Knowledge Bridging by Biotechnology Start-ups

by

David H. Hsu and Kwanghui Lim*

This draft: June 2005 (original draft: August 2004)

ABSTRACT

We empirically investigate a type of technological boundary spanning-oriented search in which firms apply knowledge from one technical domain to innovate in another, a phenomenon we term knowledge bridging. In an analysis of a novel dataset of all the biotechnology firms founded to commercialize recombinant DNA technology, we examine the consequences of knowledge bridging at the patent-firm level of analysis, as well as the determinants of this search behavior at the firm-year level of analysis. The empirical setting allows us to establish a common technological starting point and examine knowledge bridging-oriented technological search behavior over time. At the patent-firm level of analysis, we find that knowledge bridging is significantly associated with measures of patent value. In an analysis of antecedents to knowledge bridging search at the firm-year level of analysis, we find that such search behavior is correlated with firms' initial search direction and receipt of venture capital funding.

Keywords: knowledge exploration, technological boundary spanning, innovation, entrepreneurship, biotechnology, patents.

JEL: O3, M13.

---

* The Wharton School, 2000 Steinberg Hall-Dietrich Hall, University of Pennsylvania, Philadelphia, PA 19104, dhsu@wharton.upenn.edu and NUS Business School, National University of Singapore, BIZ1-04-27, Singapore 117592, Republic of Singapore, k@kwanghui.com. We thank Kathy Ku of the Stanford University Office of Technology Licensing for allowing us to access the Cohen Boyer patent records. We thank Dan Levinthal, Scott Shane, and Scott Stern for helpful suggestions. We also thank conference or seminar participants at a 2004 Academy of Management session, the 2005 Swiss Federal Institute Innovation conference, the 2005 UCLA/Oxford Entrepreneurship conference, Harvard Business School, University of Maryland, and Wharton for useful comments and discussions. Josh Lerner generously provided access to his biotechnology index. Tong Zhao, Zhuang Wenyue and Mike Gonsalves provided valuable research assistance. We acknowledge use of the NUS patent database. We thank the Mack Center for Technological Innovation at Wharton and the National University of Singapore Business School for funding this project.

I. Introduction

The notion that organizations engage in local search is well established in the management literature. Firms have a propensity to explore knowledge that is familiar and within easy reach from their existing geographic and technological positions. The underpinnings for this behavior have been explored at multiple levels of analysis, ranging from individual-level explanations of bounded rationality (March and Simon, 1958) to firm-level capabilities, routines, and learning myopia (Nelson and Winter, 1982; Levinthal and March, 1993). New venture "imprinting" by founders (Stinchcombe, 1965) and the long-lasting impact of firms' initial conditions (e.g., Baron, Burton and Hannan, 1996; Cockburn, Henderson and Stern, 2000) also suggest mechanisms by which firms' search behavior is perpetuated.

In environments in which innovation is important as the basis for competition, firms and their managers may be particularly concerned about the long term competitive effects of local search (March, 1990). Not surprisingly, then, there has been considerable interest in mechanisms associated with overcoming the constraints of local search. For example, Mowery, Oxley, and Silverman (1996) examine strategic alliances as a mechanism for overcoming local search. The mobility of engineers and scientists in the labor market is another means by which firms can overcome local search (Almeida and Kogut, 1999; Rosenkopf and Almeida, 2003).

In this paper, we assess the empirical importance of a third potential mechanism for overcoming local search: using ideas from one technical domain to create innovation in another area. While this can take place either through "porting" ideas across application areas or through recombining knowledge from different arenas, we use the term "knowledge bridging" to describe the phenomenon. We ask two related research questions: (1) Is knowledge bridging behavior at the patent level related to innovation outcomes? and (2) what are the determinants of knowledge bridging at the firm level? We therefore analyze both the consequences of knowledge bridging behavior and its antecedents.

Economists, historians of technical change, and management scholars have long been interested in the notion that innovation results from recombination of existing knowledge (e.g., Schumpeter, 1934; Basalla, 1988; Hargadon and Sutton, 1997; Fleming, 2001). However, this literature has not empirically linked knowledge bridging to firm-level innovation and commercialization efforts. Our paper aims to address this shortfall by conceptually identifying knowledge bridging as a distinct form of technical boundary spanning and empirically studying its importance at both the firm-invention and firm-year levels of analysis. Empirically studying knowledge bridging requires constructing novel measures of the concept that will allow for across-patent and across-firm comparisons. We construct these measures based on the premise that the extent to which patents and firms cite patents in different technical areas indicates the degree of knowledge bridging from those areas (and so our measures compare technical areas of innovation in relation to technical areas cited).

To investigate knowledge bridging-based technological search, we carefully chose an empirical setting in which firms were founded to exploit a given technological innovation. This design allows us to track firms' temporal patterns of knowledge bridging from their inception, while holding initial technology constant. We can then study the effect of knowledge bridging on innovative outcomes (at the patent-firm level of analysis), as well as the factors that relate to knowledge bridging use (at the firm-year level of analysis) over time. The empirical strategy is therefore similar to that used by Shane (2000) in his study of entrepreneurial opportunities arising from a common three-dimensional printing technology. As well, our study is similar in spirit to Cockburn, Henderson and Stern (2000) in that we look for residual effects net of initial (founding) conditions.

The historical circumstances surrounding the commercialization of recombinant DNA technology via open, non-exclusive licensing of the Cohen-Boyer patent by Stanford University is a fortuitous one from a research standpoint. We are able to examine all de novo start-ups founded to commercialize this technology due to generous access to the Stanford licensing program records by the Stanford Office of Technology Licensing. The recombinant DNA innovation allowed DNA from two or more sources to be combined into a single source. The commercialization of this innovation is also of considerable interest since it launched the modern biotechnology industry (Kenney, 1986).[1]

The empirical results are presented in two stages. In the first stage, we find that patents which bridge knowledge domains have more impact than those that do not, as measured by forward citation counts. Knowledge bridging patents are also more general-purpose in nature, as they are cited by patents from a broader range of technological areas. In the second stage of the analysis, we document the substantial variation in firms' use of knowledge bridging in their business development. We then investigate covariates of the degree to which firms change their technological search patterns, taking as the unit of analysis a firm-year. We find that a firm's technological position at the time of founding and the receipt of venture capital funding significantly affect firms' use of knowledge bridging-based exploratory search.

In the next section, we review the relevant literature and derive empirical predictions. Section III discusses the data and method employed, while section IV presents the empirical results. A final section concludes and discusses future directions.

II. Literature and Hypothesis Development

[1] The biotechnology industry is quite technologically dynamic, and thus represents an interesting empirical setting in its own right. As of 2003, biotechnology innovations accounted for 155 U.S. Federal Drug Administration (FDA) approved drugs, with over 370 biotechnology clinical trials and vaccines in development (BIO website, accessed May 24, 2004). Furthermore, biotechnology firms are a significant source of upstream innovation for pharmaceutical firms (Gans, Hsu and Stern, 2002): of the 691 new chemical entities approved by the FDA between 1963 and 1999, 38 percent were licensed by pharmaceutical firms, primarily from biotechnology firms (DiMasi, 2000).

The creative application of knowledge across technical domains has been researched not only across disciplinary lines, but also across levels of analysis, including the invention or idea level (e.g., Fleming, 2001), the product level (e.g., Katila and Ahuja, 2002), and the firm level (e.g., Kogut and Zander, 1992). While we are primarily interested in investigating firm-level knowledge bridging-oriented search, we recognize that the concepts we discuss are sufficiently general that they can also take place across other levels of analysis. For tractability, we begin our discussion of knowledge bridging at the idea and individual level before moving into the firm level.

A. Ideas and individuals in knowledge bridging

In this section, we propose two modes by which knowledge bridging can take place, illustrate these modes by discussing innovators in biotechnology, and discuss the related academic literature. We conclude the section with an empirical prediction relating knowledge bridging to innovative outcomes.

Consider Figure 1, which shows two types of knowledge bridging. A first type involves taking knowledge from one domain and reapplying it to another. An example of this "porting" form of knowledge bridging is the birth and development of the academic field of evolutionary economics. Borrowing key ideas from evolutionary biology—such as principles of genetic variation and selection—evolutionary economists have advanced our knowledge of how organizations evolve in a way analogous to that of living organisms. The term "porting" has been used by Baldwin and Clark (2000) to describe the concept of applying problem solving from one domain for use in another, which they argue is a basic operator for modular systems.[2] Adner and Levinthal (2000) use a similar concept in their discussion of how technological "speciation" occurs, which introduces necessary variety to an organizational gene pool. More generally, Gavetti and Rivkin (2005) argue that problem solving by thinking through analogies can be a powerful tool leading to innovative thinking (but also caution that this method of problem solving can also be a pitfall for managers).

Figure 1 also shows a second type of bridging, which involves borrowing ideas and knowledge from several areas and recombining them for innovation in another area. To illustrate this, consider the academic field of strategic management. It borrows knowledge from a number of disparate fields such as economics, sociology, history, and political science—and recombines insights and methods from these fields to create new knowledge about corporate strategy.

To illustrate knowledge bridging, we share a few examples of individuals who were able to come up with novel new ventures and scientific breakthroughs in biotechnology as a result of bridging

[2] While we use terms from the modularity and product development literatures because we believe it describes the phenomenon we wish to empirically study, we abstract away from the manufacturing product settings that characterize that literature, while retaining the concepts.

knowledge domains. Consider first the case of serial entrepreneur Alejandro Zaffaroni, who successfully launched seven biotechnology companies across different fields of the industry. One of his former colleagues remarked about Zaffaroni: "…he is reading and thinking very widely. He is totally unafraid of any new technology in any area of human creativity. He has wonderful contacts with people in many different areas, so he sees the bridges between otherwise disparate fields" (as quoted in Burt, 2004).

A second individual, Kary Mullis, invented what has become a staple of the microbiology laboratory: polymerase chain reaction (PCR) technology. Cetus Corporation hired Mullis in 1979 to synthesize oligonucleotide probes.[3] By 1983, however, oligonucleotide synthesis was becoming reliably automated, and Mullis was facing obsolescence in his job as a chemist at Cetus. With more time on his hands, Mullis began "puttering around" with oligonucleotides and became interested in ways to easily detect single base pair changes (against a known sequence) in DNA. Since a genetic mutation may indicate the presence or the potential for a disease, Mullis was interested in finding a potential diagnostic application (Mullis, 1990). Planning this experiment led Mullis to the invention of PCR in the spring of 1983. While driving to his cabin in California, Mullis came up with the breakthrough idea that using two oligonucleotide primers working in opposite directions on each strand of denatured DNA, he could create instructions to continually "amplify," or replicate, specific DNA targets (Yoffe, 1994). Mullis had been spending a lot of time writing computer programs and recognized the power of reiterated loops; he envisioned PCR to be such a loop. When he got back to Cetus, Mullis spent three months running experiments before achieving success. Mullis won the 1993 Nobel Prize in chemistry for his invention.

While Mullis relied more on his in-depth knowledge of chemistry in relation to his knowledge of computer science for the PCR invention, importing ideas and concepts from across academic fields appeared important in his discovery. In other instances, deep knowledge of two or more disparate academic fields was important for the innovative process. Consider George Church, a professor specializing in bioinformatics at Harvard Medical School: "Church's ability to bring together information technology and experimental genetics has made him a '*force majeure* in science,' according to Philip Leder, Andrus professor of genetics and head of the genetics department at Harvard Medical School. Far from being 'just a computer geek,' Leder says, Church is a polymath who 'has terrific ideas that nobody else would think of putting together, because of the many disciplines he has mastered'" (Thomas, 2004). In our typology of knowledge bridging, Zaffaroni and Church illustrate knowledge recombination, while Mullis exhibited knowledge porting in his PCR discovery.

[3] An oligonucleotide is a short chain of specifically-sequenced nucleotide bases. The oligonucleotide can bind specifically with a string of complementary nucleotide bases in single-stranded DNA, and when radioactively labeled, engineered oligonucleotides can serve as probes for detecting whether a sample of DNA contains a particular gene or nucleotide sequence.

These examples and accounts are consistent with the academic literature. Schumpeter (1934: 65-66) conceived innovation as "carrying out new combinations." Usher (1954: 21), in his classic work, argued: "There are other discontinuities that may be overcome, through some act of synthesis. The establishment of new organic relations among ideas, or among material agents, or in patterns of behavior is the essence of all invention and innovation." More recently, Weitzman (1998) developed a model of economic growth that depends entirely on idea recombination.

Economists have investigated the idea that knowledge can flow across industrial boundaries through general purpose technologies (GPTs), defined as "key functional components embodied in hardware that can be applied as elements or modular units of the engineering designs developed for a wide variety of specific operations or purposes" (David, 1990: 355). While GPTs such as machine tools, electricity, and semiconductors have been investigated through extended historical accounts (e.g., Rosenberg, 1963) and in theoretical treatments (e.g., Bresnahan and Trajtenberg, 1995), empirical studies at lower levels of aggregation are still rare (see the recent papers by Hall and Trajtenberg [2005] and Roy [2005] for exceptions).[4]

Management scholars studying knowledge flows have tended to concentrate on boundary-spanning contexts and processes. One early research stream emphasized the importance of technological "gatekeepers," individuals that facilitate inter-organizational communication and cooperation by spanning organizational and sub-unit boundaries (e.g., Allen, 1977; Tushman and Scanlan, 1981). A strategic version of the boundary-spanning concept was put forth by Burt (1992). Individuals or actors who connect structural holes in a network are in an economically privileged position because the unconnected parties have few or no alternative routes to link themselves outside of the boundary spanner. Indeed, good ideas can come out of knowledge bridging behavior (Burt, 2004).[5]

---

[4] A related strand of research has observed that knowledge spillovers exist as a result of imperfect appropriability (e.g., Griliches, 1992), and that geographic proximity may be needed to capture the spillovers (Jaffe, Trajtenberg and Henderson, 1993). The availability of knowledge spillovers is aided by U.S. institutions such as the patent system that exchanges intellectual property protection for invention disclosure and the academic promotion system based on publishing papers in publicly-available professional journals. Such knowledge spillovers help explain the significant inter-industry knowledge use patterns documented by Scherer (1982). These knowledge flows from one industry to another have been recognized as a productive source of economic growth.

[5] It is interesting to speculate whether knowledge boundary spanning primarily takes place at the individual or team level. Brokerage opportunities must be the result of connecting spaces in which there has been limited amount of connection (competition in connecting those spaces would dissipate rents). Most interconnections do not hold value (or are costly to connect relative to expected benefits). Knowledge specialization of disparate areas can lead to interconnection, but making the investment will not be worthwhile for the average person. Of course having fragmented knowledge across the organization can also be costly (Henderson and Clark, 1990). Most of the prior literature has focused on diversity of knowledge at the team rather than the individual level. Moreover, valuable knowledge can be stored at the individual, team, and/or organization levels, with perhaps different decay rates associated with each.

This discussion suggests that knowledge bridging may not just lead to good ideas, but also those that are seminal, and with many wide-ranging applications.

B. Factors affecting knowledge bridging use at the firm level

We believe knowledge bridging operates similarly at the firm level. If our hypothesis that knowledge bridging-based technological search is related to innovation is correct, one naturally would like to know what affects firms' use of this mechanism (it is difficult to empirically study this question at the patent level since patent-level covariates would not be meaningful in this regard). In this section, we review the literature on factors affecting the search behavior of firms to derive empirical predictions.

Researchers have conceptualized entrepreneurship and new venture creation as activities that recombine disparate fields. Lazear (2004) sees entrepreneurs as generalists with training in several different areas, a quality which facilitates entrepreneurial opportunity recognition. This view is consistent with Biais and Perotti (2003), who argue that entrepreneurs, being non-specialists who see functional fit across areas, need to consult external specialists for advice.

At the firm level, the phenomenon of using knowledge and ideas from one context to solve problems in another has been examined ethnographically by Hargadon and Sutton (1997). They observe that a prominent product development firm, IDEO, as it sought to solve problems for various clients, often draw on an inventory of accumulated knowledge to customize a solution for a given situation. More generally, the re-application of prior problem-solving strategies into new settings is pertinent to settings such as legal research, management consulting, or computer programming. While verbatim importing likely represents a minority of cases in these settings, one could imagine "cutting and pasting" applicable modules to new settings for productive use.[6] In sum, the same processes of knowledge bridging operating at the individual level could also be taking place at the organizational level. We now turn our attention to the possible antecedents of knowledge bridging.

Persistence of search direction. One well-documented phenomenon in organizations is the persistence of search direction. Founder imprinting and search routinization are powerful reasons for such organizational inertness. Founder imprinting can be manifested in firm philosophies, policies and procedures as they relate to organizational culture, human resource management, and research and development practices (Stinchcombe, 1965). Organizational search has its roots in individual cognitive patterns, suggesting that search patterns tend to be subject to routinization (March and Simon, 1958). Such behavior is also part and parcel of organizational life (Nelson and Winter, 1982).

[6] Management scholars have also identified a process of technology melding, taking technologies from two different domains and creating a novel application (Kodama, 1992; Levinthal, 1998).

Indeed, Helfat (1994) and Henderson and Cockburn (1994) find substantial (and varied) fixed firm effects in research and development across two different industries: petroleum and pharmaceuticals, suggesting substantial within-industry heterogeneity in R&D investment strategy. Furthermore, Boeker (1989) found that semiconductor start-ups typically maintained the corporate strategies they had at the time of founding. Cockburn, Henderson and Stern (2000) suggest that initial conditions are important in explaining the level of adoption of a firm's R&D policies in a later period (in their case, the initial extent of science-driven drug discovery by pharmaceutical firms persisted over long periods). Their results also suggest that while initial conditions are important, they are not all-encompassing in explaining variation in the adoption of strategies that may affect organizational performance.

Because individuals have different backgrounds, knowledge, and skills, they will likely respond to entrepreneurial opportunity windows in different ways (Shane, 2000). This implies that individual beliefs about exploiting even a common technological innovation may very well be varied, which can account for differences in the initial position of entrepreneurial start-ups. Given the dual forces of technological search routinization and founder imprinting, *we expect firms' initial search conditions to persist over time*.

Mechanisms of overcoming local search. Given the constraints imposed by local search, it is not surprising that scholars have investigated ways to overcome such behavior, particularly in the context of fast-changing R&D environments. For example, Mowery, Oxley, and Silverman (1996) and Stuart and Podolny (1996) examine strategic alliances as a mechanism for overcoming local search. As well, engineers and scientists are mobile, and hiring on the scientific labor market is another means by which firms can access distant knowledge (Almeida and Kogut, 1999; Rosenkopf and Almeida, 2003).[7]

Another possible mechanism for overcoming local search is to recombine existing organizational knowledge (Kogut and Zander, 1992). Doing so would inject valuable variation into the search process (by experimenting with new components or recombining old ones), perhaps in a manner akin to the act of invention itself (Fleming, 2001). This discussion points to *two predicted mechanisms of promoting knowledge bridging search: through strategic alliances and labor market mobility of scientists.*

Antecedents to knowledge bridging search. The empirical literature that addresses triggers of change in technological search is surprisingly limited. Cyert and March (1963) discuss the effects of organizational failure as a proximate cause of organizational search, and subsequent empirical research shows that poor organizational performance encourages firms to seek more risky investments, which in

[7] The efficacy of the latter mechanism is likely to be context-specific, however. For example, Zucker, Darby and Brewer (1998) find that in the early biotechnology industry, the scarce resource was specialized knowledge resident in highly accomplished university scientists. The fact that these scientists were for the most part not mobile helps explain the observed geographic concentration of the industry (large concentrations of firms located near academic centers of excellence in biology and chemistry).

turn can lead to worse future performance (Bromiley, 1991). As well, organizational search importantly depends on managerial aspiration levels (e.g., Greve, 1998 and references therein). Consequently, organizational search should depend on feedback from product development efforts and other signals from the market: product market acceptance should reduce knowledge bridging search, with the opposite holding true with product market setbacks.

It is well recognized that for start-up firms, resource constraints, such as access to financial and human capital, often limit business development. When financial resources are available to the new venture, such as following receipt of venture capital (VC) funding or during "hot" fundraising conditions, firms may enjoy more organizational slack and surplus resources, and may therefore experiment and engage in more exploratory search, with the opposite effect when such financial resources are not readily available. In addition, VCs can act in extra-financial ways influencing business development and governance in start-ups (see Hsu [forthcoming] and references therein).

Finally, there may be entry order effects associated with competitive dynamics within an industry. As time passes, we predict that the knowledge base of entrepreneurial teams will need to be more specialized and/or span more boundaries in order to reach the competitive frontier. This results because the rent-generating potential of intellectual capital dissipates as knowledge diffuses and is commoditized, thus leading to potential opportunity exhaustion. Knowing this, potential new entrants will evaluate the expected value of their innovations (as compared to their opportunity costs) before making their entry decision. There is an interdependence of own and competitors' knowledge base in shaping search direction (Stuart and Podolny, 1996). This effort at avoiding head-to-head competition with industry incumbents represents a second mechanism, reinforcing the opportunity exhaustion effects.

The discussion in this section yields *three predictions of events that will increase a firm's use of knowledge bridging-oriented technological search: (1) negative feedback in a firms' product development progress, (2) munificent financial resources, and (3) later entry into the industry*.

III. Data and Method

To test these hypotheses, we require a sample of firms that were founded to exploit a given technological opportunity. Constructing a sample of firms that is relatively uniform in the basic technology upon which they are capitalizing allows us to observe differences in initial conditions such as founder capabilities, along with the subsequent evolutionary development patterns for these firms.

The commercialization of recombinant DNA following its discovery in 1973 by University of California-San Francisco scientist Herb Boyer and Stanford scientist Stan Cohen provides a fortuitous empirical setting. The history of the landmark Cohen Boyer patent is recounted elsewhere, and so we will not duplicate those efforts here (see e.g., Reimers, 1987 and Hughes, 2001). Instead, we merely note that

Stanford University conducted an open non-exclusive licensing program of the patent (which they advertised in the scientific journals *Science* and *Nature*), and so we are able to observe with great precision de novo firms founded to commercialize recombinant DNA technology (users of the technology that did not participate in the licensing program would be infringing the patent and subject to litigation).[8] Because we are interested in studying the evolutionary development of knowledge bridging search, we track new ventures established to commercialize the Cohen-Boyer patent. Aside from the clear scientific importance of the Cohen-Boyer innovation, the patent was also clearly important commercially: over its lifetime, the patent yielded approximately $200M in licensing revenues, which implies product sales based on the innovation of some $40B.[9]

A. Method

Researchers have used two types of measures of technological search, both of which are based on patent citation data. The first type of measure is used by Rosenkopf and Nerkar (2001). In the context of optical disk drive firms these authors use citations to non-disk patents as a measure of technological exploration and non-self citations as a measure of organizational exploration. While these are useful measures of exploration, we wish to develop measures that are more specific to the concept of knowledge bridging. A second class of measures examine the degree to which a firm overlaps with another firm's technical knowledge (Mowery, Oxley and Silverman, 1996), or the extent it overlaps with a common stock of scientific knowledge in a network-centric sense (Stuart and Podolny, 1996). Both of these approaches use measures that rely on knowledge overlap across firms. In contrast, our conceptualization of knowledge bridging search emphasizes the overlap between the technical domain a firm relies upon and the technical area in which it produces new knowledge.

We therefore elect to develop new measures to better capture the knowledge bridging concept. These measures are based on the patent classes of focal patents (a measure of knowledge outputs) *in relation* to the patent classes of the patents they cite (a measure of technological knowledge inputs). Following Jaffe (1986), we believe that using patent class data is a good way of capturing the concept of technological position.

---

[8] The Cohen Boyer invention was covered by three patents, with the most important being a process patent, U.S. patent number 4,237,224, entitled "Process for Producing Biologically Functional Molecular Chimeras." This patent, which became the backbone of the Stanford Technology Licensing Office's licensing efforts of recombinant DNA, was issued on December 2, 1980, and expired 17 years later, in 1997. Stanford offered licenses to the patent for a modest fee ($10,000 annual payments, with 0.5% royalty rates on end products).

[9] Between 1980 and 2000, the patent was cited 235 times by other patents, while the average patent of this vintage in this technology class was cited 9.64 times (Jaffe and Trajtenberg, 2002). Despite the economic value of this patent, which yielded such products as recombinant growth hormone and recombinant insulin, its legal validity was not subsequently challenged.

The first step in our method is to identify start-up firms that entered as a result of opportunities to commercialize recombinant DNA technology. We rely on the Stanford University Technology Transfer Office's records of licensees to the technology. We include firms in this sample if: (1) they are de novo firms (as opposed to established pharmaceutical firms), and (2) licensed the Cohen Boyer patents at the time of founding, or within a time window of two years after their founding. This process yielded a total of 19 firms. Using the U.S. patent database, we identified all of the patents granted to these firms between January 1976 and December 2004. This resulted in a dataset of 3,652 firm-patent pairs. We then traced backward citations (references made by these patents) to all other U.S. patents.[10] We also traced all forward citations through 2004; these are citations made by other U.S. patents to those in our sample. In total, our dataset contains 26,770 backward citations and 22,676 forward citations. For each focal patent, we obtained patent class information as well as the names and addresses of each inventor (2,901 persons). Finally, we identified all other patents awarded to the same inventors, thereby building an innovation profile of each inventor over time.[11] We conduct two analyses, the first using the patent-firm as the unit of analysis and the second using the firm-year. The former analysis examines the importance of knowledge bridging, while the latter sheds light on its antecedents. The following section describes the variables and empirical tests used in the analyses. The summary statistics and descriptions of all variables are found in Table 1.

B. Patent level analyses

To assess the potential impact of knowledge bridging, we investigate two patent-level dependent variables: forward patent citations and patent "generality."[12] The forward citation measure is the count of future patents that reference the focal one (excluding self citations), evaluated at five years post focal patent issue (mean = 2.4 in the dataset), and through the year 2004 (mean = 5.3). *Patent generality* (Henderson, Jaffe, and Trajtenberg, 1998) measures the diversity of patent classes of the forward patent

---

[10] Approximately 3.5% of backward citations are to patents issued prior to 1976. These are not available electronically from the U.S. Patent Office; we therefore used the Delphion database for these data. Therefore, our dataset contains *all* backward citations regardless of dates, and so left-censoring of the data is not an issue.

[11] We found 22,491 patents awarded to inventors with these or similar names. A research assistant was assigned the arduous task of filtering this dataset row by row, identifying each unique inventor based on their names as well as the address of the company the patent was assigned to. The main difficulty encountered was with common names (did an inventor work in multiple firms or were different people with the same name work across those firms?). There are only 41 such inventor names in our database, accounting for 1,142 patents. For these cases, we set a dummy variable to 1, and this variable is included in the regressions when appropriate.

[12] One reason we investigate impact at the patent level is because the biotechnology industry (and biopharmaceuticals more generally), is notorious for the length of product (drug) development. Evaluators (e.g., venture capitalists) use patent information as a signal of value, and patent-based measures have been validated in a number of academic studies (e.g., Jaffe and Trajtenberg, 2002).

citations (mean = 0.54).[13] A high generality score suggests that a patent is being cited by a broader range of technological classes, and is likely to be more "general-purpose" in its application.

The right hand side variables include measures of *patent knowledge bridging* and controls for patent characteristics, patent application year fixed effects, and start-up fixed effects. The key variable of interest is *patent knowledge bridging*, which we define as [1 – (share of cited patents that are in the same primary class as the focal patent)] (mean = 0.52). The extent to which inventors cite patents in different areas relative to the area of the focal invention indicates the degree of knowledge bridging. High measures of *patent knowledge bridging* imply substantial use of scientific knowledge originating from outside the focal patent area.

An alternate definition of knowledge bridging is the variable *patent originality* (mean = 0.53). This variable is defined similarly to *patent generality*, but uses backward citations instead (Henderson, Jaffe, and Trajtenberg, 1998). The higher a patent's originality score, the more diverse are the citations made by that patent to different technological classes. While *patent originality* is related to *patent knowledge bridging*, the conceptual difference is important: *patent originality* measures the breadth of patent classes cited, while *patent knowledge bridging* measures the overlap between a patent's own class and those it cites.

A number of control variables are used in the analysis. *Patent scope* (mean = 2.71) is the number of international patent classes assigned to the focal patent, which has also been correlated with economic value (Lerner, 1994). Patent cohort effects, as embodied in a set of indicators of *patent application years*, are important controls when explaining forward citation rates (recent patents are likely to receive fewer forward citations than earlier ones). Finally, the variable *inventor patent experience at other firms* is defined as the number of patents issued to a focal patent's inventors while employed by *other* organizations prior to the application date of the focal patent (mean = 6.6). The measure is meant to capture the mobility of scientific knowledge through human capital.

C. Firm-year analyses

We next investigate the antecedents to knowledge bridging using firm-year regressions. Before describing the variables used in that analysis, let us first discuss a patent classification issue. For the firm-level analyses, we categorized the current U.S. primary classes of these patents (and the primary classes of the patents cited) using a modified version of the categorization proposed by Jaffe and Trajtenberg

---

[13] The measure is computed using the formula: $G_i = \left[1 - \sum_{j=1}^{J}\left(\frac{N_{ij}}{N_i}\right)^2\right]\left(\frac{N_i}{N_i - 1}\right)$, where i indexes the patent, j indexes patent classes, and N represents counts of forward citations. The expression outside of the square brackets adjusts for bias associated with small numbers of forward patent counts (Hall and Trajtenberg, 2005).

(2002, Appendix 1).[14] For our purposes, the Jaffe and Trajtenberg (2002) classes are a good starting point because it collapses an almost unwieldy number of three digit U.S. patent classes into industry family groups. While this classification for industry groups outside the biotechnology and drug discovery areas is useful for this study, the classification within biotechnology is too coarse for our purposes. Specifically, Jaffe and Trajtenberg's class 33, biotechnology, contains a high fraction of the primary patent classes in our sample. Fortunately, the U.S. Patent and Trademark Office's Manual of Classification includes 41 sub-categories (at the six digit level) for biotechnology patents, which we collapse into 25 distinct biotechnology areas based on the opinions of two experts.[15] Our categorization scheme combines the original Jaffe-Trajtenberg classification with the USPTO subclasses within biotechnology, and we use it to construct the dependent variable at the firm-year level of analysis, *firm knowledge overlap*.

At the patent level, we measure knowledge overlap as the extent to which a focal patent cites other patents in the same technological class. However, most firms are likely to obtain more than one patent in a given year, possibly in multiple technology areas. As such, for the firm-year analysis, we need to construct a multi-dimensional measure of knowledge overlap that takes into consideration the entire set of patents and patent classes cited each year. The first step is to obtain a vector of focal patents' current classes, ***a***, for a given firm-year. Similarly, assemble a vector of patent classes for the patents cited by the focal patent, ***b***. Next, construct the vector dot product, ***a•b*****,** which measures the degree to which the vectors ***a*** and ***b*** are aligned. The variable *firm knowledge overlap* is constructed by calculating the following expression: $\left[-\text{arccosine}\ (\frac{a \bullet b}{|a||b|})\right]\left[\frac{180}{\pi}\right]$. This measures the angle, in degrees, between the two vectors. A higher value of *firm knowledge overlap* (mean = -40.5 degrees) corresponds to greater similarity between the patent classes of the focal patent and those it cites. At one extreme, both the citing and cited patent classes could be parallel, meaning that the knowledge produced is directly in line with the knowledge cited (in this case, *firm knowledge overlap* = 0). At the other extreme, the citing and cited patent classes could be orthogonal, so that the knowledge produced is in a completely different direction than the knowledge cited (in this case, *firm knowledge overlap* would be –90 degrees).

We employ three sets of independent variables in the analysis: a measure of initial conditions, firm-size measures, and other firm-specific measures. Each is discussed in turn. The prior literature

[14] U.S. patents are classified at both the three digit and six digit levels (with the latter being more detailed than the former). See Lerner (1994) for more institutional details about U.S. patent classifications. Since patent classes are sometimes reclassified over time, it is important to use patent classes in the sample at a given time, which in our case is the 2004 classification. Using patent classes at the time of their issue does not significantly alter the results.

[15] Two biotechnology scientists from the University of Pennsylvania Medical School faculty were interviewed regarding the USPTO classifications. Working cooperatively, they suggested a collapse of the 41 USPTO subclasses into 25, reflecting distinct domains of biotechnology research. In this paper we present the results using their recommendations. Our results are consistent, however, whether we break biotechnology into 41 subclasses, 25 subclasses, or use the USPTO 3-digit patent classes.

suggests that taking account of initial search conditions is important. We adopt Cockburn, Henderson, and Stern's (2000) philosophy of examining organizational strategy while taking into account the potential importance of imprinted initial conditions. We do this by constructing a variable, *founder overlap*, that calculates a dot product (as before) comparing the degree to which the patent classes of each firm's founders' initial patents within the first three years of founding (a stationary vector) overlaps with those granted to the firm in a given year (a time-varying vector) (mean = -58 degrees).

Two control variables proxy for firm size. The variable, *employees in t-2*[16] (mean = 720) is a firm's headcount lagged by two years. The variable *revenues in t-2* measures gross revenues of a firm two years in the past (mean = $171.4M), but is not used in the empirical analysis because it is almost perfectly collinear with *employees in t-2*. Data to code these firm size variables is taken from the CorpTech directory of technology firms and from firms' public filings, such as the 10-K and the annual report to shareholders.

A third group of right hand side variables measure other firm-specific differences. The first two of these explore mechanisms highlighted in the literature on overcoming local search (and so may reflect organizational efforts to change their strategic direction). The variable *cumulative equity strategic alliances in t-2* is a count of the total number of equity strategic alliances a firm has undertaken up until two years before the observation year (mean = 1.05). The alliance data are obtained from Recombinant Capital. The variable *firm inventor mobility in t-2* counts the number of inventors who apply for patents at the focal firm and who also have prior experience patenting at another organization, evaluated with a two year lag (mean = 2.11).

The next two firm-specific measures gauge product development progress. Since the overwhelming majority of the firms in the dataset innovate in human therapeutics, we collect information about each firm's lagged annual count of Federal Drug Administration (FDA) drug approval progress.[17] The variable *cumulative FDA drug development approvals in t-2* tallies the cumulative number of FDA drug approvals the firm received two years before the observation year (mean = 2.2). At any time during the FDA drug approval process, firms can decide to terminate projects. We interpret these terminations as negative signals to the firm, such as disappointing clinical trials. The variable *cumulative drug development terminations in t-2* counts the stock of such terminations two years from the observation period (mean = 3.03). Data to code these variables are from the Recombinant Capital database.

---

[16] The time lags associated with this and other independent variables are meant to capture the idea that the current observation of firms' technological search will depend on events that took place in the past. In the empirical analysis, we check for sensitivity to the two year lag used in the reported specifications, and find that the results are generally robust for a one to three year lag, though because of the limited number of observations in this dataset, we report two year lags.

[17] More details of the drug approval process can be found at the FDA's website, www.fda.gov.

Because the drug discovery process is long and involved, most biotechnology firms are not able to generate working capital based on sales revenue in the short run. As a result, many (if not most) biotechnology firms are dependent on pharmaceutical firm alliances, venture capital (or other private) funding, and public offerings for financial capital. VC funding events, for example, can have organizational impact in start-ups' business development (Hsu [forthcoming]), and so we construct two variables to capture these effects. *Venture capital funding in t-2* is an indicator = 1 if the start-up firm was venture capital funded two years prior (mean = 0.76). Similarly, the variable *initial public offering in t-2* is an indicator for that event, which may act as a source of organizational slack (mean = 0.50).

Finally, because there might be entry order effects, we code the order in which firms entered the industry based on founding year and month data, and label this variable *entry order*. The earliest entrant is assigned a value of one, and later entrants are assigned increasing numbers. If two or more firms were founded in the same month and year, they are assigned the same entry order number (there is one instance in which we were not able to establish entry sequence for three firms).

IV. Empirical Results

We present the empirical results in three stages. First, we examine regression results of the innovative impact of knowledge bridging at the patent level of analysis. Second, we describe knowledge input - output matrices for firms in the sample to illustrate representative knowledge bridging patterns over time. Finally, we step through regression results of the antecedents of firm-level knowledge bridging-based technological search.

A. Patent level regression results

Table 2 initiates the patent level analysis. In the first three columns, we examine correlates of *patent forward citations* within five years of patent issue (excluding self-citations), while in the last column we analyze these forward citations through 2004.[18] Negative binomial count regressions are used throughout these regressions due to the count nature of the dependent variable. In each of the specifications, we include patent application year fixed effects to account for invention cohort effects and firm fixed effects to control for unobserved differences across firms. In addition, the results are also robust to a set of fixed effects for three digit patent classes of the focal patents, which adjusts for the possibility that inventions in different technical classes have varied "risks" for forward citations (unreported regressions). The first two columns show that *patent knowledge bridging* is positive and

[18] We report the more conservative estimates of forward citations excluding self-citations to emphasize the importance of knowledge bridging across organizational boundaries. The results are also generally robust to inclusion of self forward citations, though results using such citations on *forward citations through 2004* are statistically noisier.

significant at the 1% level, both in a parsimonious specification as well as when controlling for *patent scope* and *inventor patent experience at other firms*. The point estimates of the bridging variable across the two specifications are remarkably stable and are economically significant: a one standard deviation increase in *patent knowledge bridging* is associated with a 7.5% increase in *patent forward citations*. In the second specification, *patent scope* is also positive and statistically significant (the implied marginal effect of a one standard deviation increase in *patent scope* is 5.6%). Notice also that the variable *inventor patent experience at other firms*, the measure of exploratory search via scientific labor market mobility, is not statistically different from zero.

One might argue that an alternative definition of knowledge bridging is captured by the *patent originality* variable, which indicates the diversity of primary patent classes a focal patent cites. While this measure is 58% pair-wise (and significantly) correlated with *patent knowledge bridging*, we believe that the concepts behind the measures are distinct. *Patent originality* does not make the comparison between inputs (backward citation patent classes) and outputs (focal patent class). Nevertheless, specification (2-3) takes *patent originality* as the measure of knowledge bridging, and finds it to be both statistically significant (at the 1% level) and economically important (a one standard deviation increase in *patent originality* is associated with an 8.3% increase in the dependent variable). The statistical significance (or lack thereof) of the variables *patent scope* and *inventor patent experience at other firms* mirrors that of specification (2-2). The final column of Table 2 replicates the logic of specification (2-2), but uses as the dependent variable the count of forward citations through 2004 (excluding self citations). Due to the censoring of forward citations, it is important to include the patent application year fixed effects.[19] The results here are consistent with the prior ones.

Table 3 reports fixed effects OLS regressions with patent generality as the dependent variable. Again, in all the specifications, we include patent application year and firm fixed effects (we do not report patent class fixed effects, though the results are robust to their inclusion). In addition, the standard errors are robust and take into account clustering by firm. We replicate the logic of the prior table by examining in turn *patent knowledge bridging* and *patent originality* as measures of knowledge bridging search behavior. We also add *forward citations through 2004* to the second and fourth specifications, as the variable adjusts for the quantity of forward citations—and so addresses the alternate hypothesis that knowledge bridging is related to more general future inventions as a result of being associated with more forward cited patents. While *forward citations through 2004* is estimated with a positive and significant coefficient, the implied economic effect is modest. The main result of Table 3, however, is that in each

[19] An alternate approach is to deflate the forward citations by the average value for its scientific field-year cohort as a fixed effect, as discussed in Jaffe and Trajtenberg (2002). Because we do not use the National Bureau of Economic Research dataset for our patent data (the NUS patent project allows us to access more recent patent data), we do not use these deflators in our analysis.

case, the measure of knowledge bridging is positive and significant at the 1% level. At the means of the other independent variables, a one standard deviation increase in *patent knowledge bridging* and in *patent originality* in equations (3-2) and (3-4) are associated with 7.1% and 7.3% increases, respectively, of *patent generality*. Taken together, the patent level results suggest that knowledge bridging search is associated with the production of patents that are of higher quality and generality.

B. Firm level input-output knowledge matrices

We now turn our attention to understanding bridging knowledge domains at the firm level. In the firm level analyses, our motivation is to understand causes of knowledge bridging use, now that we have seen the importance of bridging in innovative outcomes at the patent level of analysis. As a first step, we provide some descriptive results from the raw data before moving to a more systematic regression analysis in the next section.

A first set of charts illustrate firms' backward citation patterns across time. Figure 2a shows the number of citations by U.S. patent class made by New England Biolabs' patents to various primary patent classes on a yearly basis. From this figure, it is clear that the company relies mainly on knowledge in U.S. patent class 435, and that the pattern persists throughout the 17 year period shown. By comparison, Biogen cites a broader set of technological classes (Figure 2b), including US patent classes 424, 435, 514 and 530. This pattern is also fairly stable across time, though a number of other classes are clearly represented. Extending this analysis to other firms in the sample reveals a large variation in the plots: there are large inter-firm differences that persist over time. This descriptive finding suggests that firms' initial search directions tend to persist over time.

Next, we turn our attention to a descriptive analysis of knowledge origin vs. knowledge use. Figure 3 illustrates modes of knowledge bridging at the firm level for firms in the data set. In considering types of bridging, it is useful to compare them against a base case of no bridging. In this organizational search mode, firms are drawing primarily on search within the domain in which they hope to innovate. We contrast this situation with taking knowledge from one area and applying it to another through problem solving strategies employing either porting or recombination. Each row in a table shows the primary classes of a firm's patents, while each column indicates the primary classes of the patents it cites. The patent class shown is that of the modified Jaffe-Trajtenberg scheme, which is also used in the regression analysis below. Figure 3a shows the results for Visys, which produces patents in a fairly narrow set of technical classes (primarily 33d). However these patents make citations to a broad range of other patent classes, including classes 14, 19, 33d, 33e. This represents a case of bridging from those other classes into class 33d. Figure 3b shows another example, that of Therion. It produces patents in a number of different classes, but most of them cite patents that are in class 33j. This illustrates a case of bridging from class 33j to other technological areas. Finally, Figure 3c shows the patenting behavior of New England Biolabs. This firm mainly produced patents in class 33i, but these cite patents in class 33d more often than those in class 33i itself. Interestingly, the firm produces only a small number of patents in class 33d, even though 33d is so heavily cited by patents in class 33i. A likely interpretation of this pattern is that New England Biolabs is taking knowledge from class 33d and transferring it into patentable inventions within class 33i.

The differences we observe across firms suggest the need for a more systematic analysis of the antecedents of firms' use of knowledge bridging search as well as the importance of accounting for fixed firm effects due to different organizational styles. Consequently, we now turn to a regression analysis examining correlates of knowledge bridging at the firm-year level.

C. Firm-year level regression results

Table 4 presents firm fixed effects OLS regressions of *firm knowledge overlap*.[20] In a first specification, we examine the role of *founder overlap*, our measure of imprinting, while controlling for *employees in t-2*. The estimated coefficient on *founder overlap* is positive and statistically significant at the 1% level, suggesting that founders' knowledge and search trajectory are importantly imprinted on the future direction of the firm. Larger firms as measured by employees are less likely to engage in knowledge bridging search.

In the second column, we investigate two mechanisms identified in the literature as instrumental in overcoming local search, strategic alliances and scientific labor market mobility. We enter *cumulative equity strategic alliances in t-2* and *firm inventor mobility in t-2* as the only two covariates of *firm knowledge overlap*. Both coefficients are positive (at the 1% and 10% statistical significance levels, respectively), suggesting that use of these firm strategies reduce knowledge bridging-oriented search in this data set. While these estimates do not remain statistically significant once other effects are controlled for (as we shall soon see), these findings are interesting nevertheless. Incidentally, the alliance effect is positive in the parsimonious specification across multiple definitions of alliance activity: those involving up-front payments, are research and development based, or include any type of alliance.

A third specification examines a number of start-up firm measures. The first two variables proxy for product development progress: *cumulative FDA drug approvals in t-2* and *cumulative drug terminations in t-2*. These two estimates are not statistically significant in the regression. *VC funding in t-*

[20] Despite censoring points (at -90 and 0 degrees) associated with the dependent variable, we chose to report fixed effects OLS regressions. A random effects tobit model results in estimates similar to those described in this section (a fixed effects tobit model is not available because the fixed effect cannot be conditioned out of the likelihood function). As well, we do not present cross sectional tobit models due to the longitudinal nature of the data set, although the results produced are similar.

2 is an indicator variable for VC receipt, and it is estimated with a large, negative effect (the estimate is statistically significant at the 8% level). The result suggests that VC funding is related to knowledge bridging search, and is consistent with the proposition that more munificent funding environments are associated with organizational slack, which in turn is associated with more exploratory technological search. Due to the lag structure of the variable, we do not believe that reverse causality is at play here (the argument that VCs select firms that use knowledge bridging search).[21]

In the final column of Table 4, we control for all effects simultaneously. The *founder overlap* and *VC funding in t-2* results remain robust (the latter at the 6% level, the former at the 1% level) in this specification. Furthermore, the estimated coefficients are economically significant: a one standard deviation increase in *founder overlap* is associated with a 5.8% increase in *firm knowledge overlap* at the means of the other right hand side variables. A discrete change into being VC-funded is associated with a 100% decrease in *firm knowledge overlap* at the means of the other right hand side variables.

We report on two robustness tests of these results. First, the results are robust to using the U.S. three digit primary technical classifications (instead of our modified classification scheme). Secondly, to account for possible prior selection effects (not all firms are equally likely to be able to use knowledge bridging search) we run Heckman two stage selection models. In the first stage, we estimate the probability of access to exploratory search mechanisms, which we proxy as alliance behavior (the regressors are variables that do not affect the primary *firm knowledge overlap* equation but may be correlated with access to search mechanisms: number of founders and number of founders who were faculty members at top university faculties). The results are also robust to this selection model.

Finally, to estimate *entry order* effects (which are not time varying, and so get absorbed by the fixed effects in the above regressions), we use cross sectional tobit models (described here, but not reported in tables). The *entry order* estimate in a fully-specified model is statistically significant (at the 5% level), but carries the opposite sign of what we expected. The result suggests that later entrants make less use of knowledge bridging search. The result should be interpreted very cautiously, however, because we use cross sectional methods in longitudinal data. Essentially, the effect is identified from only 19 observations in the cross section. A possible explanation for the observed empirical pattern may be the following (though it is hard to examine empirically): knowledge bridging may involve importing small amounts of information as the basis for recombination, and so later entrants may be recombining small

[21] We also test an alternate variable, *hot biotech funding environment in t-2*, based on Lerner's (1994) index of biotechnology funding environment (including funds from VC, initial public offerings and other forms of external funding for biotechnology firms) as an industry level proxy for funding environment munificence. The estimate is negative but not statistically significant. In addition to being available for only a subset of the time spanned by our observations (resulting in fewer usable observations), we prefer a firm rather than an industry level measure. As well, the index aggregates several types of funding sources, and is cumulatively more variable than the supply of VC going into the industry (which has implications for the lag structure of the variable).

amounts of external knowledge (while still relying intensively on the dominant body of knowledge relevant to a field). Nevertheless this phenomenon merits further investigation, most appropriately in a larger dataset.

V. Discussion and Conclusions

We investigate a type of technological boundary spanning-oriented search in which firms apply knowledge from one technical domain to innovate in another, a phenomenon we term knowledge bridging. We develop and explore measures capturing the concept of knowledge bridging-oriented search. In a carefully chosen empirical setting of firms founded to commercialize recombinant DNA technology, we find that patents exhibiting a high degree of knowledge bridging are associated with higher patent forward citation rates and generality scores. Our firm-year level of analysis reveals strong variation in the degree to which firms bridge knowledge domains, and shows that firms' initial search trajectories dampen knowledge bridging use, while the receipt of venture capital spurs such use. We do not find empirical evidence that strategic alliances and inventor labor market mobility predict knowledge bridging use, though an apples-to-apples comparison to the prior literature is difficult because different measures are used across the studies. Overall, this paper is the first to empirically link knowledge bridging to firm-level innovation and commercialization efforts.

It is important to discuss a number of interpretational issues. These concern the use of patent-based measures and interpreting the process by which knowledge bridging takes place. With regard to the patent-based measures, the costs and benefits to patent-based measures have been extensively discussed elsewhere (see for example, Jaffe and Trajtenberg, 2002). The main issue here is whether our knowledge bridging measures adequately capture the phenomenon of re-applying technical knowledge from one domain to another. We can make an analogy to academic journal articles: those that cite work from a variety of fields are likely to have borrowed, recombined (and possibly extended) external knowledge.

A second issue involving patent citation data is that inventors might strategically cite prior art across technical domains to appear more novel, thus improving the likelihood of receiving a patent. Inventors have an incentive not to over-cite in this manner, however, since doing so will tend to enlarge the relevant prior art (thus narrowing the scope of the patent). Patent examiners are charged with assembling the "right" citations, as patent citations are used as a legal device to circumscribe patent scope through identification of prior art (though their incentives to do this accurately are not clear). The ideal way to test for this effect would be to assemble a sample of patent applications, some of which are granted, others of which are not—and look for differences based on prior art. Without conducting a well-designed study on the topic, we are not prepared to speculate on potential bias arising from this issue.

A final issue relates to interpretational and methodological issues related to patent references and citations. Alcacer and Gittelman (2003) argue that patent examiner-imposed citations may be an important phenomenon. If true, then our calculation of the knowledge bridging measures may not accurately represent search behavior by scientists and organizations. Because the data on patent examiner-imposed citations are only available since 2001, we are not able to empirically examine the extent to which this phenomenon holds in our sample. We are ultimately concerned, however, with knowledge *use*, and as long as each patent does depend on other patents it cites for prior technical knowledge, we are not so concerned about whether a patent examiner or the inventor herself was responsible for adding those citations to the patent.[22]

Another set of issues relate to the process by which knowledge bridging-oriented invention takes place. First, the debate on the extent to which social interaction is necessary for invention (including knowledge bridging invention) is a long-standing one (e.g., Gilfillan [1935] versus Usher [1954]), and relates to the individual versus team nature of invention and innovation. While anecdotes supporting either view can be offered, it is difficult empirically to adjudicate between these views using patent data, as we only observe successful inventions which are granted patents. In any case, we know of no systematic effect in this realm that would bias our results.

Second, with respect to the mechanisms of achieving knowledge bridging, our measures of strategic alliances and inventor labor mobility are not significantly related to knowledge bridging use. As we caution before, however, it is difficult to directly compare our results against prior studies, not only because of the different measures used, but also because of the different types of exploratory search highlighted in this literature. One explanation of our finding on the scientific labor market front is found in the evidence and discussion by Zucker, Darby and Brewer (1998). They argue that during eras of new scientific knowledge (such as in the post Cohen-Boyer biotechnology industry), human capital-based knowledge is important and not easily mobile—as in the case of Ph.D. scientists in the early biotechnology industry—though this effect may dissipate over time as specialized knowledge diffuses.

We end with some thoughts on ways to extend this research. First, while we have taken a first step at empirically accounting for prior *access* to exploratory search mechanisms (which of course is a pre-requisite to using any form of boundary-spanning activity), we believe that this issue needs more systematic attention in this literature. Second, while we purposefully investigated knowledge bridging in a well controlled empirical setting, it would be useful to examine the phenomenon in other arenas to better

[22] In addition, Thompson and Fox-Kean (2005) critique Jaffe, Trajtenberg and Henderson's (1993) patent matching procedure (using patent citations) to infer geographic localization of knowledge spillovers. Jaffe et al.'s (1993) findings importantly rest on the ability to control for the pre-existing geographic distribution of inventive activity, which is done through a patent matching procedure pairing citing patents with comparable non-citing ones. The empirical design in our paper, however, does not rely on constructing patent citation-based matched samples.

understand the generalness of the results. Finally, a more micro-level analysis of how venture capitalists or firm managers facilitate or encourage knowledge bridging would be interesting. For example, to what extent do firm policies such as allowing scientists to engage in the broader open science community (e.g., Henderson and Cockburn, 1994) or setting aside time for research personnel to engage in their own scientific endeavors (such as at 3M or Google Labs) result in a more knowledge bridging? Exploring these and other firm-level mechanisms for successful cross-fertilization of ideas would deepen our understanding of the phenomenon.

References

Adner, R. and D. Levinthal (2000). "Technological Speciation and the Emergence of Emerging Technologies," in G. Day and R. Shoemaker, eds. Wharton on Emerging Technologies. John Wiley & Sons.

Alcacer, J. and M. Gittelman (2003). "How Do I Know What You Know? Patent Examiners and the Generation of Patent Citations," working paper, New York University.

Allen, T.J. (1977). Managing the Flow of Technology. MIT Press: Cambridge, MA.

Almeida, P. and B. Kogut (1999). "Localization of Knowledge and the Mobility of Engineers in Regional Networks," *Management Science*, 45: 905-917.

Baldwin, C. and K. Clark (2000). Design Rules: The Power of Modularity. MIT Press, Cambridge, MA.

Baron, J.N., M.D. Burton, and M.T. Hannan (1996). "The Road Taken: Origins and Evolution of Employment Systems in Emerging Companies," *Industrial and Corporate Change*, 5: 239-275.

Basalla, G. (1988). The Evolution of Technology. Cambridge University Press, Cambridge, UK.

Biais, B. and E.C. Perotti (2003). "Entrepreneurs and New Ideas," Toulouse and University of Amsterdam working paper.

Boeker, W. (1989). "Strategic Change: The Effects of Founding and History, *Academy of Management Journal*, 32: 489-515.

Bresnahan, T.F. and M. Trajtenberg (1995). "General Purpose Technologies: Engines of Growth?" *Journal of Econometrics*, 65: 83-108.

Bromiley, P. (1991). "Testing a Causal Model of Corporate Risk Taking and Performance," *Academy of Management Journal*, 34: 37-59.

Burt, R. (1992). Structural Holes: The Social Structure of Competition. Harvard University Press, Cambridge, MA.

Burt, R. (2004). "Structural Holes and Good Ideas," *American Journal of Sociology*, 110: 349-399.

Cockburn, I., R. Henderson, and S. Stern (2000). "Untangling the Origins of Competitive Advantage," *Strategic Management Journal*, 21: 1123-1145.

Cyert, R.M. and J.G. March (1963). A Behavioral Theory of the Firm. Blackwell Publishers: Malden, MA.

David, P.A. (1990). "The Dynamo and the Computer: An Historical Perspective on the Modern Productivity Paradox," *American Economic Review Papers and Proceedings*, 80: 355-361.

DiMasi, J.A. (2000). "New Drug Innovation and Pharmaceutical Industry Structure: Trends in the Output of Pharmaceutical Firms," *Drug Information Journal* 34: 1169-1194.

Fleming (2001). "Recombinant Uncertainty in Technological Search," *Management Science*, 47: 117-132.

Gans, J.S., D.H. Hsu and S. Stern (2002). "When Does Start-up Innovation Spur the Gale of Creative Destruction?" *RAND Journal of Economics*, 33: 571-586.

Gavetti, G. and J. Rivkin (2005). "How Strategists Really Think: Tapping the Power of Analogy," *Harvard Business Review*, April.

Gilfillan, S.C. (1935). The Sociology of Invention. MIT Press: Cambridge, MA.

Greve, H.R. (1998). "Performance, Aspirations, and Risky Organizational Change," *Administrative Science Quarterly*, 43: 58-86.

Griliches, Z. (1992). "The Search for R&D Spillovers," *The Scandinavian Journal of Economics*, 94: Supplement 29-47.

Hall, B. and M. Trajtenberg (2005). "Uncovering GPTs Using Patent Data," in Antonelli, Foray, Hall, and Steinmueller (eds.), *Festschrift in Honor of Paul A. David*, Edward Elgar.

Hargadon, A. and R. Sutton (1997). "Technology Brokering and Innovation in a Product Development Firm," *Administrative Science Quarterly*, 42: 716-749.

Helfat, C.E. (1994). "Evolutionary Trajectories in Petroleum Firm R&D," *Management Science*, 40: 1720-1747.

Henderson, R.M. and K.B. Clark (1990). "Architectural Innovation: The Reconfiguration of Existing Product Technologies and the Failure of Established Firms," *Administrative Science Quarterly*, 35: 9-30.

Henderson, R. and I. Cockburn (1994). "Measuring Competence? Exploring Firm Effects in Pharmaceutical Research," *Strategic Management Journal*, 15: 63-84.

Henderson, R., A. Jaffe, M. Trajtenberg, (1998). "Universities as a Source of Commercial Technology: A Detailed Analysis of University Patenting, 1965-1988," *The Review of Economics and Statistics*, 80: 119-127.

Hsu, D.H. (forthcoming). "Venture Capitalists and Cooperative Start-up Commercialization Strategy," *Management Science*.

Hughes, S.S. (2001). "Making Dollars out of DNA: The First Major Patent in Biotechnology and the Commercialization of Biotechnology," *Isis*, 92: 541-575.

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

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

Jaffe, A. and M. Trajtenberg (2002). Patents, Citations, and Innovations: A Window on the Knowledge Economy. MIT Press: Cambridge, MA.

Katila, R. and G. Ahuja (2002). "Something Old, Something New: A Longitudinal Study of Search Behavior and New Product Introduction," *Academy of Management Journal*, 45: 1183-1194.

Kenney, M. (1986). Biotechnology: The University-Industry Complex. Yale University Press: New Haven, CT.

Kodama, F. (1992). "Technology Fusion and the New R&D," *Harvard Business Review*, 70: 70-78.

Kogut, B. and U. Zander (1992). "Knowledge of the Firm, Combinative Capabilities, and the Replication of Technology," *Organization Science*, 3: 383-397.

Lazear, E. (2004). "Balanced Skills and Entrepreneurship," *American Economic Review*, May, 94: 208-211.

Lerner, J. (1994). "The Importance of Patent Scope: An Empirical Analysis," *RAND Journal of Economics*, 25: 319-333.

Levinthal, D.A. and J.G. March (1993). "The Myopia of Learning," *Strategic Management Journal*, 14: 95-112.

Levinthal, D.A. (1998). "The Slow Pace of Rapid Technological Change: Gradualism and Punctuation in Technological Change," *Industrial and Corporate Change*, 7(2): 217-247.

March, J.G. (1990). "Exploration and Exploitation in Organizational Learning," *Organization Science*, 2(1): 71-87.

March, J.G. and H.A. Simon (1958). Organizations. John Wiley, New York.

Mowery, D.C., J.E. Oxley and B.S. Silverman (1996). "Strategic Alliances and Interfirm Knowledge Transfer," *Strategic Management Journal*, 17: 77-91.

Mullis, Kary B. (1990). "The Unusual Origin of the Polymerase Chain Reaction," *Scientific American*, April, pp. 56-65.

Nelson, R.R. and S.G. Winter (1982). An Evolutionary Theory of Economic Growth. Harvard University Press, Cambridge, MA.

Reimers, N. (1987). "Tiger by the Tail," *Chemtech*, 17: 464-471.

Rosenberg, N. (1963). “Technological Change in the Machine Tool Industry, 1840-1910,” *Journal of Economic History*, 23: 414-443.

Rosenkopf, L. and A. Nerkar (2001). “Beyond Local Search: Boundary-Spanning, Exploration, and Impact in the Optical Disk Industry,” *Strategic Management Journal*, 22: 287-306.

Rosenkopf, L. and P. Almeida (2003). “Overcoming Local Search Through Alliances and Mobility,” *Maangement Science*, 49: 751-766.

Roy, P. (2005). “Leveraging Existing Technology: The Role of Alliances in Cross-Application across Technology Domains in Entrepreneurial Firms,” working paper, Wharton School.

Scherer, F.M. (1982). “Inter-industry Technology Flows and Productivity Growth,” *The Review of Economics and Statistics*, 64: 627-634.

Schumpeter, J. (1934). <u>The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest and the Business Cycle</u>. Harvard University Press: Cambridge, MA.

Shane, S. (2000). “Prior Knowledge and the Discovery of Entrepreneurial Opportunities,” *Organization Science*, April, 448-469.

Stinchcombe, A.L. (1965). “Social Structure and Organizations,” in <u>Handbook of Organizations</u>, J. March (ed.), Rand-McNally: Chicago, IL, 142-193.

Stuart, T.E. and J.M. Podolny (1996). “Local Search and the Evolution of Technological Capabilities,” *Strategic Management Journal*, 17: 21-38.

Thomas, P. (2004). “DNA as Data,” *Harvard Magazine*, January-February.

Thompson, P. and M.E. Fox-Kean (2005). “Patent Citations and the Geography of Knowledge Spillovers: A Reassessment,” *American Economic Review*, 95: 450-460.

Tushman, M.L. and T.J. Scanlan (1981). “Boundary Spanning Individuals: Their Role in Information Transfer and Their Antecedents,” *Academy of Management Journal*, 24: 289-305.

Usher, A.P. (1954). <u>A History of Mechanical Inventions</u>, 2nd edition, Harvard University Press.

Weitzman, M. (1998). “Recombinant Growth,” *Quarterly Journal of Economics*, 63: 331-360.

Yoffe, E. (1994). “Is Kary Mullis God? Nobel Prize Winner’s New Life,” *Esquire*, July.

Zucker, L., M. Darby and M. Brewer (1998). “Intellectual Human Capital and the Birth of U.S. Biotechnology Enterprises,” *American Economic Review*, 88: 290-306.

Figure 1: Exploratory search & innovation via knowledge bridging

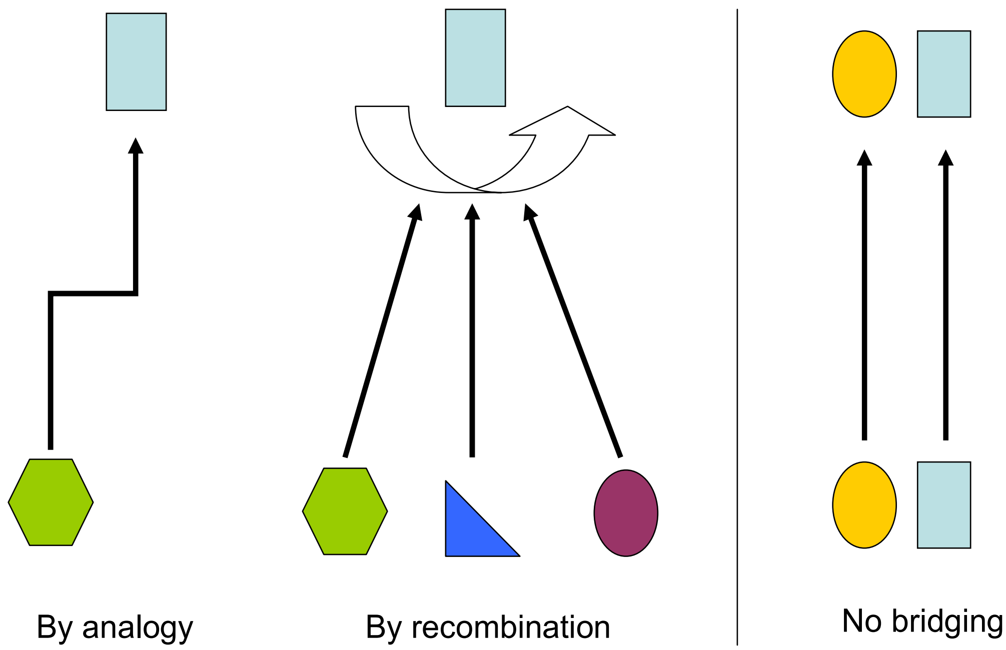

Figure 2: Referenced patent classes over time by firms in the sample

Figure 2a: Primary U.S. Patent Class Cited Each Year by New England Biolabs Patents

Primary	No of Citations (by Patent Application Year)															
US Class	1986	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002
257		1														
428			1													
435	2		11	35	39	1	26	49	19	6	19	30	35	34	20	8
514									1			1		1		
530									2			2		2		
536					1		2	4			1				2	

Figure 2b: Primary U.S. Patent Class Cited Each Year by Biogen Patents

Primary	No of Citations (by Patent Application Year)																						
US Class	1980	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
blank																				1	1		
100										1													
106										1													
210																			1				
340							1																
379					1																		
424	1	5		2	6		5		11	4	7	3	1	13	29		2	1	6	18	19		6
435	5	3	2	8	19	6	11	9	10	4	13	10	11	7	59	2	8	6	8	55	7	2	2
436					1				1		1	1		1	7		2	1					
504										1													
514					3	2		1	3	2	7	5	1	1	21	1	5	7	20	49	15		
521									1														
525																					1		
526																			1		1		1
530		1	3	2	5		8	5	8	2	7	6	7	7	27			1	3	1	9		2
536							1		4			1			5		4			1	1		
544																				2			
560															1								
588										1													
604														1									
623									1	1	1												
715											1												
800							1															1	

Figure 3: Modes of knowledge bridging by firms in the sample

Figure 3a: Citing versus Cited Patent Classes for Visys

Patent classes cited by Visys

Visys' patents, by class

Citing Class	Cited Class													
	14	15	19	22	31xa	32	33d	33e	33n	33o	39	43	44	61
14	11						6							
33d	31	4	19		1		58	8	2	3				1
44				1		1					1	10	15	

Figure 3b: Citing versus Cited Patent Classes for Therion

Patent classes cited by Therion

Therion's patents, by class

Citing Class	Cited Class					
	31i	31k	33d	33e	33f	33j
31i	1	2				2
31k		2				
33e						2
33f			1	2	1	2
33j						6
33k						1
33o						1

Figure 3c: Citing versus Cited Patent Classes for New England Biolabs

Patent classes cited by NE Biolabs

NE Biolabs' patents, by class

Citing Class	Cited Class													
	14	15	31xa	33d	33e	33f	33i	33k	33l	33n	33o	33p	46	69
14	1			3	5		6							
15		2	1	1	2		3			1				
33d				6	2	2	8	1	1				1	
33e	1			6	3		4			1				
33f				1		1	2							
33i	8	4	2	113	57	13	72			16	1	1		1
33n					2									

Table 1
Summary Statistics and Variable Definitions

VARIABLE	DEFINITION	MEAN	SD
Patent-level measures			
Patent forward citations within 5 years	Count of forward citations (excluding self-citations) within 5 years of patent issuance	2.43	3.65
Patent forward citations through 2004	Count of forward citations (excluding self-citations) as of December 2004	5.25	12.51
Patent generality	Concentration of forward-citing patent classes (see Jaffe and Trajtenberg, 2001), adjusted as per Hall (2005)	0.54	0.33
Patent knowledge bridging	1 – (backward citing patents in common with the primary patent class of the focal patent as a share of backward citations of the focal patent)	0.52	0.38
Patent originality	Concentration of backward-cited patent classes (see Jaffe and Trajtenberg, 2001), adjusted as per Hall (2005)	0.53	0.33
Patent scope	Number of primary classes assigned to the patent, evaluated as of December 2004	2.71	1.43
Inventor patent experience at other firms	Number of patents issued to focal patent's inventors when employed by other organizations as of the application date of the focal patent	6.63	11.92
Patent application year	Set of dummy variables = 1 for patent application years from 1979 through 2003		
Firm-level measures			
Firm knowledge overlap	Angle degree between firms' primary patent classes and those it cites in a given year, based on a vector dot product	-40.51	19.38
Founder overlap	Angle degree between firm founders' primary patent classes in firm's first three years (a stationary vector) and the firm's patent classes in a given year based on a vector dot product	-57.51	21.52
Employees (t-2)	Number of employees in time t-2	719.64	1287.59
Cumulative FDA drug approvals (t-2)	Cumulative number of FDA drug approvals in time t-2	2.22	4.55
Cumulative FDA drug terminations (t-2)	Cumulative number of FDA drug terminations in time t-2	3.03	6.45
Cumulative equity strategic alliances (t-2)	Cumulative number of equity-based strategic alliances in time t-2	1.05	1.52
Firm inventor mobility (t-2)	# of inventors who apply for patents at the focal firm and who also have prior experience patenting at another organization, as of t-2	2.11	4.47
Venture capital funding (t-2)	Dummy = 1 if the firm received venture capital funding, as of time t-2	0.76	0.43
Initial public offering (t-2)	Dummy = 1 if the firm went public (IPO), as of time t-2	0.50	0.50
Entry order	Sequential order of entry (1 is first).	9.84	5.51

Table 2
Patent Forward Citation Fixed Effects Negative Binomial Regressions
(Patent level of analysis)

	Dep. Var.: Patent forward citations within 5 years of patent issue (self citations excluded)			*Dep. Var.*: Patent forward citations through 2004 (self citations excluded)
Independent variables	**(2-1)**	**(2-2)**	**(2-3)**	**(2-4)**
Patent knowledge bridging	0.191*** (0.068)	0.191*** (0.068)		0.103** (0.056)
Patent originality			0.242*** (0.085)	
Patent scope		0.038*** (0.017)	0.058*** (0.019)	0.027** (0.014)
Inventor patent exp. at other firms		-0.000 (0.003)	0.001 (0.003)	-0.002 (0.003)
Patent application year fixed effects	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes
Constant	0.730 (0.735)	-0.596 (1.014)	-12.990 (528.556)	1.762*** (0.243)
Log likelihood	-3763.934	-3756.649	-3068.572	-6429.550
# observations	1887	1884	1525	2949

*, ** and *** denote statistical significance at the 10%, 5% and 1% level, respectively.

Table 3
Patent Generality Fixed Effects OLS Regressions
(Patent level of analysis)

	Dependent variable: Patent generality OLS regressions (standard errors are robust & clustered by firm)			
Independent variables	**(3-1)**	**(3-2)**	**(3-3)**	**(3-4)**
Patent knowledge bridging	0.187*** (0.035)	0.186*** (0.033)		
Patent originality			0.223*** (0.023)	0.222*** (0.026)
Patent scope		0.004 (0.004)		0.005 (0.004)
Forward citations through 2004		0.001*** (0.000)		0.001* (0.000)
Inventor patent exp. at other firms		-0.000 (0.001)		-0.000 (0.001)
Patent application year fixed effects	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes
Constant	0.576*** (0.008)	0.462*** (0.047)	0.521*** (0.013)	0.411*** (0.061)
R^2	0.12	0.13	0.13	0.13
# observations	1495	1494	1258	1257

** and *** denote statistical significance at the 5% and 1% level, respectively.

Table 4
Knowledge Overlap Firm Fixed Effects OLS Regressions
(Firm-year level of analysis)

	Dependent Variable = Firm knowledge overlap			
Independent Variables	**(4-1)**	**(4-2)**	**(4-3)**	**(4-4)**
Founder overlap	0.236***			0.269***
	(0.089)			(0.095)
Employees in t-2	0.003***			0.002
	(0.001)			(0.001)
Cumulative equity strategic alliances in t-2		4.043***		0.673
		(1.503)		(2.747)
Firm inventor mobility in t-2		0.538*		0.302
		(0.296)		(0.289)
Cumulative FDA drug dev. approvals in t-2			0.137	0.255
			(0.637)	(0.602)
Cumulative drug dev. terminations in t-2			0.649	0.494
			(0.406)	(0.395)
Venture capital funding in t-2			-18.006*	-26.242*
			(10.596)	(14.066)
Initial public offering in t-2			8.125	-0.707
			(5.319)	(14.066)
Firm fixed effects	Yes	Yes	Yes	Yes
Constant	-14.894**	-29.594	-27.514	-6.864
	(6.183)	(9.274)	(7.743)	(13.837)
R^2	0.16	0.15	0.16	0.21
# observations	184	312	312	184

*, ** and *** denote statistical significance at the 10%, 5% and 1% level, respectively.