Research Policy 38 (2009) 994–1005

Contents lists available at ScienceDirect

Research Policy

journal homepage: www.elsevier.com/locate/respol

Measuring knowledge spillovers: What patents, licenses and publications reveal about innovation diffusion

Andrew J. Nelson

Lundquist College of Business, University of Oregon, Eugene, OR 97405, United States

ARTICLE INFO

Article history:
Received 13 February 2008
Received in revised form
10 September 2008
Accepted 23 January 2009
Available online 4 March 2009

Keywords:
Patents
Licenses
Publications
Innovation diffusion
Knowledge spillovers

ABSTRACT

Measurement of knowledge spillovers remains an important challenge. While patent citation analyses are one common empirical approach, questions persist about their efficacy and potential biases. In an effort to assess various measures of knowledge diffusion, this paper compares patent data surrounding recombinant DNA technology to licenses and publications building on the same technology. Evaluation of these measures highlights errors of both omission and over-representation in each measure, and reveals potential biases tied to organizational age and location. The results suggest that studies of knowledge diffusion can be strengthened dramatically by drawing upon multiple indicators.

1. Introduction

Knowledge diffusion and knowledge spillovers have received significant attention in the economics, management and public policy literatures (e.g., Audretsch and Feldman, 1996; Breschi and Lissoni, 2001; Feldman and Kelley, 2006; Fritsch and Franke, 2004; Griliches, 1992; Jaffe, 1989; Krugman, 1991a,b; Owen-Smith and Powell, 2004; Romer, 1986). But, the growth of this literature has added new urgency to a fundamental associated question: How can knowledge spillovers be measured? Perhaps the most common empirical approach relies on patent citations to serve as indications of knowledge flows (e.g., Acs et al., 1992, 1994; Jaffe, 1989; Jaffe et al., 1993, 2000; Thompson and Fox-Kean, 2005). Recent research, however, has questioned the efficacy and validity of patent citation measures (e.g., Brouwer and Kleinknecht, 1999; Graham and Higgins, 2007; Sampat, 2005). These studies, in the aggregate, suggest that patents and patent citations may both under-represent innovation by failing to capture all innovative activity and over-represent innovation by capturing inventive activities that are of little economic import.

While some studies have compared patent counts and patent citations to R&D expenditures and/or survey data in order to assess the efficacy of patent indicators (e.g., Acs and Audretsch, 1989; Duguet and MacGarvie, 2005; Jaffe et al., 2000; Kaiser, 2002; Scherer, 1983), technology licenses and publications have been underutilized as comparable measures of knowledge diffusion. These additional measures are advantageous for several reasons. I argue that patent-linked technology licenses may be a highly accurate measure of downstream knowledge utilization since the competing economic interests on the part of licensor and licensee act as a system of "checks and balances" to ensure correctness. The importance of publications, meanwhile, is reflected in survey data on the sources and channels of knowledge that contribute to firms' innovation efforts. In reporting on the results of a survey in the U.S. manufacturing sector, Cohen et al. (2002) found that publications are the dominant channel by which knowledge flows from the public sector. Similarly, Agrawal and Henderson (2002) reported that the MIT faculty whom they interviewed consider publications to be two-and-one-half times more important than patents as a knowledge channel. When a single invention is both patented and published (c.f. Murray, 2002; Murray and Stern, 2007), publications can be especially informative.

To assess the extent to which these different measures capture knowledge spillovers[1] and to explore the potential biases of each measure, this paper introduces a novel dataset of matched patents, licenses and publications flowing from recombinant DNA technology—an advantageous selection since the core technology was patented and published, and subsequently licensed broadly.

[1] I use the term "spillovers" somewhat loosely here, since many economists argue that a spillover exists only when the originating party is not compensated and/or there is no investment required on the part of the absorbing party.

E-mail address: ajnelson@uoregon.edu.

0048-7333/$ – see front matter
doi:10.1016/j.respol.2009.01.023

I present data on each diffusion measure over time in order to assess patterns of involvement on the organizational level along with errors of omission by individual measures. Notably, I find that direct patent citations dramatically understate the extent of technology diffusion compared to licenses and publications. But, the severity of biases in each measure depends on timing, firm age, and geography. To conclude, I discuss what the different results for each measure imply for our studies of firm innovation performance and knowledge diffusion, and I build a case for the simultaneous use of multiple indicators.

2. Patents as measures of innovation and knowledge diffusion

Patents are one of the most prevalent measures of innovation, and for good reason: they are easily accessible in electronic form; by definition, they are linked to inventiveness; they are classified by category and sub-categories; they identify individuals and organizations; and they contain a trace of what knowledge they build upon through the citation of prior art. This last feature, in particular, makes patents useful for tracing knowledge flows. Nevertheless, there are lingering questions as to whether patents accurately capture innovations and whether citations are good measures of knowledge flows.

The most straightforward use of patents involves a simple count of the number of patents that an organization produces, usually coded by time, industry and other characteristics. In an early study, Comanor and Scherer (1969) investigated how pharmaceutical patent counts in 1952–1957 were correlated with product introductions from 1955 to 1960, finding that patent counts appeared to be a predictor. More recently, Narin et al. (1987) found a high correlation between research inputs, patent counts, and patent citations. Similarly, Ahuja and Katila (2001) found a strong correlation (almost .9) between R&D inputs and patent counts.

Many researchers, however, question the use of patents to assess innovation since patents may reflect two types of errors. On one hand, patent data count only those inventions that inventors choose to patent and they therefore understate the full extent of inventive activity (Acs and Audretsch, 1989; Griliches, 1990; Hall et al., 2005; Scherer, 1983). Surveys by Levin et al. (1987) and Cohen et al. (2000) indicated that firms often employ alternative mechanisms such as secrecy to protect inventions. Arundel and Kabla's (1998) survey of large European firms indicated that errors of omission in patent data may be very large, with significant variation across industrial sectors. But, the full extent of these errors remains unknown. As Hall et al. (2005:18) noted, "Unfortunately, we have very little idea of the extent to which patents are representative of the wider universe of inventions, since there is no systematic data about inventions that are not patented."

On the other hand, patent data may overstate innovation since patents are tied to inventions and not to innovations. (I define inventions as new technological developments and innovations as those inventions that are economically useful and diffused.) Based on Meinhardt's (1946) data, Moser (2005) relayed that only 5–20% of patents become economically useful innovations.

For these reasons, other researchers have employed alternative measures altogether. For example, in a novel and painstaking data-gathering effort, Moser (2005) matched patent data to exhibition data from two 19th-century technology fairs in order to assess which inventions were patented. She found that patent laws influenced the direction of inventive activity, such that inventors in countries without patent laws focused on industries where secrecy was an effective means of protection. Moser's results suggest that the efficacy of patents as a measure will depend upon regulatory environments and the specific industry under consideration.

Another popular measure of innovation is a firm's research and development expenditures (e.g., Scherer, 1983). Cockburn and Griliches (1988:422) matched the survey data on patenting compiled by the Yale group (Levin et al., 1987) to National Bureau of Economic Research data on research and development. They found that patent count data are "subject to much error" and that data on R&D expenditures were better measures than patents of the inputs into firm innovation processes. In a later piece, Griliches (1990) argued that firm R&D expenditures capture a very early stage in the innovation process, while patents capture a later (albeit still early) stage in the process. But, the relationship between R&D expenditures and innovation is somewhat tenuous since R&D expenditures only capture input into the innovation process and not output. Moreover, R&D data alone provide little indication of the role of external sources of knowledge in firm innovation activities.

Hagedoorn and Cloodt (2003) compared four measures of innovation performance: patent counts, patent citations, R&D expenditures and new product announcements. While they pointed to some advantages of a composite construct made up of multiple measures, they also argued that the significant overlap between indicators means that any of the four measures can be taken as reflective of innovation in the broad sense. By contrast, in attempting to replicate Hagedoorn and Cloodt's results as closely as possible, Graham and Higgins (2007) found little support for the claim that patents are a valid indicator of innovative performance.

Beginning with Trajtenberg (1990), a number of researchers have employed citation-based patent measures, rather than strict counts. These citation-based measures assess the number and characteristics of downstream patents that reference a particular focal patent. Trajtenberg argued that citations provide a better indication of a patent's technological and economic value. Moreover, the use of citations addresses issues of knowledge flows, since an organization that cites a foregoing patent is presumed to have made use of knowledge in that foregoing patent. A primary concern with citation-based measures, however, centers on whether all subsequent innovations that build directly upon a patent actually contain a citation to that patent.

On one hand, there is reason to believe that patent citations would be quite accurate in their references to prior art. United States patent applicants have a "duty of candor" to disclose prior art on which an invention is based. The United States Patent and Trademark Office (USPTO) guidelines read: "This section should also contain a description of information known to you, including references to specific documents, which are related to your invention" (USPTO, 2007). Citations to prior art are typically citations to patents (and publications) that codify this existing knowledge base. In a comparison of international patents, Michel and Bettels (2001) found that US patent applications contain *many* more references than do applications to non-US patent offices—a result that Michel and Bettels attribute to the USPTO requirement to supply a complete list of the state of the art. Since a patent can be deemed unenforceable if the inventor or attorney/agent fails to properly disclose relevant references (Allison and Lemley, 1998; Sampat, 2005), Michel and Bettels (2001:192) claim that "rather than running the risk of filing an incomplete list of references, [applicants] tend to quote each and every reference even if it is only remotely related to what is to be patented."

On the other hand, Kesan (2002) argued that relevant prior art often is not disclosed. Both the level of proof required to show a willful failure to disclose and the associated expense of reexamination are high enough that patent applicants may face limited risk from a failure to disclose all prior art (see also Merges, 1999). Moreover, patent applicants may have disincentives to search for prior art. Since damages for "willful infringement" of a patent are significantly higher than damages for unintentional infringement, applicants may be better off if they are unaware of competitors' patents (Lemley and Tangri, 2003). To the extent that patents do not cite prior art upon which they build, citation measures will fail to fully capture the diffusion and impact of knowledge protected by an earlier patent.

The situation is confounded by the fact that many citations are added by patent examiners rather than by the inventor herself or her agent (Alcácer & Gittelman, 2006). Such citations, which may account for up to 40% of all citations in a patent, may indicate tenuous connections between a prior patent and the knowledge employed by a later inventor (Sampat, 2005). Only in the past few years has the USPTO made data available on which citations are added by examiners. Of course, since patent attorneys have long added citations in service of their inventor-clients, examiner-added citations may represent a simple extension of such third-party additions. In fact, they may "clean" citation data by better representing previous related knowledge upon which an invention builds.

We do have limited survey data on the extent to which patent citations actually represent knowledge flows. Jaffe et al. (2000) surveyed both citing and cited inventors around specific inventions and then matched the survey responses. They found that citations do indicate knowledge flows, but with a substantial amount of noise. Duguet and MacGarvie (2005) employed a survey about the innovative activities of French firms. They found that the correlation between patent citations and firms' statements about knowledge acquisition varied by geography and by channels of knowledge acquisition. For the French firms surveyed, patent citations were best at capturing knowledge sources from European firms and obtained via cooperative R&D, equipment purchases, and mergers and acquisitions. By contrast, patent citations were poor at capturing knowledge sourced from outside Europe and via other channels such as analyses of competing products, personnel hiring and exchanges, and communications with suppliers and customers.

Like the Jaffe et al. (2000) and Duguet and MacGarvie (2005) surveys, my interest lies in assessing the extent to which citation analyses capture knowledge flows. But rather than rely on survey data, my study compares various complementary indicators of knowledge diffusion – patents, licenses and publications – in order to assess the potential omissions and biases of each measure.

3. Technology licenses and publications as measures of knowledge diffusion

The comparison of technology licenses, publications and patents permits an assessment of the magnitude and direction of potential biases in citation analyses. Technology licenses are contractual agreements that grant organizations permission to use a particular piece of patent-protected knowledge held by another organization. Licenses are an appealing measure of innovation for two reasons. First, licenses often provide an indication of the economic value of those innovations that make use of the license. In order to receive a license, a licensee typically must pay (1) an upfront fee and/or (2) an annual fee and/or (3) a percentage of annual revenue on related products. Those licenses that include the third element – related-product revenue – are particularly informative since they permit a researcher to assess if and when a licensee has released a product based on the patented technology. Moreover, if the percentage figure is known, the revenue data permit a precise measurement of the total firm revenue for these same related products.

A second reason that licenses are an appealing measure is that the competing interests of licensor and licensees serve to enforce a system of "checks and balances" that results in an accurate capture of all relevant organizations. On one hand, the organization issuing a license for use of its patent(s) has a vested interest in ensuring that all of those organizations that are using the patented knowledge have in fact signed a license. This enforcement is important in order to maximize revenue by capturing the full population of users. Moreover, failure to strictly enforce a license requirement against any one organization would lead, in a game-theoretic sense, to a decreased likelihood that *all* organizations would sign the license; if organizations observe that the license requirement is not enforced, their incentive to comply is reduced.

As a result of this interest in strict enforcement, organizations in the technology-licensing business are rigorous in investigating the activities of a broad range of firms in order to determine if these firms' innovations are making use of a patent held by the licensing organization. In fact, interviews that I conducted with licensing professionals in fields ranging from the life sciences to digital audio indicate that out-licensing organizations employ dedicated personnel for the express purpose of ensuring that all organizations using their technology are, in fact, signed up as licensees.

On the other hand, potential licensees act as a "check" on the out-licensor's desire to capture as many organizations as possible. Licenses cost the licensee some amount of money in the form of an upfront fee and/or an annual fee and/or a share of related-product revenue. This cost is one that potential licensees are willing to pay only if they truly are making use of the patented technology of the licensor.[2] In short, then, strict investigation and enforcement on the part of the licensor balanced against the economic interests of the licensee create a system of "checks and balances" to counteract both the under- and over-representation of follow-on firms that patent citations alone may reflect.

It must be emphasized these in-licensing organizations are making use of the out-licensing organization's patent. Since the only way for the out-licensor to enforce a license requirement is via patent protection (and the threat of a patent infringement lawsuit against those that do not sign a license), licensing organizations are, by definition, those organizations that *intend* to make use of the patented technology covered by the license. Similarly, the subset of licensing organizations that pay a percentage of related-product revenue to the licensor are, by definition, those organizations that *have*, in fact, made use of the patented technology covered by the license.

Theoretically, organizations that pay a percentage of related-product revenue should also be amongst the organizations that cite back to the focal patent. Failure on the part of these product-releasing organizations to cite the patent can only be explained in two ways: either some organizations are using the technology in downstream applications that they have not patented, or some organizations are not citing back to the focal patent even though they are making use of the technology in products that they have patented.

This study also makes use of a third measure of knowledge diffusion: publications. Survey data on sources of knowledge emphasize the importance of publications (Cohen et al., 2002; Agrawal and Henderson, 2002). In fact, Branstetter and Ogura (2005) present evidence from the University of California that suggests that citations to publications (as opposed to citations to patents) provide a much broader view of knowledge spillovers from academic science. (For discussion of patent citations to publications, see also Hicks et al., 2001; Narin et al., 1997.) But, while publications may be an important *source* of information, downstream organizations must also publish (and cite) in order for downstream publications themselves to provide information on knowledge diffusion. In universities and other public science organizations, in which individual researchers

[2] More specifically, a firm will take out a license if the cost of the license does not exceed the cost of a patent infringement lawsuit times the probability of such a lawsuit. Thus, "non-using" firms may take out a license if the cost is relatively low and they fear that a lawsuit is likely, or "using" firms may fail to take out a license if the license cost is relatively high and they determine that a lawsuit is unlikely.

are presumed to subscribe to a reward system based upon prestige built through publication (Merton, 1973), publications may be common. With commercial firms, however, such downstream publication may be rare since publications, by definition, openly share knowledge with others. Firms may be unlikely to engage in such sharing since doing so would allow competitors to gain insight into a firm's inventive activities—potentially exploiting the results for their own benefit and to the detriment of the inventing firm (Dasgupta and David, 1994). Nevertheless, in industries such as biotechnology, we know that firms engage in significant publishing activity since the need to engage with inter-organizational knowledge sharing networks outweighs the danger of alerting competitors to their activity (Cockburn and Henderson, 1998; Powell et al., 1996). Moreover, publications can actually be part of a firm's patenting strategy. As Hicks (1995) pointed out, firms often patent and then immediately publish on the same material in order to preclude other firms from patenting in the same area. This strategy is one reason that patent-publication pairs are not uncommon, particularly in biotechnology (Murray, 2002; Murray and Stern, 2007).

As with patent citations, publication citations may indicate existing knowledge upon which the current publication builds (Cole, 2000). A long line of literature in the sociology of science has dissected this claim and led to a richer understanding of what publication citations represent. MacRoberts and MacRoberts (1986, 1989), for example, develop the useful concept of "influence." They write, "When an author makes use of another's work either directly or through secondary sources, and this is evident in the text, he has been influenced by that work" (MacRoberts and MacRoberts, 1989: 342). While some citation analyses are based upon the assumption that authors cite the work that influences them, MacRoberts and MacRoberts (1986) found that authors cite only about 30% of their formal influences. Moreover, Edge (1979) maintains that publication citations only capture the influence of other formal publications, ignoring informal communications and tacit knowledge that may be far more important as influences.

MacRoberts and MacRoberts (1989, 1996) point to a number of other concerns with publication citation analyses, including biases in citation patterns, motivations to cite that extend beyond intellectual influence (e.g., "political" motivations to cite), "negative" citations (in which a publication is cited as wrong or misleading), variation between specialties, and authors' ignorance of some relevant literature. Together, these concerns with citation patterns, along with the different organizational incentives for publishing in the first place, suggest that while publication citations contain much useful information, they are unlikely to fully capture knowledge flows. The challenge, therefore, is to assess the extent and characteristics of errors in each measure.

4. Data and methods

My interest in tracing knowledge diffusion led me to take an invention-centric focus and to investigate the spread of single newly invented technique across organizations and over time. To assess differences between various indicators of diffusion, I employ a novel dataset consisting of matched patents, licenses and publications surrounding a single invention: recombinant DNA (rDNA). Stan Cohen of Stanford University and Herb Boyer of the University of California at San Francisco (UCSF) developed rDNA technology in 1973. Cohen and Boyer published their findings in the *Proceedings of the National Academy of Science* in 1973 and Stanford, acting as the agent for itself and UCSF, applied for patent protection in 1974.

Cohen and Boyer's technique, commonly known as gene splicing, permits the insertion of foreign DNA into a host organism. Together with monoclonal antibodies and polymerase chain reaction, it is considered to be a foundational technology for the biotechnology industry and it has been used to create synthetic insulin, human growth hormone, and drugs for diabetes, anemia and blood clots. Its discovery also led to a Nobel Prize in Chemistry. Thus, while the study's design is limited to a single invention, this invention is nearly unmatched in impact and importance.

The USPTO ultimately issued three patents on Cohen and Boyer's discovery, covering both process and product for different organisms. These patents were issued between 1980 and 1982, though the 1980 patent, which also received the most subsequent citations, was strong enough for Stanford to enforce a license requirement starting in that year. (Stanford only required a license of firms, not nonprofit organizations; for complete details on the licensing arrangement, see Reimers, 1987; Feldman et al., 2005.) This 1980 patent expired in 1997. In the analyses that follow, I group the three Stanford rDNA patents together. To compile the patent data, I employed a custom-programmed data-scraping tool to download and parse information from the USPTO website on all issued patents through 2007 that referenced any of the three core rDNA patents, capturing a total of 346 patents. I refer to these 346 patents as "direct-citing" or "one-step" patents since they directly cite the core rDNA patents and are therefore only one step removed from them. I also collected data on all patents that referenced any of these 346 one-step patents, which yielded an additional 3547 patents. I refer to these 3547 patents as "two-step" patents since they are two citation steps removed from the core rDNA patents.

The Stanford University Office of Technology Licensing (OTL) managed for Stanford and UCSF all licensing activities related to the Cohen–Boyer rDNA discovery. This office was gracious to provide me with the full list of licensees for the technology, including the date upon which a license was signed, the dates during which it was active, and the associated fees and related-product revenues.

Given the importance of the Cohen–Boyer patent to both Stanford and UCSF, and the need to police users of the technology in order to ensure that all were paying their due, Stanford employed two full-time employees who were charged with determining whether an organization was using rDNA technology and, if so, with forcing them to sign a license if they had not done so already. Stanford's enforcement of this license requirement hinged on their ability to prosecute the Stanford-held rDNA patents; in other words, if an organization were not drawing upon the Stanford patents, they would have no reason to sign a license. Therefore, discrepancies between the in-licensing organizations and the organizations that cite these core patents have only two explanations: either some organizations used rDNA in downstream applications that they did not patent, or some organizations did not cite back to any of the core rDNA patents even though they were making use of the technology in their own patented inventions. Given the setting of this investigation in the field of biotechnology, I reject the first explanation. Firms that develop any products based in biotechnology are very likely to patent these products; in biotechnology more so than any other industry, patents (versus secrecy, for example) are the preferred method for protecting intellectual property (Cohen et al., 2000). Thus, I contend that differences between licensing organizations and patent-citing organizations are indicative of imperfections in the ability of patent citations to capture downstream users.

To compile publication information, I accessed the ISI SciSearch database. I downloaded all publications through 2007 that referenced the core 1973 rDNA publication, capturing a total of 756 publications.

For all four measures (direct-citing patents, two-step patents, licenses and publications), I then created a master list of all organizations that appeared in any of the datasets. Given misspellings and alternative labels for some organizations, this process required the manual review of each record to determine organizational matches.

(See Melamed et al., 2006, for a discussion of related issues for patents alone.) Over the 35-year time period covered, there were a number of mergers, acquisitions and name changes amongst the organizations in the master list. I was liberal in my matching criteria. For example, Schering-Plough acquired DNAX in 1982 and acquired Canji in 1996. Neither Schering-Plough nor Canji holds a patent that directly references back to the original rDNA patents. But, DNAX does hold such a patent. I therefore count Schering-Plough, Canji, and DNAX as having a match between a license and a direct patent citation. I was also liberal in counting all issued patents through 2007, though I truncate the patent trend lines in graphs at 2003 in order to account for the lag between patent application and issue.

To provide a sense of the demography of involved organizations, I coded each organization as a firm, a university or a non-university public research organization (PRO). In addition, I researched each organization to determine its date of founding and the location of its headquarters. While it would be desirable to capture additional firm characteristics such as R&D expenditures, the early time period (beginning in 1973) and the large number of non-public companies prohibit such an analysis.

In my analyses, I first assess the number and type of active organizations captured by each measure over time. I then compare each pair of indicators – patents versus licenses, patents versus publications, and licenses versus publications – to assess differences between the measures in the organizational activity and features that they capture and neglect.

5. Results

Table 1 conveys the demographic composition of organizations captured by each measure. Direct patent citations capture the fewest organizations while two-step patent citations capture the most organizations. Publications capture almost three times as many organizations as do direct-citing patents, though they capture only about two-thirds as many firms. Licenses capture far more firms than either direct patent citations or publications, though they miss other types of organizations since Stanford/UCSF did not require universities and other PROs to sign a license.

Fig. 1 illustrates patenting, publication and licensing trends for rDNA over time. It is clear from the graph that each measure provides a very different picture of activity around the technology. Whereas publication citations peak 4 years after the invention of the technique (and 4 years after the first publication), citations in patent applications peak 21 years after the original patent application (15 years after the first patent issued), with a second local peak 8 years after application, in 1982. Like direct patent citations, two-step patent citations also peak in 1995, though they emerge later and exhibit a second local peak in 2000. Licenses, by contrast, peak in 1980, the first year in which Stanford required them, with a second local peak in 1993. Thus, the four measures provide different indications of the intensity of activity at various points in time.

Figs. 2 and 3 break out, respectively, the one-step patent and publication measures in Fig. 1 according to the demography of the organizations involved. (Stanford only required firms to sign licenses and therefore the license data is attributable entirely to firms.) It is little surprise that firms are responsible for the vast majority of patents while universities and other public research organizations produce the majority of publications. Nevertheless, the extent of "cross-realm" activity (firm publications and university/PRO patents) is worth noting. In particular, the considerable number of firm publications provides an initial indication that publication measures are relevant towards capturing firm-level phenomena.

Table 1
Number of organizations of various types captured by each measure.

	Universities	Other PROs	Firms	Total
Direct patent citations	31	14	90	135
Two-step patent citations	102	73	633	807
Direct publication citations	201	113	55	369
Licenses	–	–	464	464

5.1. Organizations involved

While Figs. 1 through 3 count the total number of patents, publications and new licenses in a given year, another approach is to consider the number of different organizations that each measure captures in a given year, along with related trends over time. Fig. 4 displays, for each year, the number of organizations patenting, publishing, and maintaining an active license. The number of organizations active in patenting and publishing is highly correlated to the total count of patents and publications in a year; for patenting, the correlation coefficient is .93, while for publishing the correlation coefficient is .90. While patents and publications are measures of research output, holding an active license in a year (typically accomplished by paying a modest annual maintenance fee) is akin to engaging in the *pursuit* of an output—or, at least, maintaining the option of engaging in such a pursuit. For this reason, the measure of active licenses differs considerably from the number of new licenses signed in a year. While some companies dropped the license after a period of time, apparently abandoning their related rDNA research efforts, the *s*-curve of the active licensees indicates a continually expanding number of firms that were actively pursuing applications of the technique. Thus, licenses provide an indication of current and ongoing research activities by a firm, whereas patents and publications only show the single year that such activities result in that particular output.

Fig. 5 breaks out the one-step patenting activity in Fig. 4 according to organization type. Firms are responsible for the majority of patenting over time (66.7%). PROs account for 10.4% of patenting organizations and universities account for 23.0% of patenting organizations. Amongst two-step citations (not included in Fig. 5), the proportion of PROs remains about the same (8.9%). But the proportion of universities falls to 12.5%, with firms comprising the remaining 78.6% of organizations. Overall, the average organization with a one-step patent produces 2.56 patents that directly reference back to one of the core rDNA patents. By organization type, the average patenting firm produces 2.78 patents that directly cite one of the core rDNA patents; the average patenting PRO produces 2.36 such patents; and the average patenting university produces 1.65 such patents.

Fig. 6 breaks out the publishing activity in Fig. 4 according to organization type. Universities are responsible for the majority of publishing over time (54.5%). PROs comprise 30.6% of publishing organizations and firms comprise 14.9% of publishing organizations. Overall, the average organization that directly cites the core rDNA publication produces 2.05 such publications. By organization type, the average publishing firm produces 1.55 such publications; the average publishing PRO produces 1.73 such publications; and the average publishing university produces 2.70 such publications.

In sum, the comparisons within and between Figs. 1 through 6 indicate that patents, licenses and publications provide alternative images of diffusion that differ in timing and in organizational demography. At the same time, the measures are not orthogonal, meaning that they do not capture unrelated aspects of innovation performance. For example, even if interest is limited to a subgroup such as firms developing successful follow-on innovations, comparison of indicators reveals shortcomings in individual measures. I discuss direct comparisons between the measures in the sections that follow.

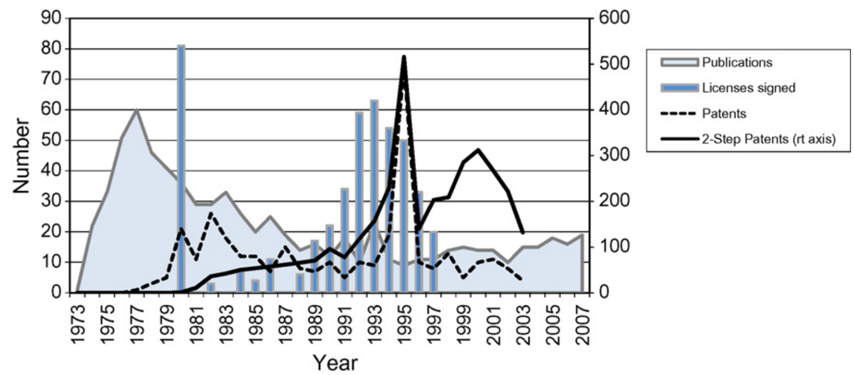

Fig. 1. Number of publications in year/number of patent applications in year/number of licenses initiated in year.

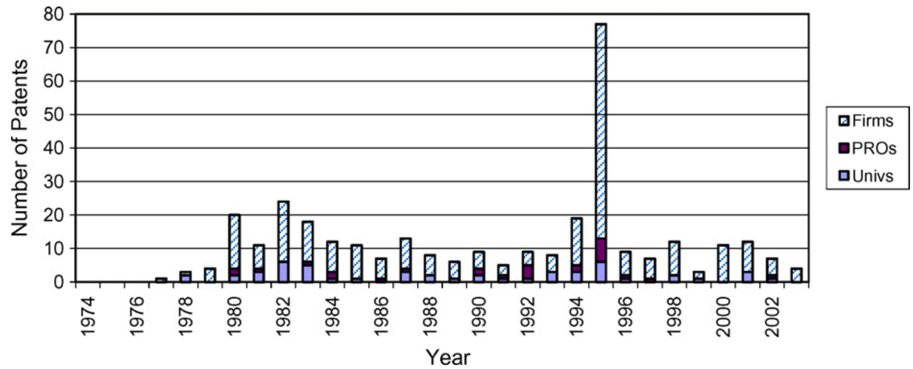

Fig. 2. Patents per year by type of organization listed as assignee on the patent (year of application for issued patents). The sum on this graph is greater than the total number of patents due to multiple assignees.

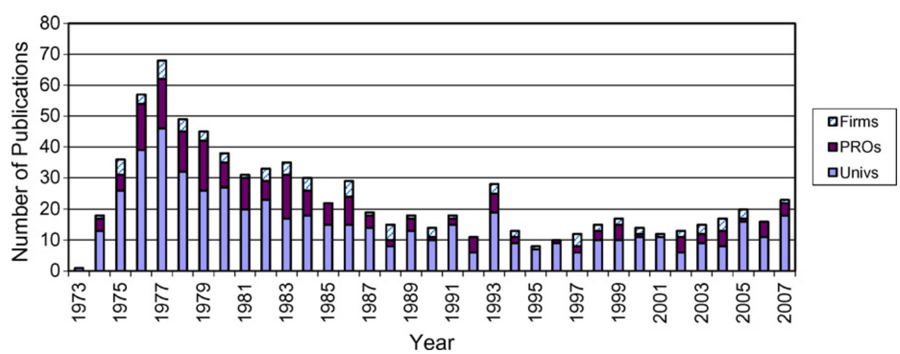

Fig. 3. Publications per year by type of organization listed as assignee on the publication.

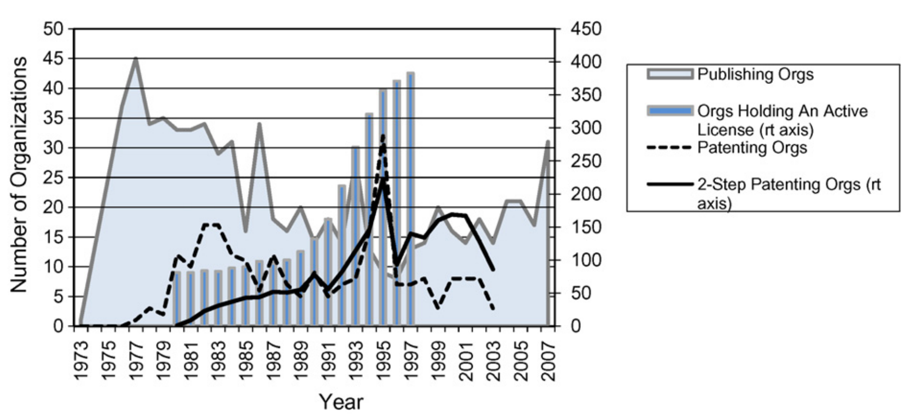

Fig. 4. Number of organizations patenting/publishing in a given year, and number of organizations maintaining active licenses in a given year.

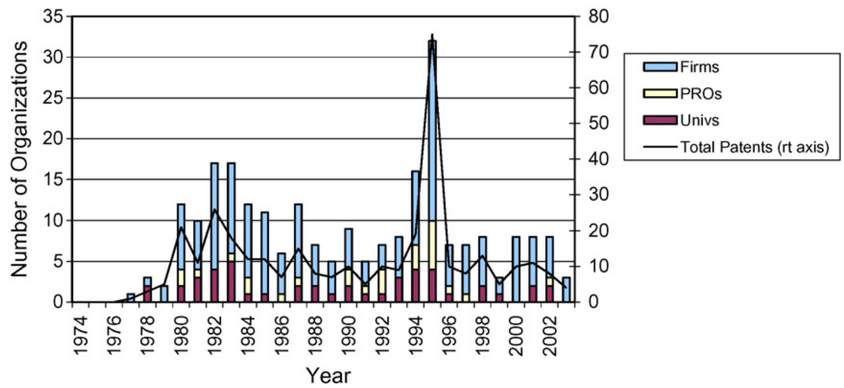

Fig. 5. Number of organizations patenting in a given year, by organization type.

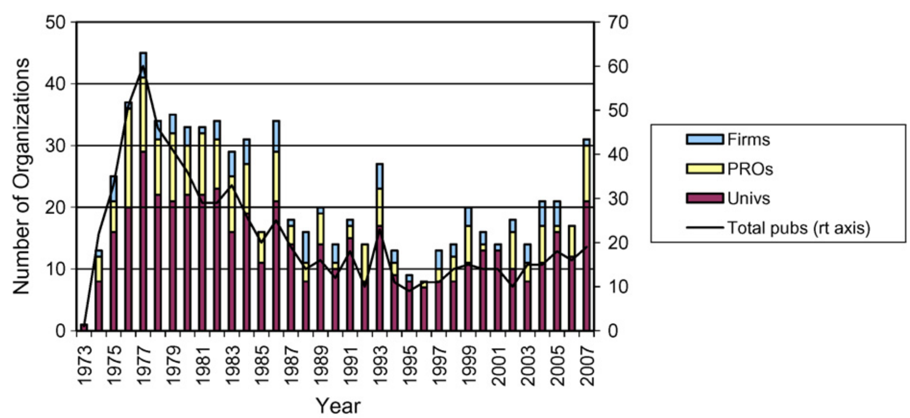

Fig. 6. Number of organizations publishing in a given year, by organization type.

5.2. Licenses versus patents

Direct patent citations miss the vast majority of the organizations that build upon the focal rDNA patents: 464 organizations signed licenses for rDNA while only 135 organizations hold patents that directly cite back to original rDNA patents (upon which the license is based).[3] The overlap between the two samples is 55 organizations. Therefore, patents fail to capture 409 organizations in the license sample, or 88.1% of the organizations that are building upon the patents (per the license data). This is an extreme error of omission.

An argument can be made, of course, that patent citation data reflect those organizations that actually have made a downstream invention of value, whereas licensing data only reflect those organizations that *attempted* or *intended* to make something. Fortunately, the licensing data facilitate a test of this proposition: since licensing contracts were structured to provide a share of revenue for related products, it is possible to identify which firms paid back a share of revenue to Stanford/UCSF, indicating that they did, in fact, release products. The data indicate that 117 of the 464 licensing organizations released revenue-generating products during the license period (1980 through 1997). Of these 117 organizations, only 21 have patents at any point time (1973 through 2007) that directly cite one of the three Cohen–Boyer patents. Thus, even if organizations of interest are limited to firms that actually released related products, direct patent citations still miss a large proportion (82.1%) of the relevant organizations. In these cases, direct patent citations are grossly under-representing related downstream innovations and the organizations associated with them.

The downstream innovations that patent citations miss are important. Of the top ten licensees in terms of revenue returned to Stanford, three of them (Johnson & Johnson, Novo-Nordisk, and Chiron) do not have patents that directly cite back to the core rDNA patents. (The revenue paid back to Stanford may be taken as a direct proxy for the sales of related drugs.) Collectively, these firms, which all signed licenses in 1980, paid over $26 million in revenues based on four blockbuster drugs (Feldman et al., 2005). Thus, the innovations that patent citations miss are not on the fringe of the field, but instead are amongst the most successful.

Since the one-step patent citation measure only picks up those organizations that directly cite one of the Stanford rDNA patents (that also form the basis of the license), it is informative to assess how much additional activity is captured via two-step patent citations. As discussed earlier, two-step citations pick up many more patents (3547 versus 346) and many more organizations (807 versus 135). Two-step patent citations also pick up all of the top 10 licensees (including the three missed by direct patent citations). Nevertheless, two-step patent citations still miss considerable activity, capturing only 195 of the 464 licensing organizations

[3] The Stanford Office of Technology Licensing lists 468 licensees and this figure is used in other studies of recombinant DNA licensing (Feldman et al., 2005). But, in four cases, the OTL uses two separate entries for the same company. Clorox and Eastman Kodak renegotiated their licenses, leading to two entries for each company to reflect the different periods in which specific licensing terms were in effect. In 1998, Beckman Instruments acquired Coulter Corporation and Beckman Instruments changed its name to Beckman Coulter. The company has two entries to reflect this name change even though it is Beckman that held the license in both time periods. Finally, there are two entries for Ciba-Geigy, as a result of its later merger to form Novartis and its purchase of Corning's share of an earlier joint venture, Ciba-Corning Diagnostics Corp. But, since Ciba-Corning and Novartis each have their own entries, it seems reasonable to only maintain one entry for Ciba-Geigy.

(42.0%) and only 54 of the 117 product-releasing organizations (46.2%).

Moreover, the use of two-step patent citations compounds errors in the other direction, capturing numerous organizations that did not commercially develop rDNA technology (according to the license data). For example, Cooligy Inc. holds seven two-step patents, all of which cite patent 6103199, which in turn cites a core rDNA patent. Patent 6103199, held by Aclara Biosciences, describes a device that uses microfluidic processes to dispense samples in the separation of biomolecules. This biomolecular application led to the inclusion of the rDNA patent citation. Cooligy Inc., however, designs and manufactures a product that cools computer chips using microfluidic channels. They, and their seven two-step patents, have no link to rDNA. In total, two-step patent citations pick up 377 firms through 1997, when the licensing requirement expired. But, only 54 of these 377 firms actually released rDNA-related products (per the licensing data), indicating that use of this measure may introduce significant noise into the population of downstream users.

Direct patent citations also capture many organizations that apparently did not commercialize rDNA. There are 55 organizations that directly cite back to the focal patent and hold a license. But, only 21 of these organizations released products. Thus, the remaining 34 organizations hold patents that never materialized into revenue-generating products, indicating a significant amount of patenting even when no related product is released. There are, of course, many reasons that firms may patent besides the protection of actual products released into the marketplace (Levin et al., 1987; Cohen et al., 2000). The present data indicate that in 61.8% of direct-citation cases (34 of 55), patents are not tied to related products.

The data also show that 80 organizations have patents with direct citations to the core rDNA patents but had no license, revealing potential imperfections in the licensing data. Only 35 of the 80 organizations were firms, however, so only 35 needed a license. (Stanford/UCSF did not require licenses of public research organizations.) Of these 35 organizations, 11 applied for their patents after the license requirement had expired in 1997. Still, that leaves 24 organizations that were never picked up by the Stanford OTL, but probably should have been, with a broad distribution over the time period. License data, it appears, are also an imperfect reflection of who has accessed a technology with commercial intentions. In this case, license data miss 4.9% of the relevant organizations (from a total population composed of organizations with a patent and/or a license).

What are the characteristics of the organizations that are missed by patent data but captured by licensing data? They are all firms, since Stanford only required for-profit organizations to take out a license. They are also over-represented in the later years, as indicated in Figs. 7 and 8. Fig. 7 tracks organizations by year of license initiation, displaying the number of new licenses signed each year

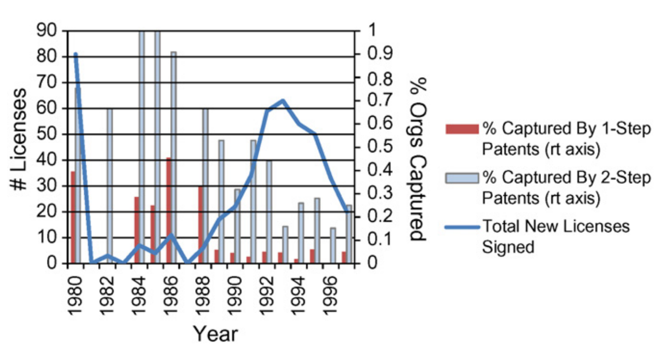

Fig. 7. Organizations signing up for a license in a given year.

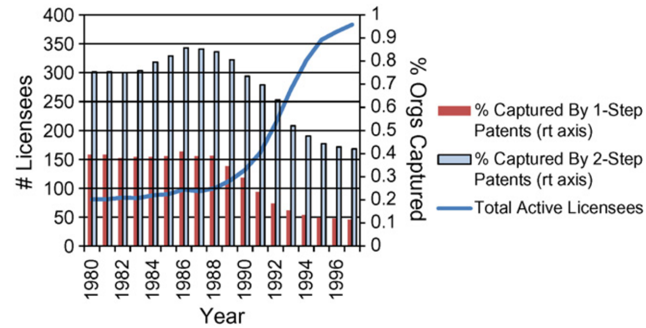

Fig. 8. Organizations actively holding a license in a given year.

and the percentage of these licenses captured via one-step and two-step patent citations. Fig. 8 employs the same schema, but tracks organizations that actively maintain a license in a given year. In all cases, the percentage of organizations captured tapers off in later years. Practically speaking, this finding means that studies focused on innovations that follow closely on the heels of their predecessors may capture a relatively representative sample by relying on patent data to indicate timing of diffusion. But, those studies that cover a longer time period for a technology will grossly understate innovation and the extent of diffusion in later years.

Geographically, there is little difference between licensing organizations captured by patent measures and those missed by patent measures. In both cases, the average distance from Stanford/UCSF is far (2217 miles for patent holders versus 1880 miles for license holders, and 2217 miles for licensees with a direct-citing patent versus 1836 miles for licensees without a direct-citing patent). These distances reflect the dominance of the distant Boston biotechnology community as well as the presence of international organizations in both samples. The differences in average distance are not significant. In terms of potential local biases, direct patent citations pick up only four of the 62 licensing organizations that are within 50 miles (6.5%), while they pick up 51 of the 402 licensing organizations that are more than 50 miles away (12.7%). These differences, however, are not significant.

In terms of organization age, patent data bias the sample towards older organizations. The average founding year of a firm in the patent data is 1944. The average founding year of a firm in the licensing data is 1966 ($p < .001$) Within the licensing data, the average founding year of a firm with both a patent and a license is 1935. The average founding year of an organization with a license but no direct patent is 1970 ($p < .001$).

These differences remain with two-step patents. Within the licensing data, the average founding year of a firm with both a license and a patent or two-step patent is 1953. The average founding year of an organization with a license but no patent or two-step patent is 1975 ($p < .001$).

The difference in founding years also remains if the population is limited to those organizations actually released products. The average founding year of an organization that both released a product (per the license data) and appears in the patent data is 1942. The average founding year of an organization that released a product (per the license data) but is not in the patent data is 1967 ($p < .01$). Together, these findings indicate that reliance on patent citations as a downstream measure of firm innovation biases a sample towards older firms.

Despite this bias, however, it must be noted that there is little evidence that patent data are biased against startups in particular. 179 of the 464 licensing organizations were startups at the time that they signed their license. (I define a startup as an organization that is less than 5 years old.) Direct patent citations capture 16 of these organizations, or 8.9%, and capture 11.9% of all licensing organizations (startups and established firms). These differences are not significant.

5.3. Patents versus publications

As Figs. 1 through 6 indicate, publications provide a very different picture of knowledge diffusion than do patents—one that emerges earlier and that highlights the involvement of public science organizations. Publications also indicate broader diffusion than one-step patents, both in terms of outputs – there are 756 publications versus 346 patents – and in terms of involved organizations: there are 369 organizations with citing publications and 135 organizations with direct-citing patents. (Two-step patent citations capture 807 organizations and 3547 patents.) The direct patent group is not a subset of the publication group, but instead represents a relatively unique sample. In fact, only 43 of the organizations hold both a citing publication and a direct-citing patent.

Since the patent and publication populations are relatively distinct and since there are a significant number of firms that publish, publications pick up 32 firms that direct patent citations miss (and 27 firms that two-step patent citations miss). Inclusion of these 32 firms would increase the overall firm population by 35.6% compared to the use of direct patent citations alone. Publications are also remarkably effective in picking up product-releasing firms. Direct patent citations pick up 90 firms, 21 (23.3%) of which released products (per the license data). Publication citations pick up 55 firms, 14 (25.5%) of which released products (per the license data). Thus, the two measures are comparable in terms of the percentage of product-releasing firms that they capture amongst the full set of firms captured.

Finally, publication citations add a significant number of organizations to the population captured via direct patent citations alone. Reliance on direct patent data misses 326 (70.7%) of the total organizations involved if the researcher employs both publications and patents, while reliance on two-step patents misses 270 (25.1%) of the total organizations involved if the researcher employs both publications and two-step patents. Particularly if one believes that public science organizations play a unique role in diffusion (e.g., Owen-Smith and Powell, 2004), the tendency of patent data alone to under represent such organizations may be problematic.

Another difference between patents and publications concerns timing. Patent data construct a much later diffusion curve than do publication data (see Fig. 1). In part, this delay may reflect a natural increase in patenting activity around the time of commercial launch, consistent with Basberg's (1982) findings around patenting in the Norwegian whaling industry. In part, the later timing also results from the fact that the rDNA patents issued much later than the publication appeared. A high degree of uncertainty surrounded life sciences patenting in the 1970s (as indicated in part by the Diamond versus Chakrabarty case that brought the issue before the US Supreme Court) and the first rDNA patent did not issue until 1980 (Hughes, 2001).

Historical context is seen to have a major influence on the timing of patent data throughout the sample. For example, the peak in 1995 (see Fig. 1) may be traced to the TRIPs Agreement, which brought United States patent law in line with that of other nations through the General Agreement on Tariffs and Trade (GATT). Article 33 of the TRIPs Agreement specified that the term of patent protection should be 20 years from the date of filing, whereas US patent law had previously granted protection for 17 years from the date of patent grant. The change in US patent law took effect on June 8, 1995. Patents filed before this date were granted a monopoly based on the longer of the two possible patent terms (17 years from the issue date or 20 years from the earliest filing date). Of the 75 direct-citing rDNA patents filed in 1995, 62 were filed between May 30 and June 7. Only three direct-citing patents were filed in the remainder of the calendar year after June 7. In discussing TRIPs, Branstetter and Ogura (2005) note that overall applications submitted to the USPTO more than doubled in May and June of 1995, and that the surge appears to have been driven by the desire to benefit from the two potential patent terms. Thus, policy shifts exert a major influence on trends indicated by patent data. More generally, this finding reinforces Moser's (2005) contention that regulatory environments shape the efficacy of patents as a measure. In the present case, it is not only broad national differences that matter, but also specific within-country legislative changes that can interfere with the validity of longitudinal data. Absent the 1995 spike, citations in patent applications would peak in 1982, 8 years after the initial rDNA application and 2 years after the first rDNA patent issued.

These patent data anomalies do not imply that publication data are superior to patent data. (Indeed, while publications capture more organizations overall, they miss 93 organizations that only hold a direct-citing patent and do not publish.) First, while publication data are less prone to policy shocks, they are highly susceptible to changes in journal populations and standards. Second, the publication process differs in important respects from the patent process. Patent applications impose application and attorney fees and are subject to review only by a single organization (the USPTO). Rejected papers, by contrast, may be submitted to alternative journals and still appear as publications. Therefore, issued patents may reflect, on average, higher quality additions to the knowledge base, depending on the patent review process and the specific journal publications to which they are compared. Similarly, since patents are intended to protect economic interests, they may serve as better innovation indicators than publications. Finally, it is critical to note that citations in patents and publications serve somewhat different purposes. While patent citations indicate important examples of prior art against which technical novelty is judged, publication citations indicate influences upon the published work. In fact, citing publications may not build on this cited work in any related technical fashion, as evidenced by the significant number of review article citations and social science citations to the rDNA technique.

As with the patent/license comparison, it is possible to compare the geographic distances and founding dates of organizations captured by each measure. It is not a surprise that publications capture older organizations than do patents given the dominance of universities in the publication sample and the relatively long history that most universities enjoy. An analysis limited only to firms, however, finds no statistically significant differences between the age of firms captured by publications and the age of firms captured by patents. Following Gittelman (2007), it might also be expected that publishing organizations would be farther from the point of invention than patenting organizations. Across all organizations, this expectation holds; the average distance from Stanford/UCSF is 3397 miles for publishers versus 2217 miles for patenters ($p < .0001$). When the analysis is limited to firms, however, the difference in distances converges and is no longer significant.

Finally, it is worth noting the discrepancy between patent citations to the core rDNA patents versus patent citations to the core rDNA publication. The core patents and core publication contain roughly the same information. Indeed, all three of the core rDNA patents cite the publication. Thus, one might expect downstream patents to be indifferent as to whether they cite the patent or the publication as the relevant prior art. But, only four patents (besides the original core rDNA patents) actually cite the core publication, reflecting a strong tendency for downstream patents to cite other patents as prior art, even where there is a matched publication. (Unfortunately, it is impossible to determine which citations to the core rDNA patents were added by examiners. Sampat, 2005, suggests that examiners are at a comparative disadvantage in adding citations to publications versus patents.)

5.4. Licenses versus publications

Both publications and licenses provide a broad image of diffusion with 369 and 464 involved organizations respectively. Since licenses are given only to firms, however, they miss the considerable involvement of non-profit organizations in knowledge diffusion. Again, to the extent that such organizations may represent crucial nodes in a diffusion network (Owen-Smith and Powell, 2004), their omission may be critical. Moreover, licenses also miss some firms. Publications pick up 55 firms and 38 of these 55 appear amongst the 464 licensees. Thus, there are 17 firms that have published and do not appear in the license database. Those 17 firms would add 3.5% more to a sample of firms pulled only from license data. On the other hand, since publications are dominated by universities and other public science organizations, as indicated in Fig. 6, their use in isolation also provides a limited sample that is skewed towards particular types of organizations.

Publications and licenses also differ in their indications of the timing of diffusion. As evidenced in Fig. 4, licenses only reflect activity during the license period, which is bounded by patent issue on one end and by patent expiration on the other end. Thus, reliance on license data provides a skewed image of diffusion timing that is shaped as much by intellectual property considerations as it is by actual knowledge-access patterns.

Finally, as with the other comparisons, it is possible to assess the geography and age of firms captured by licensing versus publication data. Consistent with Gittelman's (2007) contention that market-oriented activities are more geographically clustered than are general knowledge-building activities reflected in publications, the average distance from Stanford/UCSF of a licensing firm is 1880 miles while the average distance for a publishing firm is 2832 miles ($p < .0001$). Moreover, publication data pick up much older firms than do licensing data. The average founding date of a firm in the publication dataset is 1926, while the average founding date of a firm in the full licensing dataset is 1965 ($p < .001$). But, there is no evidence that startups differ from established firms in their likelihood of holding publications that cite the core rDNA publication. Of the 179 startups in the licensing population, 11 (6.7%) hold such publications. Of the 285 non-startups, 26 (9.1%) hold such publications. The differences are not statistically significant.

6. Discussion

In this paper, I have attempted to assess the efficacy of various measures of knowledge diffusion by comparing rDNA patent citations, publications citations, and licensing data. The results show dramatic differences in magnitude and timing for the different indicators. Fig. 9 compares the various measures in terms of organizations captured, firms captured, and firms with products captured by each measure. To summarize, direct patent citations – the most common measure of knowledge spillovers – appear to be the most restrictive measure of diffusion, capturing the smallest number of organizations. Judged by the licensing data, direct patent citations miss 88% of all organizations building upon a core patent and 82% of those organizations that release a product based on a core patent. Moreover, in 62% of cases, one-step patents are not tied to actual released products. Thus, direct patent citations appear to be a "noisy" measure, capturing a significant number of firms who never release directly related products.

Two-step patent citations capture the greatest number of organizations in general and the greatest number of firms. But, they still miss the majority of licensing organizations and they are far less effective than licenses in capturing firms that release products. As a result, two-step patent citations also are a very noisy measure that captures numerous organizations with questionable linkages to rDNA. Finally, with both direct and two-step patents, changes in policy through the TRIPs Agreement had a dramatic effect on the overall level of patenting, thereby interfering with longitudinal analyses.

Licenses provide the only reliable assessment of actual products released and they capture more organizations than either publication citations or direct patent citations. Moreover, the incentives on the part of licensor and licensees ensure that licenses are a highly accurate representation of downstream technology users. Only firms signed licenses, however, meaning that this measure does not capture any of the significant university/PRO activity. Moreover, it does not capture activity outside of the license period.

Finally, publication citations are most effective in picking up universities/PROs and they add a significant number of firms missed by direct patent citations. But, they undercount firms overall and their connection to economically useful innovations is tenuous.

Practically speaking, the obvious points for innovation research are that different measures provide different patterns and that, ceteris paribus, more measures are better since they capture a greater amount of activity. As Fig. 9 indicates, combinations of indicators are almost always superior to any single indicator. Amongst the various combinations, however, it is worth noting that the simultaneous use of patents and publications – an increasingly common measure – is the least effective combination.

The analyses also revealed systematic biases between measures. Publication citations portray a world of innovation diffusion that emerges earlier, features older organizations, and is more geographically dispersed than that of patent citations. In turn, patent citations paint a picture of innovation diffusion that is smaller, poorer, older, and emerges earlier than that of licenses. But, it is important to note that the results do not appear to be biased against startups in particular, and instead reflect more general features of the population of organizations.

One obvious question resulting from these analyses concerns why direct patent citations miss so many organizations and so much activity. The answer lies in legal function of patent citations, which has on occasion been misinterpreted by innovation scholars. Organizations are required to cite techniques such as rDNA only when the technique is "material to the patentability" of present patent, meaning that citation affects the novelty or non-obviousness of the technique. Thus, patent citations cannot be expected to capture all downstream uses or users of a technology, particularly as that technology becomes well-known and widely diffused. In other words, while the licensing data highlight shortcomings in the ability of patent citations to capture downstream applications of rDNA, the data do not necessarily imply legal failings on the part of patent applicants. Instead, they point to the importance of understanding the role and legal context of patents (Gittelman, 2008).

The growing discrepancy between measures over time also reflects Merton's (1968) idea of "obliteration by incorporation." As Merton explains, scientists primarily make use of more recent contributions, which have developed earlier ideas. As a result, "earlier and often much weightier scientific contributions tend to be obliterated. . .by incorporation into later work" (Merton, 1968: 28). Citing Kessler, Merton notes that a serious student of physics can safely ignore the original foundational work of Newton, Faraday and Maxwell. Moreover, he argues that this dynamic is especially active in science (as opposed to the humanities). Undoubtedly, much of the fading tail on patent and publication citations to rDNA has to do with the inability of citations to reflect use, rather than a lack of use or diffusion.

A core finding of the present study is that different measures capture different aspects of diffusion, suggesting that diffusion itself is a concept in need of specification and elaboration. Depending upon the study, diffusion can refer to knowledge of the existence of a technique, to application of the technique in scientific experi-

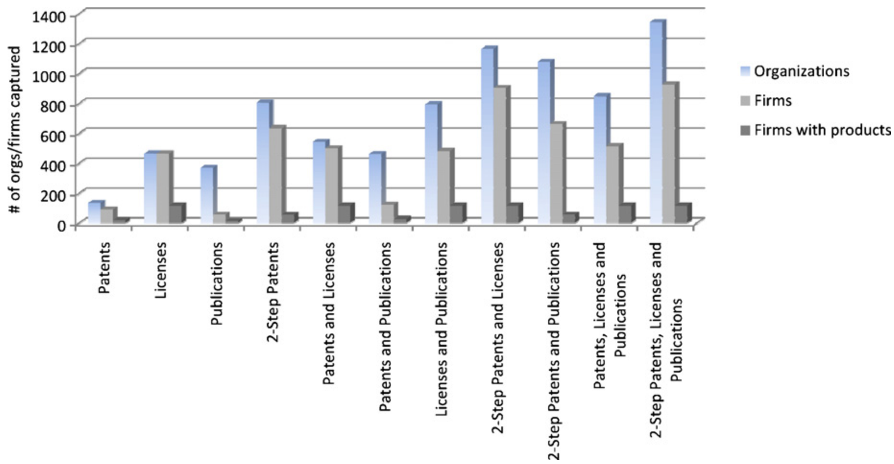

Fig. 9. Number of unique organizations/firms/firms-with-products captured by each measure.

ments or in product development, or to the sale of products based upon the technology. (Bozeman, 2000, offers a useful discussion of the various definitions of technology diffusion.) It is clear that certain measures are more effective at measuring certain aspects. For example, licenses appear to be the most reliable measure of product sales. But, it is also doubtful that any single indicator is an adequate measure of even one specific perspective on diffusion.

The limited scope of this study demands, of course, that these conclusions be taken in context. In particular, the life-sciences setting of this study and the study's focus on a single core technology must be acknowledged as limitations. As such, it is important to question the generalizability of this analysis. On one hand, the core invention studied garnered a Nobel Prize and is nearly unmatched in importance since it laid the foundation for an entire industry. Moreover, previous investigations indicate that patent protection is more important in this industry than in any other (Cohen et al., 2000) and that applicants insert a higher number of citations in biomedical (and chemical) patents than in other technology fields (Sampat, 2005). These findings imply that the present study is a liberal test of patent efficacy, and that patent citation data presumably would miss even more activity in other technological realms.

On the other hand, the desire to obtain a long time series led to selection of a technology that emerged when norms around intellectual property and universities were in flux (Colyvas, 2007). Moreover, the biotechnology industry is unique in the emphasis that firms place on publication. In other industries, the inclusion of publication citations might not yield as large a boost in the population of downstream users. Finally, the normative commentary on the use of licensing data must be tempered with the acknowledgement that such data can be difficult to obtain. In fact, in some industries, licensing data may be lacking simply because licensing is not a common practice amongst organizations. These limitations signal the need for greater cross-industry analyses of markers and practices of knowledge diffusion.

The intention of this study has not been to dismiss any particular measure used to study innovation diffusion. Rather, the results suggest that the ease of collecting and analyzing certain types of data should not blind us to alternative measures, nor assuage our caution at interpreting results, nor guide the ways in which we choose to study innovation. For example, in the field of university technology transfer, researchers, policy analysts, and practitioners often emphasize university-assigned patents as a key measure of successful technology transfer. By contrast, the present analysis indicates that patents fail to capture most formal technology transfer activities—much less the critical informal activities. Innovation is difficult to measure and specific indicators can provide important insights. But, in our rush to employ them, we should neither be blinded to their biases nor lose sight of the critical perspectives that alternative measures provide.

Acknowledgements

I am thankful for the advice and superb suggestions of Woody Powell, four anonymous reviewers, Mike Russo, Riitta Katila, Geerte Hesen, and members and reviewers with the Technology and Innovation Management section at the Academy of Management 2008 Annual Meeting. Thanks are also due to the Stanford Office of Technology Licensing for graciously providing licensing data.

References

Acs, Z.J., Audretsch, D.B., 1989. Patents as a measure of innovative activity. Kyklos 42 (2), 171–180.

Acs, Z.J., Audretsch, D.B., Feldman, M.P., 1992. Real effects of academic research: comment. American Economic Review 82 (1), 363–367.

Acs, Z.J., Audretsch, D.B., Feldman, M.P., 1994. R&D spillovers and recipient firm size. Review of Economics and Statistics 76 (2), 336–340.

Agrawal, A., Henderson, R., 2002. Putting patents in context: exploring knowledge transfer from MIT. Management Science 48 (1), 44–60.

Ahuja, G., Katila, R., 2001. Technological acquisitions and the innovation performance of acquiring firms: a longitudinal study. Strategic Management Journal 22 (3), 197–220.

Alcácer, J., Gittelman, M., 2006. Patent citations as a measure of knowledge flows: the influence of examiner citations. The Review of Economics and Statistics 88 (4), 774–779.

Allison, J.R., Lemley, M.A., 1998. Empirical evidence on the validity of litigated patents. American Intellectual Property Law Association Quarterly Journal 26, 185–269.

Arundel, A., Kabla, I., 1998. What percentage of innovations are patented? Empirical estimates for European firms. Research Policy 27 (1), 127–141.

Audretsch, D.B., Feldman, M.P., 1996. R&D spillovers and the geography of innovation and production. The American Economic Review 86 (3), 630–640.

Basberg, B.L., 1982. Technological change in the Norwegian whaling industry. Research Policy 11 (3), 163–171.

Bozeman, B., 2000. Technology transfer and public policy: a review of research and theory. Research Policy 29 (4–5), 627–655.

Branstetter, L., Ogura, Y., 2005. Is Academic Science Driving a Surge in Industrial Innovation? Evidence From Patent Citations. Working paper 11561, NBER.

Breschi, S., Lissoni, F., 2001. Knowledge spillovers and local innovation systems: a critical survey. Industrial and Corporate Change 10 (4), 975–1005.

Brouwer, E., Kleinknecht, A., 1999. Innovative output, and a firm's propensity to patent. An exploration of CIS micro data. Research Policy 28 (6), 615–624.

Cockburn, I., Griliches, Z., 1988. Industry effects and appropriability measures in the stock market's valuation of R&D and patents. The American Economic Review 78 (2), 419–423.

Cockburn, I., Henderson, R., 1998. Absorptive capacity, coauthoring behavior, and the organization of research in drug discovery. The Journal of Industrial Economics 46 (2), 157–182.

Cohen, S., Chang, A.C.Y., Boyer, H.W., Helling, R.B., 1973. Construction of biologically functional bacterial plasmids *in vitro*. Proceedings of the National Academy of Sciences 70 (11), 3240–3244.

Cohen, W.M., Nelson, R.R., Walsh, J.P., 2000. Protecting Their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (or not). Working paper 7552, NBER.

Cohen, W.M., Nelson, R.R., Walsh, J.P., 2002. Links and impacts: the influence of public research on industrial R&D. Management Science 48 (1), 1–23.

Cole, J.R., 2000. A short history of the use of citations as a measure of the impact of scientific and scholarly work. In: Cronin, B., Atkins, H.B. (Eds.), The Web of Knowledge: A Festschrift in Honor of Eugene Garfield. Information Today Inc., Medford, NJ, pp. 281–300.

Colyvas, J., 2007. From Divergent Meanings to Common Practices: the early Institutionalization of Technology transfer in the life science at Stanford University. Research Policy 36 (4), 456–476.

Comanor, W., Scherer, F.M., 1969. Patent statistics as measures of technical change. Journal of Political Economy 77 (3), 392–398.

Dasgupta, P., David, P., 1994. Toward a new economics of science. Research Policy 23 (5), 487–522.

Duguet, E., MacGarvie, M., 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.

Edge, D., 1979. Quantitative measures of communication in science: a critical review. History of Science 17 (36), 102–134.

Feldman, M., Kelley, M.R., 2006. The ex ante assessment of knowledge spillovers: Government R&D policy, economic incentives and private firm behavior. Research Policy 35 (10), 1509–1521.

Feldman, M., Colaianni, A., Liu, K., 2005. Commercializing Cohen–Boyer 1980–1997. Working Paper 05-21, DRUID.

Fritsch, M., Franke, G., 2004. Innovation, regional knowledge spillovers and R&D cooperation. Research Policy 33 (2), 245–255.

Gittelman, M., 2007. Does geography matter for science-based firms? Epistemic communities and the geography of research and patenting in biotechnology. Organization Science 18 (4), 724–741.

Gittelman, M., 2008. A note on the value of patents as indicators of innovation: implications for management research. Academy of Management Perspectives 22 (3), 21–27.

Graham, S.J.H., Higgins, M.J., 2007. The Impact of Patenting on New Product Introductions in the Pharmaceutical Industry. MPRA Paper 4574, Munich Personal RePEc Archive.

Griliches, Z., 1990. Patent statistics as economic indicators: a survey. Journal of Economic Literature 28 (4), 1661–1707.

Griliches, Z., 1992. The search for R&D spillovers. The Scandinavian Journal of Economics 94 (Supplement), S29–S47.

Hagedoorn, J., Cloodt, M., 2003. Measuring innovation performance: is there an advantage in using multiple indicators? Research Policy 32 (8), 1365–1379.

Hall, B.H., Jaffe, A., Trajtenberg, M., 2005. Market value and patent citations. RAND Journal of Economics 36 (1), 16–38.

Hicks, D., 1995. Published papers, tacit competencies and the corporate management of the public/private character of knowledge. Industrial and Corporate Change 4 (2), 401–424.

Hicks, D., Breitzman, T., Olivastro, D., Hamilton, K., 2001. The changing composition of innovative activity in the US—a portrait based on patent analysis. Research Policy 30 (4), 681–703.

Hughes, S.S., 2001. Making dollars out of DNA: the first major patent in biotechnology and the commercialization of molecular biology, 1974–1980. ISIS 92 (3), 541–575.

Jaffe, A., 1989. Real effects of academic research. American Economic Review 79 (5), 957–970.

Jaffe, A., Trajtenberg, M., Henderson, R., 1993. Geographic localization of knowledge spillovers as evidenced by patent citations. The Quarterly Journal of Economics 108 (3), 577–598.

Jaffe, A., Trajtenberg, M., Fogarty, M., 2000. The Meaning of Patent Citations: Report of the NBER/Case-Western Reserve Survey of Patentees. Working Paper 7631, NBER.

Kaiser, U., 2002. Measuring knowledge spillovers in manufacturing and services: an empirical assessment of alternative approaches. Research Policy 31 (1), 125–144.

Kesan, J.P., 2002. Carrots and sticks to create a better patent system. Berkeley Technology Law Journal 17 (2), 138–173.

Krugman, P., 1991a. Increasing returns and economic geography. Journal of Political Economy 99 (3), 483–499.

Krugman, P., 1991b. Geography and Trade. MIT Press, Cambridge, MA.

Lemley, M., Tangri, R.K., 2003. Ending Patent Law's Willfulness Game. Working paper 142, UC Berkeley Public Law and Legal Theory.

Levin, R., Klevorick, A., Nelson, R., Winter, S., 1987. Appropriating the returns from industrial research and development. Brookings Papers on Economic Activity 3, 783–820.

MacRoberts, M., MacRoberts, B., 1986. Quantitative measures of communication in science: a study of the formal level. Social Studies of Science 16 (1), 151–172.

MacRoberts, M., MacRoberts, B., 1989. Problems of citation analysis: a critical review. Journal of the American Society for Information Science 40 (5), 342–349.

MacRoberts, M., MacRoberts, B., 1996. Problems of citation analysis. Scientometrics 36 (3), 435–444.

Meinhardt, P., 1946. Inventions, Patents and Monopoly. Stevens and Sons, London.

Melamed, R., Shiff, G., Trajtenberg, M., 2006. The 'Names Game': Harnessing Inventors Patent Data for Economic Research. Discussion Paper 5833, CEPR.

Merges, R.P., 1999. As many as six impossible patents before breakfast: property rights for business concepts and patent system reform. Berkeley Technology Law Journal 14, 577–615.

Merton, R., 1968. Social Theory and Social Practice, 1968 Enlarged Edition. The Free Press, New York.

Merton, R., 1973. The Sociology of Science. The University of Chicago Press, Chicago.

Michel, J., Bettels, B., 2001. Patent citation analysis: a closer look at the basic input data from patent search reports. Scientometrics 51 (1), 185–201.

Moser, P., 2005. How do patent laws influence innovation? Evidence from nineteenth-century world's fairs. The American Economic Review 95 (4), 1214–1236.

Murray, F., 2002. Innovation as co-evolution of scientific and technological networks: exploring tissue engineering. Research Policy 31 (8), 1389–1403.

Murray, F., Stern, S., 2007. Do formal intellectual property rights hinder the free flow of scientific knowledge? An empirical test of the anti-commons hypothesis. Journal of Economic Behavior and Organization 63 (4), 648–687.

Narin, F., Noma, E., Perry, R., 1987. Patents as indicators of corporate technological strength. Research Policy 16 (2), 143–155.

Narin, F., Hamilton, K., Olivastro, D., 1997. The increasing linkage between U.S. technology and public science. Research Policy 26 (3), 317–330.

Owen-Smith, J., Powell, W.W., 2004. Knowledge networks as channels and conduits: the effects of spillovers in the Boston biotechnology community. Organization Science 15 (1), 5–21.

Powell, W.W., Koput, K., Smith-Doerr, L., 1996. Interorganizational collaboration and the locus of innovation: networks of learning in biotechnology. Administrative Science Quarterly 41 (1), 116–145.

Reimers, N., 1987. Tiger by the tail. Chemtech 17 (8), 464–471.

Romer, P., 1986. Increasing returns and long-run growth. Journal of Political Economy 94 (5), 1002–1037.

Sampat, B., 2005. Examining Patent Examination: An Analysis of Examiner and Applicant Generated Prior Art. Working paper, Columbia University.

Scherer, F.M., 1983. The propensity to patent. International Journal of Industrial Organization 1 (1), 107–128.

Thompson, P., Fox-Kean, M., 2005. Patent citations and the geography of knowledge spillovers: a reassessment. American Economic Review 95 (1), 450–460.

Trajtenberg, M., 1990. A penny for your quotes: patent citations and the value of innovations. RAND Journal of Economics 21 (1), 172–187.

United States Patent and Trademark Office, 2007. A Guide to Filing A Utility Patent Application. US GPO, Washington DC.