various weighted logit models. Models I and II include all data in the analysis, and thus include assignees that are firms, universities, and government research labs. Both the constrained (I) and unconstrained (II) models produce a coefficient estimate for MOBILITY that is positive and strongly statistically significant. Because we are especially interested in the mobility among firms, as opposed to mobility between universities and industry, we estimated a logit that contains only patents assigned to industry. The results of this estimation are depicted in model III. The coefficient estimate on MOBILITY is positive and significant, both in the statistical and economic senses. The fourth column depicts the implied marginal effects of the coefficient estimates from model III. Note that having an inventor on patent A who at one time worked for or collaborated with the assignee of patent B, increases A's probability of citing B by about .08.

Many of the other coefficient estimates have the anticipated signs. For example, A is more likely to cite B if B is more recent, wider in scope, more general, and more important; most coefficient estimates are statistically significant but this is because of the large number of observations. As column four indicates, however, few coefficients imply

an important economic or behavioral relationship between the corresponding independent variables and the probability of a citation.

VI. Discussion

We find evidence consistent with a story that, in one important nanotechnology subfield, when inventors move among firms they spread knowledge. In particular, we find if we consider two patents in the "Chemicals, misc." subclass, A and B, where A and B are assigned to different firms and where A is granted after B, patent A is more likely to cite patent B if the patent A firm employs an inventor who earlier worked for the patent B firm.

This evidence is consistent with other stories as well, and future work will focus on ruling out the "mobility causing knowledge" diffusion story. Omitted variables make it unwise to infer from our results that mobility is causing citations. For example, different firms share research agendas for reasons unrelated to worker mobility. In addition, inventors have varied skill sets which match better with some research agendas than others. Thus, we would expect that inventors when they move tend to remain in the same research areas and the fact that two firms share an inventor is an indication of, and not a cause of, similar research agendas. And of course firms with similar research areas will cite one another's patents.

We also know that firms which are engaged in similar R&D programs congregate in the same geographical cluster. Moreover, workers who switch jobs tend to remain in the same geographical area. Thus the fact that two firms share the same inventor is an indication of geographical proximity and, again, similar research programs and a higher

propensity to cite one another's work. This concern may be at least partly addressed by including dummies indicating whether the citing and cited firms are in the same geographical cluster. By focusing on patents in a single subclass, we had hoped to reduce this kind of heterogeneity, but this subclass is large and may be quite heterogeneous. In future work, we will experiment with more narrowly defined and homogenous technologies.

Finally, we realize that there are limits to what we can learn about the role of inventor mobility in the diffusion of knowledge from patent citations and patent based measures of worker mobility. Patents capture only a small part of the economically important knowledge created and diffused through the economy. Citations are noisy measures of the important precursors of innovations. In addition, because we observe inventors only when they patent, we miss much inventor flow between employers. This latter concern is less troubling, however, if most diffusion occurs via the movement of "star" industrial inventors. Star inventors publish frequently, making them easier to follow through their career.

Appendix A. SOUNDEX coding system

A SOUNDEX code for a surname is an upper case letter followed by 6 digits. For example the SOUNDEX code for Kim is K500000, while that for Marschke is M620000. The first letter is always the first letter of the surname. The rules for generating a SOUNDEX code are:

1. Take the first letter of the surname and capitalize it.
2. Go through each of the following letters giving them numerical values from 1 to 6 if they are found in the Scoring Letter table (1 for B, F, P, V; 2 for C, G, J, K, Q, S, X, Z; 3 for D, T; 4 for L; 5 for M, N; 6 for R; 0 for Vowels, punctuation, H, W, Y).
3. Ignore any letter if it is not a scoring character. This means that all vowels as well as the letters h, y and w are ignored.
4. If the value of a scoring character is the same as the previous letter then ignore it. Thus if two ‘t’s come together in the middle of a name they are treated exactly the same as a single ‘t’ or a single ‘d’. If they are separated by another non-scoring character then the same score can follow in the final code. The name PETTIT is P330000. The second ‘T’ is ignored but the third one is not since a nonscoring ‘I’ intervenes.
5. Add the number onto the end of the SOUNDEX code if it is not to be ignored.
6. Keep working through the name until you have created a code of 6 characters maximum.
7. If you come to the end of the name before you reach 6 characters, pad out the end of the code with zeros.
8. Optionally you can ignore a possessive prefix such as ‘Von’ or ‘Des’.

See "Using the Census SOUNDEX," General Information Leaflet 55 (Washington, DC: National Archives and Records Administration, 1995) for the detailed method.

References

Adams, James D. 1990. "Fundamental Stocks of Knowledge and Productivity Growth." *The Journal of Political Economy*, 98(4), pp. 673-702.

Adams and Jaffe (1996) "Bounding the Effects of R&D" *Rand Journal of Economics* (27), 700-721.

Agrawal, Ajay K., Cockburn, Iain M., and McHale John 2003 "Gone but Not Forgotten: Labor Flows, Knowledge Spillovers, and Enduring Social Capital," NBER Working Paper no. 9950

Agrawal, A. and Henderson, Rebecca. (2001) "Putting Patents in Context," *Management Science*, 48(1): 44-60.

Arrow, K.J. "Economic Welfare and the Allocation of Resources for Invention." In R.R. Nelson, ed., *The Rate and Direction of Inventive Activity: Economic and Social Factors*, *NBER Special Conference Series*. Vol. 13, Princeton University Press, New Jersey, 1962.

Audretsch and Feldman, 1996. "R&D Spillovers and the geography of innovation and production," *American Economic Review*, 86, pp. 630-640

Cohen, Wesley M., Nelson, Richard R., and Walsh, John P. 2000. "Protecting Their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (or Not)," NBER Working Paper no. 7552

Darby, Michael R. and Zucker, Lynne G. (2003). "Grilichesian Breakthroughs: Inventions of Methods of Inventing and Firm Entry in Nanotechnology," NBER Working Papers 9825.

Duguet, Emmanuel and MacGarvie, Megan. 2005. "How well do patent citations measure flows of technology? Evidence from French innovation surveys," *Economics of Innovation and New Technology*, 14(5), 375-393

Hall, B., Jaffe, A. and Trajtenberg, M. "Market Value and Patent Citations: A First Look." NBER Working Paper No. 7741, 2000.

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

Jaffe, Adam B. (1986) "Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits and Market Value." *American Economic Review*, 76, 5, pp. 984-1001.

Jaffe, Adam B. (1989) "Real Effects of Academic Research." *American Economic Review,* 79 957-970

Jaffe, Adam B., Fogarty, Michael S., and Banks, Bruce A. (1998) "Evidence from Patents and Patent Citations on the Impact of NASA and Other Federal Labs on Commercial Innovation," *Journal of Industrial Economics*, 46 (2) , 183–205

Jaffe, A.B., Henderson, R., and Trajtenberg, M. (1993) "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations." *Quarterly Journal of Economics,* 108(3), pp. 577-598.

Jensen, Richard and Thursby, Marie (2001) "Proofs and Prototypes for Sale: The Tale of University Licensing," *American Economic Review*, 91(1): 240-259.

Kim, Jinyoung and Marschke, Gerald, 2005, "Labor Mobility of Scientists, Technological Diffusion, and the Firm's Patenting Decision," *The Rand Journal of Economics*, forthcoming, 2005.

Manski, C. F. and Lerman, S.R. (1977) "The Estimation of Choice Probabilities from Choice Based Samples," *Econometrica*, (45), pp. 1977-1988.

Moen, Jarle. (2005) "Is Mobility of Technical Personnel a Source of R&D Spillovers?," *Journal of Labor Economics*, 23(1), pp. 81-114.

Polanyi, Michael. *Personal Knowledge: Towards a Post-Critical Philosophy* (University of Chicago Press, Chicago, 1958)

Polanyi, Michael. *The Tacit Dimension* (Doubleday, Garden City, NY, 1966)

Singh, Jasjit. (2006) "Asymmetry of Knowledge Spillovers between MNCs and Host Country Firms," *Journal of International Business Studies,* forthcoming.

Thursby, Jerry and Thursby, Marie. (2002) "Who is Selling the Ivory Tower? Sources of Growth in University Licensing," *Management Science*, 48(1): 90-104.

Trajtenberg, Manuel, "The Names Game: Using Inventors Patent Data in Economic Research," Working paper, 2004.

Trajtenberg, Manuel. "A Penny for Your Quotes: Patent Citations and the Value of Innovations." *The Rand Journal of Economics.* Vol. 21, No. 1, pp. 172-87, Spring 1990.

Zucker, Lynne G. and Darby, Michael R.. (2007) "Socio-Economic Impact of Nanoscale Science: Initial Results and NanoBank." In: Roco, M.C., Bainbridge, W.S. (Eds.), *Nanotechnology: Societal Implications*. Springer, The Netherlands.

Zucker, Lynne G. and Darby, Michael R.. (2001) "Capturing Technological Opportunity via Japan's Star Scientists." *Journal of Technology Transfer*, 26, 37-58.

Zucker, Lynne G., Darby, Michael R., and Armstrong, Jeffery S. (1998) "Geographically Localized Knowledge: Spillovers or Markets." *Economic Inquiry* 36: 65-86.

Zucker, Lynne G.; Darby, Michael R.; Armstrong, Jeff S. "Commercializing Knowledge: University Science, Knowledge Capture, and Firm Performance in Biotechnology" NBER Working Paper No. 8499, October 2001

Zucker, Lynne G., Darby, Michael R., and Brewer (1998) "Intellectual Human Capital and the Birth of U.S. Biotechnological Enterprises." *American Economic Review* 88, 290-306

Figure 1
All Patents by Field

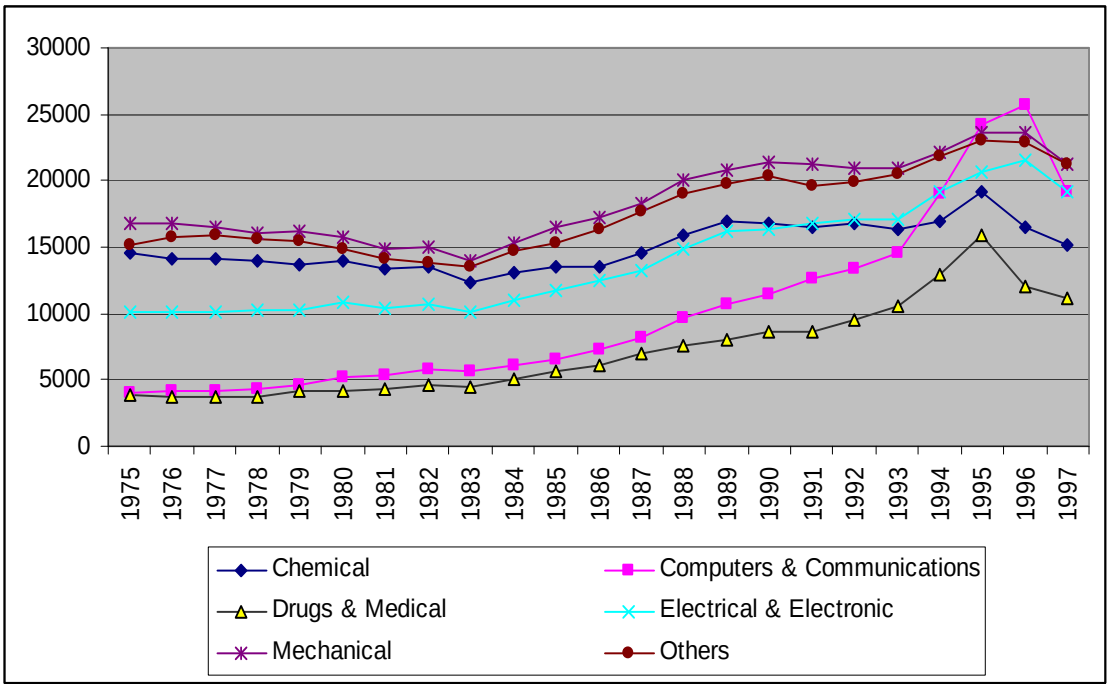

Figure 2
Nanotechnology Patents by Field

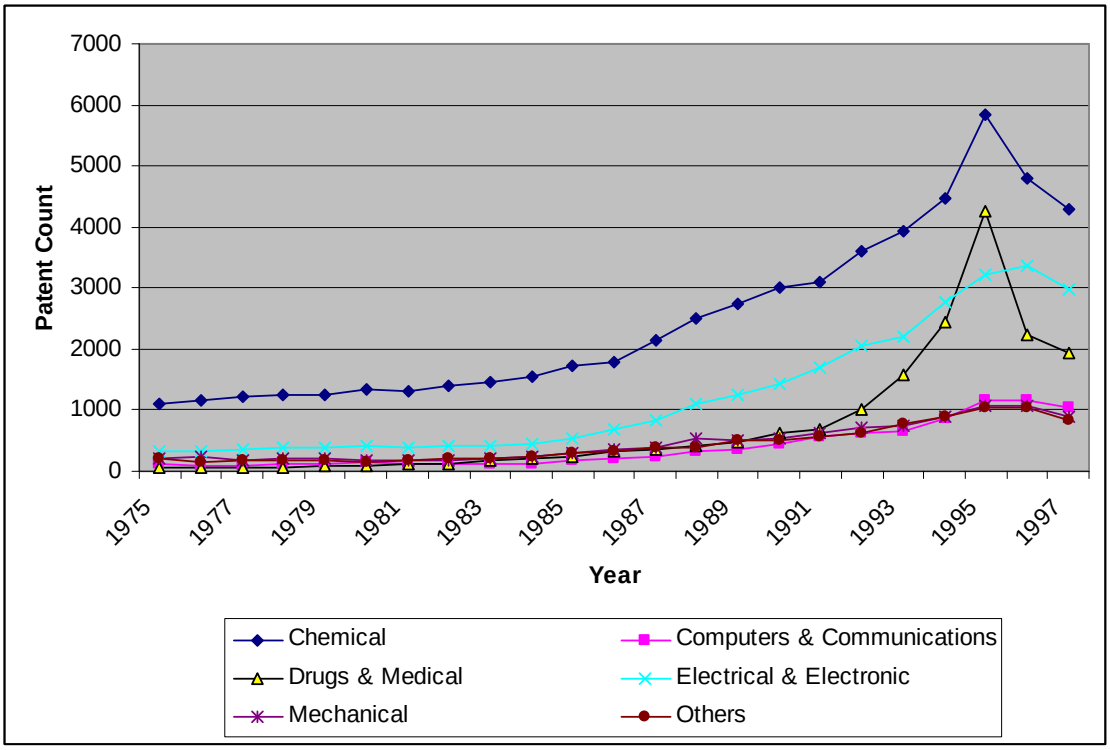

Table 1
Descriptive Statistics
Nanobank Patents from Subclass 19 ("Chemicals, Miscellaneous")

Variable	Definition	Obs.	Mean	Std. Dev.
CITE	Indicator equal to one if patent cites another subclass 19 patent	469181	.335	.472
MOBILITY	Indicator equal to one if citing patent and assignee of cited/cite-able patent share an inventor	469181	.006	.076
LCLAIM_A	Log of the number of claims made by citing patent	394353	2.721	.756
LCLAIM_B	Log of the number claims made by cited/cite-able patent	436496	2.327	.794
CITELAG	Number of years between application dates of citing and cited/cite-able patent	469181	8.111	5.446
CRECEIVE_A	Number of citations received by citing patent in 5 years following grant date	469181	4.170	7.770
CRECEIVE_B	Number of citations received by cited/cite-able patent in 5 years following grant date	469181	10.335	17.216
CMADE_A	Number of citations made by citing patent	469181	16.517	36.939
CMADE_B	Number of citations made by cited/cite-able patent	439206	8.084	8.478
GENERAL_A	Generality, citing patent	297967	.326	.292
GENERAL_B	Generality, cited/cite-able patent	421936	.391	.279
ORIGINAL_A	Originality, citing patent	463655	.439	.278
ORIGINAL_B	Originality, cited/cite-able patent	429271	.380	.279

Table 2
Estimating Determinants of Citing Patent in Same Class
Weighted Logit
("Probability that Patent A cites Patent B")

Variable	Model I All	Model II All	Model III Industry→ Industry	$\frac{d\Pr(CITE=1)}{dx}$
Intercept	-9.55 (.003)	-10.701 (.133)	-10.781 (.110)	
MOBILITY	7.010 (.448)	7.390 (.493)	7.331 (.496)	.080
CITELAG		-.154 (.009)	-.153 (.010)	-.00006†
$CITELAG^2$		.014 (.001)	.014 (.001)	.00002†
LCLAIMS_A		.096 (.023)	.119 (.024)	.000007
LCLAIMS_B		.113 (.014)	.108 (.014)	.000006
CMADE_A		.015 (.001)	.014 (.001)	.000009†
CRECEIVE_A		.016 (.002)	.019 (.001)	.000006†
CMADE_B		-.020 (.003)	-.018 (.003)	-.000008†
CRECEIVE_B		.027 (.002)	.028 (.002)	.00001†
GENERAL_A		.408 (.057)	.468 (.054)	.000008†
ORIGINAL_A		-.046 (.043)	-.040 (.038)	-.000001†
GENERAL_B		.941 (.039)	.867 (.043)	.000017†
ORIGINAL_B		-.343 (.065)	-.413 (.075)	-.000009†
Obs. Wald χ^2 (p val.) Log likelihood	469181 245.3 (.0000) -355.476	240599 5435.6 (.0000) -185.162	225454 5302.1 (.0000) -173.367	

Robust standard errors in parentheses. Marginal effects are from model III.
† $d\Pr(CITE=1) / d\ln x$