LEARNING FROM WHAT OTHERS HAVE LEARNED FROM YOU: THE EFFECTS OF KNOWLEDGE SPILLOVERS ON ORIGINATING FIRMS

HONGYAN YANG
Hong Kong Polytechnic University

COREY PHELPS
HEC Paris

H. KEVIN STEENSMA
University of Washington, Seattle

Although research suggests knowledge spillovers often benefit imitators at the expense of originators, we investigate how originating firms may also benefit from their own spillovers. When an originating firm’s spillovers are recombined with complementary knowledge by recipient firms, a spillover knowledge pool is formed, containing opportunities for the originator to learn vicariously from recipients. In a longitudinal study of 87 telecommunications manufacturers, we found that a firm’s rate of innovation and the extent to which these innovations integrate knowledge from the spillover knowledge pool is greater when this pool is larger in size and similar to the firm’s knowledge base.

In the early 1980s, scientists at Eastman Kodak began their pioneering work on a molecule that eventually led to Kodak’s core innovation in organic light-emitting diode (OLED) technology in 1985. During the next 15 years, over 30 firms, including Sony and Xerox, exploited Kodak’s efforts by combining its core discovery with other complementary knowledge to generate additional innovations. Rather than depleting innovative opportunities and limiting Kodak’s ability to advance OLED technology, the innovative efforts of these recipient firms seem to have increased Kodak’s opportunities for innovation and enhanced its subsequent innovativeness. Since 1985, Kodak has developed additional innovations embodying OLED technology—most of which further build on the advances made by those firms that built on Kodak’s core OLED technology. We explore the proposition that Kodak learned vicariously from the efforts of other firms and exploited a pool of knowledge that was created when recipient firms combined Kodak’s original OLED invention with complementary knowledge.

The use of Kodak’s OLED technology by firms other than Kodak represents a knowledge spillover, a phenomenon that occurs when recipient firms (e.g., Sony, Xerox) use an originating firm’s (e.g., Kodak’s) knowledge in their innovative pursuits (Griliches, 1992). Because knowledge is partially a public good, knowledge spillovers are largely beyond the control of originating firms (Arrow, 1962; Mansfield, 1985). Knowledge produced by an originating firm may be used by recipient firms for little incremental cost, effectively reducing the costs of innovation for recipients (Arrow, 1962). Knowledge spillovers stimulate increasing returns to knowledge production and contribute to society by enhancing economic growth (Romer, 1990).

Although an extensive literature examines how knowledge spillovers benefit both society and recipient firms (e.g., Cohen & Levinthal, 1990; Griliches, 1992), little research has addressed the potential benefits of spillovers for originating firms. Except when examining cases in which firms’ intent was to create an industry standard (Spencer, 2003), researchers have generally assumed that originating firms have nothing to gain from knowledge spillovers (Kogut & Zander, 1992). Indeed, spillovers facilitate competitive entry into a tech-
knowledge spillovers and their recombination with stemming from increased competition. However, gains from learning necessarily outweigh the losses have learned from them. We do not claim that the originators can also learn from what other firms ability to profit from innovation, our results suggest late competition and hamper an originating firm's knowledge base. Although spillovers may stimu-
in size and more similar to the firm's existing greater when the spillover knowledge pool is larger knowledge from a spillover knowledge pool is the extent to which these innovations build on period. We find that a firm's rate of innovation and predic-
tions using data on a panel of 87 telecommu-
nications equipment manufacturers over a ten-year period. We find that a firm's rate of innovation and search, Organizational Learning, and Innovation

Innovation is a problem-solving process in which solutions to economically valuable problems are discovered via search (Dosi, 1988). The creation of new knowledge most often involves the novel recombination of existing elements of knowledge (Fleming, 2001), or the reconfiguration of the ways in which knowledge elements are linked (Henderson & Clark, 1990).1

Both experiential and heuristic search guide a firm's quest for valuable solutions. In experiential search, innovators draw on their existing knowledge (Gavetti & Levinthal, 2000). Recombination attempts involve altering one element of a known solution at a time and then observing the resulting change in performance. Firms learn from their own experience, myopic to others' search efforts and outcomes (Levinthal & March, 1993). Because "bounded rationality" biases individuals toward searching salient areas of their existing knowledge, firms generally exploit their existing knowledge in their innovation efforts (March, 1991). Exploiting existing knowledge through experiential search can lead to the development of efficient organizational routines (Nelson & Winter, 1982) and positive, timely, and predictable returns (March, 1991).

However, continually exploiting existing competencies can limit firms to incremental advancements and suboptimal solutions (Fleming, 2001). Integrating external knowledge into a search process can overcome these limitations by increasing the number and variety of knowledge components available for recombination, which increases the

1 The concept of recombination is an evolutionary metaphor (Nelson & Winter, 1982). The use of this metaphor is well established in studies of technological innovation (Basalla, 1988; Fleming, 2001; Nelson & Winter, 1982). This literature uniformly and abstractly characterizes knowledge as discrete elements or components that serve as the grist for the mill of innovation. This convention of terminology is independent of the variety of ways in which research has operationalized knowledge. In keeping with this research tradition, we adopt the terms "knowledge elements" and "knowledge components" and use them interchangeably.
number of combinatorial possibilities and potential solutions (Fleming, 2001; Rosenkopf & Nerkar, 2001). Despite these benefits, integrating external knowledge is more uncertain and costly, and less successful on average than deriving solutions from one’s existing knowledge base (March, 1991).

Heuristic search can reduce the uncertainty and costs of using external knowledge in a search process. Heuristic search occurs when members of a firm cognitively evaluate alternative knowledge components and combinations and assess their implications for solution performance. An innovator’s cognitive representation is used to identify potentially valuable combinations quickly, and the combinations are then investigated via experiential search (Gavetti & Levinthal, 2000). Because a firm can evaluate a solution without directly implementing it, heuristic search is cheaper than experiential search, reduces the risks of experimentation, and increases the efficiency of exploring external knowledge (Gavetti & Levinthal, 2000).

Learning vicariously from other firms is a type of heuristic search (Cyert & March, 1963). Organizations learn vicariously by observing the behavior and associated performance outcomes of other organizations and then modeling or imitating behaviors that seem successful and avoiding behaviors that seem unsuccessful (Cyert & March, 1963). By observing the innovative activities of other organizations and the outcomes of those activities, a firm can develop a cognitive model of how and why a new combination of knowledge is formed without attempting the combination. This cognitive model can be used as a guide for future solution search by identifying potentially valuable knowledge components and combinations, detecting elements and combinations to avoid, and providing insight into the organizational routines that led to the creation of the innovation (Sorenson et al., 2006). Thus, a firm’s innovation efforts, including its routines and the outcomes of these routines, can serve as templates for other firms’ innovative pursuits (Hoetker & Agarwal, 2007). In the next section, we suggest that an originating firm’s spillover knowledge pool provides viable opportunities for the originating firm to learn vicariously from the innovation efforts of recipient firms.

Knowledge Spillovers, Knowledge Pools, and the Spillover Knowledge Pool

A knowledge spillover is a flow of knowledge from an originating firm to a recipient firm. Knowledge has spilled over from an originating firm only when recipients use it in their innovation pursuits (Griliches, 1992). Although firms originate and receive knowledge flows concurrently, originators and recipients are conceptually distinct in a given spillover process.

All knowledge is not equally accessible to a firm (Jaffe, 1986). For example, a firm can more easily access and exploit knowledge that is developed by other firms in its industry as compared to knowledge developed by firms outside of its industry (Henderson & Cockburn, 1996). Similarly, a firm can more easily exploit a pool of external knowledge developed by other firms pursuing similar technologies (Jaffe, 1986) or located in the same geographic region as the firm (Jaffe, Trajtenberg, & Henderson, 1993). Rather than conceiving of knowledge pools as equally accessible by all firms within a boundary (i.e., an industry, a technological domain, a region), we envision an external knowledge pool that is unique and specific to each originating firm and bounded by the recombinatorial activities of firms that exploit its spillovers.

For recipient firms to exploit the knowledge of others, they often need to combine this knowledge with additional knowledge from their own idiosyncratic context (Sorenson et al., 2006). Because each firm’s knowledge context is unique, the manner in which recipient firms exploit knowledge spillovers varies among recipients and differs from the manner in which the originating firm exploited the original knowledge. Through the recombinatorial activity of recipient firms, the knowledge produced by an originating firm is linked directly to external knowledge.

Following prior research on technological innovation and recombinatorial search (e.g., Fleming, 2001), we characterize knowledge spillovers as discrete knowledge components. An originator’s spillover knowledge pool represents all external knowledge components that have been linked directly to its knowledge by recipient firms through spillover. Consider a simple example: An originating firm’s knowledge component $a$ spills over to recipient firms 2 and 3. Firm 2 combines component $a$ with knowledge components $b$ and $c$. Firm 3 combines knowledge component $a$ with components $d$ and $e$. By combining component $a$ with components $b$, $c$, $d$, and $e$, new knowledge combinations $g$ and $f$ are created. The boundary of the originating firm’s spillover knowledge pool is distinct, and the contents of the pool have been contributed by recipient firms. Knowledge components $b$, $c$, $d$, $e$, $g$, and $f$ constitute the knowledge in the originating firm’s spillover knowledge pool. Knowledge component $a$ is not part of the spillover knowledge pool, because the originating firm produced it and it is thus part of the originator’s knowledge base. Although firms in a similar technological domain or within the
same geographic region as the originator may have created other knowledge components, if this knowledge has not been connected to the originating firm’s knowledge through spillover, it is not contained in the originator’s spillover knowledge pool. Thus, in contrast to knowledge pools based on industry, technological domain, or geographic location, a spillover knowledge pool is unique to each firm.

In the case of Kodak’s OLED technology, recipient firms combined Kodak’s original organic light-emitting invention with knowledge outside of Kodak’s existing base of knowledge. In so doing, these recipients created a pool of external knowledge directly associated with Kodak’s original OLED invention. This knowledge pool was specific to Kodak and directly linked to its existing knowledge base through the innovative efforts of recipient firms.

For a firm to learn vicariously from the actions of others, such actions must be both salient and germane to it (Ingram, 2002). Relative to knowledge outside of an originating firm’s spillover knowledge pool, the combinations and knowledge components contained in that pool will be more salient and applicable to its innovation efforts. Researchers and engineers often become emotionally and intellectually attached to their work and actively monitor how their ideas diffuse and are extended by others (Garud & Rappa, 1994). Consequently, originating firms tend to observe the innovation efforts of recipient firms. Indeed, originating firms and their inventors often track the innovative efforts of other organizations—including those that fail—through social networks, technical publications, conferences, patent filings, and reverse-engineering efforts (Appleyard, 1996; Henderson & Clark, 1994). These channels facilitate both the flow of knowledge from originators to recipients and the flow of knowledge back to originators once recipients have combined the spillover with other knowledge.

For example, in an interview with us, a research scientist in a large electronics company said the following regarding the development of an electronic switch in his company:

It turns out that there are many different ways to do this [electronic switch]. Our company came up with a couple of good ideas at the beginning and built along those lines. Others have developed more than 8 or 10 or 12 different unique ideas around it. We have absolutely kept a careful eye on that scientific literature and subsequently developed upon some of the good ones.

Several interview subjects confirmed that they and their colleagues used a variety of sources, such as the Science Citation Index, conference presentations, journal publications, patent applications, and informal conversations, to track how their knowledge was exploited by others. A senior scientist at a government research lab stated:

When acquaintances in the field come up to me and say, “I saw that so and so is using your technology in their work” or something like that, I’d go to a conference meeting and hear them explain how to do that.

A nanotechnology researcher at a corporate research lab told us:

I have one paper from two years ago which has 50 citations now, and I look and I say, okay, those are the people I know, and then I see somebody that I don’t know who has cited it several times, and then I start looking at that work, and what have they done.

Finally, an R&D manager from an electronics company explained the motivation:

If it is a problem that we’re working on and if somebody else cites our article, or somebody else makes an advancement on our idea, we need to know what others are doing. That is the huge reservoir from which any research scientist draws new ideas and new developments.

Characteristics of Spillover Knowledge Pools and Their Effect on Originating Firms

Our core proposition is that a firm’s spillover knowledge pool influences its innovativeness by highlighting novel combinations of knowledge and the organizational routines of the recipient firms that generated the innovations in the spillover process. These novel combinations and routines represent templates, which can be learned vicariously and incorporated into the originating firm’s heuristic search. Because the spillover knowledge pool is the direct extension of the originating firm’s knowledge, the firm’s members can more easily understand and exploit knowledge in the pool than knowledge outside it, all else being equal. Indeed, external knowledge that is related in some fashion to a firm’s existing competences is more easily assimilated and exploited by the firm than is unrelated knowledge (Lane & Lubatkin, 1998). Two characteristics of a firm’s spillover knowledge pool—size and similarity— influence the firm’s innovativeness, in terms of both its rate of innovation and the extent to which its innovations build on its spillover knowledge pool.
**Pool size.** The spillover knowledge pool highlights new recombinatorial opportunities associated with an originating firm’s existing knowledge base. As opposed to exhausting recombinatorial opportunities, recombinations spawn ever more recombinatorial opportunities (Fleming & Sorenson, 2001). As the spillover knowledge pool of an originating firm grows larger in terms of knowledge components, the number of future recombinatorial opportunities associated with the firm’s knowledge base increases (Fleming & Sorenson, 2001).

In general, larger spillover knowledge pools provide greater potential for an originating firm to learn vicariously from the recombinatorial activity of other firms. Observing how recipient firms combine its knowledge with other knowledge components enhances the originating firm’s search routines, enabling it to more efficiently and effectively search both internal and external knowledge for possible solutions.

Learning vicariously from recipient firms improves an originating firm’s ability to exploit existing competencies. When recipients combine knowledge from the originator with other components, they may do so in ways the originator has not explored. Observing how recipient firms exploit the knowledge can assist the originating firm in identifying novel combinations of well-understood components or combinations to avoid. One interviewee described how a competitor had improved upon his firm’s original research and how these improvements influenced his innovative efforts:

> He basically advanced it beyond what we had done in very clever ways. . . . I have imported his technology because, you know, if somebody is good, you want to do what they are doing.

When recipient firms combine an originating firm’s knowledge with knowledge previously unfamiliar to the originating firm, such knowledge, in effect, becomes increasingly familiar to the originating firm. The innovation efforts of recipients can help an originator identify potentially promising knowledge and combinations, which can then be investigated more thoroughly via experiential search. Because the originator has access to a working template that incorporates some of its own knowledge, the uncertainty associated with integrating this novel knowledge with its own knowledge declines (Fleming & Sorenson, 2001).

For example, recipient firms exploited Kodak’s OLED technology by combining it with additional technology to enhance the color and duration of the molecule. The innovative activities of these recipient firms advanced the technological trajectory of OLED and also provided insight to Kodak as to how to further advance its original discovery.

The extent to which a firm learns from its spillover knowledge pool is evident in the knowledge embodied in its subsequent innovations. Although knowledge in a firm’s existing knowledge base is the most accessible to it and the easiest for it to exploit, utilizing only internal knowledge is limiting (Fleming, 2001). A firm’s spillover knowledge pool is a source of external knowledge that is more accessible and easier to exploit than external knowledge not in the pool. Larger spillover knowledge pools increase the efficiency of the originating firm’s innovative pursuits because it can draw upon a larger body of related external knowledge from which to build its innovations. Firms with larger spillover knowledge pools can rely to a greater extent on the knowledge in their spillover knowledge pools as opposed to having to search external knowledge sources that are less accessible and more costly to search. In contrast, firms with smaller spillover knowledge pools will be compelled to search less accessible external knowledge sources beyond their somewhat limited spillover knowledge pools. Because of the general bias toward exploiting knowledge sources that are relatively accessible and efficient to search, the larger a firm’s spillover knowledge pool, the greater the extent to which the firm’s subsequent innovations will build on knowledge from its spillover knowledge pool.

Learning from the spillover knowledge pool is also evident in a firm’s rate of innovation. An originating firm expands and refines its search routines by learning from its spillover knowledge pool. As the spillover knowledge pool grows larger, an increasing number of recombinatorial opportunities become salient for the originating firm. Because spillover knowledge pools are a particularly efficient source of recombinatorial opportunities for originating firms, larger pools lead to higher levels of innovative output, all else being equal.

**Hypothesis 1a.** The larger an originating firm’s spillover knowledge pool, the more the firm’s subsequent innovations integrate knowledge from the spillover knowledge pool.

**Hypothesis 1b.** The larger an originating firm’s spillover knowledge pool, the greater the firm’s subsequent innovative output.

**Pool similarity.** Beyond varying in size, spillover knowledge pools also vary in the similarity between the knowledge in the pool and the originating-
ing firm’s existing knowledge base. Similarity refers to the extent to which the knowledge in a firm’s spillover knowledge pool is located in the domains in which the originating firm has expertise. For some originating firms, the knowledge in their spillover pools is relatively different (i.e., distant) from that in their existing knowledge bases. For other originating firms, the knowledge in their spillover pools is more closely related (i.e., similar) to their existing knowledge bases. The similarity of the knowledge in the spillover knowledge pool to the originating firm’s existing knowledge influences the extent to which the vicarious learning opportunities associated with the knowledge pool are valuable and accessible to the originator. For example, a chief technology officer of a large manufacturing firm told us:

We constantly search what other organizations are doing with our technology in the technological areas that we are interested in. . . . Of course, we are generally interested in the state-of-the-art innovations in these technological areas.

When the knowledge in a spillover knowledge pool is similar to an originator’s existing knowledge, the novel combinations within the pool incorporate external knowledge from knowledge areas that are familiar to the originating firm. Such combinations are particularly salient to and easily understood by its members. Through its members observing the recombinatorial activities of recipient firms in domains of knowledge similar to the originating firm’s, the originator gains deeper understanding of its existing competencies and identifies new ways to exploit those competencies.

Many interviewees explained how the innovations of recipient firms in domains similar to their own enhanced their innovation. The vice president of R&D for a pharmaceutical firm stated:

Discoveries are made with our drugs . . . that are completely unexpected and that lead you in a totally different direction.

He then provided an example of how another organization showed that a drug developed by his company was effective against a similar disease and how this enhanced understanding of the compound at his firm:

That’s just something that there’s no way that you could have known by yourself, and it completely changed my thinking as a physician and as a scientist about the action of the molecule against which [drug name] acts.

In contrast, absorbing external knowledge that is dissimilar to existing competencies can be quite challenging and costly (Cohen & Levinthal, 1990; Lane & Lubatkin, 1998). Firms must expend greater effort and more resources to integrate dissimilar knowledge, often encountering diseconomies of scale in their innovation efforts (Ahuja & Lampert, 2001). Because originators do not have direct access to the routines of the recipient firms that helped produce the innovations, they find it even more difficult to learn vicariously from combinations that incorporate unfamiliar knowledge (Jensen & Szulanski, 2007). A digital communications researcher described his reaction to seeing how others extended his ideas to areas beyond his familiarity:

I always think, “Gee, I wish I had done that.” I never have the domain knowledge to sort of jump and move right over there.

The similarity of an originating firm’s spillover knowledge pool to its knowledge base influences both its rate of innovation and the extent to which these innovations integrate knowledge from the spillover knowledge pool. The more dissimilar the external knowledge is to the existing knowledge, the greater is the recombinatorial uncertainty and the less likely are successful recombinations (Fleming & Sorenson, 2001). Thus, spillover knowledge pools that are more similar to the existing knowledge of an originating firm lead to higher rates of innovation than pools that are less similar. Moreover, because exploiting external knowledge that is similar to existing knowledge is relatively efficient, an originating firm’s innovations build on knowledge from its spillover knowledge pool to a greater extent when the knowledge in the pool is similar to the originating firm’s existing knowledge base.

Hypothesis 2a. The more similar the knowledge in an originating firm’s spillover knowledge pool to its existing knowledge, the more its subsequent innovations integrate knowledge from the spillover knowledge pool.

Hypothesis 2b. The more similar the knowledge in an originating firm’s spillover knowledge pool to its existing knowledge, the greater its subsequent innovative output.

METHODS

To improve our understanding of the phenomenon and inform our theory development, we interviewed ten research scientists and R&D managers in the corporate, governmental, and academic sec-
tors. We used this evidence to inform, illustrate, and validate our theoretical arguments. To test our hypotheses, we identified and examined the knowledge spillovers and innovations generated by a sample of firms. We observed knowledge creation using the production of novel technological inventions with patents and used patent citations to an originating firm’s patents to assess knowledge spillovers. We thus built on a large body of research that has used patents as a proxy for firm innovation (see Hagedoorn & Cloodt, 2003) and numerous studies that have used patent citations as indicators of knowledge flows and spillovers (e.g., Almeida, 1996; Hoetker & Agarwal, 2007; Jaffe et al., 1993).

Although knowledge spillovers can occur through a host of mechanisms, including technical publications, conferences, employee mobility, social networks, and reverse-engineering efforts, patents and their citations represent observable knowledge flows, regardless of the diffusion mechanism (Jaffe et al., 1993; Jaffe, Trajtenberg, & Fogarty, 2002).

Patent data have many advantages for our study. First, knowledge creation is instantiated in inventions (Trajtenberg, 1990), which provide a trace of an organization’s knowledge creation activities. Patents provide a measure of novel invention that is externally validated through the patent examination process (Griliches, 1990). Because obtaining and maintaining patent protection is time-consuming and costly, patent applications represent an inventor’s positive expectation of the economic significance of an invention (Griliches, 1990). Although patents reflect a codifiable portion of a firm’s technological knowledge, they correlate with measures that incorporate tacit knowledge, such as expert ratings of the firm’s technical competency (Narin, Noma, & Perry, 1987) and the introduction of new products (Brouwer & Kleinknecht, 1999). Trajtenberg (1990) concluded that patents are perhaps the most valid and robust indicators of knowledge creation.

A second advantage is that patents contain citations to prior patents, which represent valid and reliable indicators of knowledge diffusion. These “prior art” citations represent the preexisting technological components that were combined in a novel way to yield the patented invention (Basalla, 1988). Patent applicants are required by law to include a list of relevant citations in their applications and have incentives to do so (Griliches, 1990). The patent examiner reviewing the application is ultimately responsible for the citations contained in a granted patent. Examiners often add citations to applications (Alcacer & Gittelman, 2006), suggesting that applicants are not necessarily aware of all cited patents. Although examiner-added citations may add noise to measuring spillovers, many studies have shown that patent citations are valid indicators of actual knowledge flows (cf. Duguet & MacGarvie, 2005; Jaffe et al., 2002).

Empirical Context and Sample

Both theoretical and practical considerations influenced our choice of empirical setting. First, the setting had to be technologically intensive, because such industries have higher rates of knowledge creation and knowledge diffusion. Second, the industry needed to demonstrate significant cumulative technological advances to permit us to observe sufficient knowledge spillovers. Third, because we used patent data for multiple measures, and systematic differences in the use of patents existed among industries (Levin, Klevorick, Nelson, & Winter, 1987), we needed to study an industry in which firms actively patented their inventions. Accordingly, we chose to sample firms from the global telecommunications equipment manufacturing industry (SIC 366).

Telecommunications equipment manufacturers produce and market hardware and software that enable the transmission, switching, and reception of voice, images, and data over both short and long distances using digital, analog, “wireline,” and wireless technology. This industry was an appropriate context for our study for three reasons. First, beginning in the early 1970s, the telecommunications equipment industry entered a period of rapid technological change. Average R&D intensity increased steadily between the late 1970s and 2000, and the industry has consistently been designated as “high technology” in Bureau of Labor Statistics studies (e.g., Hecker, 1999). Second, technological knowledge has diffused quickly and advances have accumulated: on average, patents associated with telecommunications equipment technologies diffuse more rapidly (are cited more quickly) than other technologies, are cited more often than other technologies, and contain a larger number of citations to other patents (Hall, Jaffe, & Trajtenberg, 2001). Finally, research has shown that telecom
equipment firms routinely patent their inventions (Griliches, 1990; Levin et al., 1987). Hagedoorn and Clootd (2003) found that patents were a particularly good measure of innovation in this industry.

Many practical considerations guided the construction of our sample. Because we needed to control for unobserved sources of firm differences in innovativeness, we required sufficient time-varying data on the same set of firms. To minimize “left- and right-censoring” regarding the collection of patent data and to ensure access to firm financial data, we limited the sample period to 1987–97. As we explain below, our independent variables require a ten-year window of patent data prior to each firm-year observation, requiring the collection of patents applied for in the mid 1970s (the beginning of our patent data sources). Furthermore, collecting financial data on many international firms prior to 1987 proved difficult. Given the lag between the application for a patent and its eventual granting, we ended the sample in 1997, allowing eight years to elapse between the end of the sample and the end of our patent data collection. Nearly all patent applications are decided upon by the U.S. Patent and Trademark Office (USPTO) within seven years of application (Hall et al., 2001).

We limited the sample frame to public firms to ensure the availability and reliability of financial data. To minimize survivor bias, we selected the final sample of 87 firms by rank-ordering them by industry sales at the beginning of the sample period.

Data Sources

We obtained U.S. patent data from Delphion and the NUS Patent Database for the period 1977–2005. Using patents from a single country maintained consistency, reliability, and comparability across firms (Griliches, 1990). U.S. patents are a very good data source because of the rigor and procedural fairness used in granting them, the strong incentives for firms to obtain patent protection in the world’s largest market, the high quality of services provided by the USPTO, and the U.S. reputation for providing effective intellectual property protection (Pavitt, 1988).

We took significant care in aggregating the patents of subsidiaries to the firm level. We initially identified all divisions, subsidiaries, and joint ventures of each firm in the sample (using Who Owns Whom and the Directory of Corporate Affiliations) as of 1976. We then traced each firm’s history to account for name changes, division names, divestments, acquisitions, and joint ventures to obtain information on the timing of these events. This process yielded a master list of entities that we used to identify all patents belonging to sample firms during the period of study.


Operationalizing the Spillover Knowledge Pool

We constructed a firm’s spillover knowledge pool in year \( t - 1 \) using patent citation data. To begin, we identified all patents applied for and assigned to firm \( i \) in the ten years prior to, but not including, year \( t - 1 \) (i.e., \( t - 2 \) to \( t - 11 \)). This procedure resulted in a list of patents for the focal firm, each identified by a unique number. Next, the universe of U.S. patents applied for in year \( t - 1 \) and subsequently granted) was identified from the NUS Patent Database. Third, all patents in this annual universe assigned to firm \( i \) were removed. Fourth, all remaining patents in this annual universe that contained a citation to any of firm \( i \)’s stock of patents were identified. This procedure yielded a list of firm \( i \)’s forward-citing patents (not owned by firm \( i \)) in year \( t - 1 \). We refer to this as the list of forward citing patents in year \( t - 1 \). Fifth, all patent citations contained in the patents identified in step four were identified. From this list of prior art (backward) citations, all citations to patents owned by firm \( i \) were removed. We refer to this as the list of backward-citing patents. To identify the spillover knowledge pool for firm \( i \) in year \( t - 1 \), we joined the lists of forward-citing and backward-citing patents and removed all redundant patent numbers. All patents contained in a firm’s spillover knowledge pool were unique patents from firms other than firm \( i \). These patents reflected the technological components that firm \( i \)’s patents were related to as a result of recipient firms’ recombinatorial efforts.

To illustrate our operationalization of the spillover knowledge pool, consider the following example shown in Figure 1. Assume that only one patent in firm \( i \)’s ten-year patent stock, patent \( a \), is cited by two other firms (\( j \) and \( k \)) in their new patents \( b \) and \( c \), respectively. Patents \( b \) and \( c \) represent the forward citations of patent \( a \). Also assume that patents \( b \) and \( c \) cite, in addition to patent \( a \), three other patents: \( d \), \( e \), and \( f \), which are the other backward citations of \( b \) and \( c \). Patents \( b \), \( c \), \( d \), \( e \), and \( f \) form the spillover knowledge pool of firm \( i \) at year \( t - 1 \).
**Dependent Variables**

**Innovative output**. We measured innovative output as the number of successful patent applications for firm \( i \) in year \( t \). Patents were counted in the year of application to capture the precise timing of knowledge creation (Griliches, 1990).

**Knowledge integration**. We measured the extent to which the innovative output of an originating firm built on the knowledge from its spillover knowledge pool as the proportion of prior art patents contained in firm \( i \)'s patents of year \( t \) that belonged to its spillover knowledge pool in year \( t-1 \). Because this measure is a share rather than a count of citations, it captures a firm’s propensity to build on knowledge in its spillover knowledge pool. The extent to which a firm uses elements of knowledge (e.g., patents) contained in its spillover knowledge pool reflects that it is searching and exploiting this knowledge pool.

**Independent Variables**

**Pool size**. This variable was measured as the total number of unique patents in firm \( i \)'s spillover knowledge pool at year \( t-1 \). This variable was log-transformed because of skewness.

**Pool similarity**. An index developed by Jaffe (1986) was used. For each firm-year, we constructed an index that measures the distribution of a firm’s patents across primary patent classes and the distribution of a firm’s spillover knowledge pool across primary classes. We used a moving ten-year window to establish each firm's patenting profile. This distribution locates the firm in a multidimensional technology space, captured by a \( J \)-dimensional vector \( D_i \) = \( \{d_{i1}, \ldots, d_{ij}\} \), where \( d_{ij} \) represents the fraction of firm \( i \)'s patents that are in patent class \( j \). The assumption underlying this approach is that the distribution of a firm’s patents across patent classes reflects the underlying distribution of its technological knowledge (Jaffe, 1986). The second distribution locates a firm’s spillover knowledge pool in a multidimensional technology space, captured by a \( J \)-dimensional vector \( E_i \) = \( \{e_{i1}, \ldots, e_{ij}\} \), where \( e_{ij} \) represents the fraction of the patents contained in firm \( i \)'s spillover knowledge pool that are in patent class \( j \). The similarity of firm \( i \)'s spillover knowledge pool in year \( t-1 \) was calculated as:

\[
\text{Pool similarity}_{i,t-1} = \left( \sum_{j=1}^{J} d_{ij} e_{ij} \right)^{1/2} \left( \sum_{j=1}^{J} d_{ij}^2 \right)^{1/2} \left( \sum_{j=1}^{J} e_{ij}^2 \right)^{1/2}.
\]

This measure is bounded between 0 and 1, with larger values representing increasing similarity.

**Control Variables**

**Technological opportunity**. “Technological opportunity” refers to differences in the set of possibilities for technological advances that exist within technologies and industries over time (Klevorick et al., 1995). Thus, some firms may be active in richer technological areas than other firms. Drawing on Patel and Pavitt (1997), we controlled for firm-specific differences in technological opportunity in year \( t-1 \) as follows:

\[
\text{Technological opportunity}_{i,t-1} = \sum_{j=1}^{J} \left( \text{patents}_{i,t-1} \times P_{j,t-1} \right),
\]

where \( \text{patents}_{i,t-1} \) refers to the number of patents granted in patent class \( j \) in year \( t-1 \), and \( P_{j,t-1} \) is...
the proportion of firm $i$’s patents applied for in class $j$ in year $t - 1$. The number of patents granted in a patent class in a given year is a proxy for the underlying rate of technical change in that area of technology (Patel & Pavitt, 1997). We divided this variable by 1,000 to reduce its scale.

**Number of recipient firms**$_{it} - 1$. Spillovers facilitate entry into specific technological domains (Jaffe, 1986), which can lead to crowded areas of innovative activity. The extent to which an originating firm competes in crowded technological domains may increase its incentives to innovate and lead to greater innovation rates (Stuart, 1999). We controlled for this potential confound using a variable that indexes the number of unique companies whose patents, applied for in year $t - 1$, cite firm $i$’s ten-year stock of patents. We divided this variable by 1,000 to reduce its scale.

**Patent stock quality**$_{it} - 1$. The number of (forward) citations a patent receives is a valid indicator of its value or quality (Trajtenberg, 1990). Firms that produce valuable patents are at greater risk of having larger spillover knowledge pools. To control for a firm’s patent stock quality, we identified the patents it applied for and was granted in the ten years prior to and including year $t - 1$. We then identified all forward citations this stock of patents received by 2005. We computed firm $i$’s patent stock quality in year $t - 1$ as the number of forward citations. We divided this variable by 1,000 to reduce its scale.

**Firm size**$_{it} - 1$. Prior research has proved inconclusive in determining whether small or large firms are more innovative, perhaps because firm size has both negative and positive effects on innovation (Teece, 1992). We controlled for the influence of firm size using firm $i$’s sales in billion $\text{US}$ in year $t - 1$.

**R&D**$_{it} - 1$. A firm’s R&D expenditures are investments in knowledge creation (Griliches, 1990) and contribute to its ability to absorb external knowledge (Cohen & Levinthal, 1990). We controlled for the influence of R&D expenditure in billion $\text{US}$ for firm $i$ in year $t - 1$.

**Current ratio**$_{it} - 1$. The availability of slack resources can increase exploratory search and lead to greater innovative performance (Nohria & Gulati, 1996). We controlled for the slack resources of firm $i$ in year $t - 1$ using its current ratio, calculated as current assets over current liabilities.

**Firm technological diversity**$_{it} - 1$. Technologically diverse firms may be more innovative (Garcia-Vega, 2006) and more able to absorb external knowledge (Cohen & Levinthal 1990). We measured technological diversity using Hall’s (2005) adjusted Herfindahl index:

$$\text{Firm technological diversity}_{it} - 1 = \left[ 1 - \sum_{j=1}^{I} \frac{N_{jit}}{N_{it-1}} \right] \times \frac{N_{it-1}}{N_{it-1} - 1},$$

where $N_{it-1}$ is the number of patents in firm $i$’s knowledge base at year $t$. $N_{jit} - 1$ is the number of patents in primary technology class $j$ in firm $i$’s knowledge base at year $t - 1$. This variable may take on values from 0 (no diversity) to 1 (maximum diversity).

**Estimation**

We employed two dependent variables and estimated our models using panel regression methods appropriate to each dependent variable. All explanatory and control variables were lagged one year, which accounted for the delay in converting innovation inputs into outputs, reduced concerns about reverse causality, and avoided simultaneity. The panel was unbalanced as some firms were acquired or restructured, making within-firm comparisons difficult. Our first dependent variable, innovative output, was a count variable that could take on only nonnegative integer values. The use of linear regression to model such data can result in inefficient, inconsistent, and biased coefficient estimates (Long, 1997). Although Poisson regression is appropriate to model count data, our data were significantly overdispersed, violating a basic assumption of the Poisson estimator (Hausman, Hall, & Griliches, 1984). Thus, we used negative binomial regression to model the count data (Hausman et al., 1984). The negative binomial model is a generalization of the Poisson model and allows for overdispersion by incorporating an individual, unobserved effect into the conditional mean (Hausman et al., 1984). We included year dummies to control for unobserved systematic period effects. We also employed firm fixed effects to control for unobserved, temporally stable firm differences in innovation. We used Allison and Waterman’s (2002) unconditional fixed effects estimator rather than the more conventional conditional maximum likelihood estimation procedure developed by Hausman et al. (1984). However, our results do not

---

3 Allison and Waterman (2002) criticized Hausman et al.’s (1984) conditional negative binomial fixed effects model as not being a “true” fixed effects method because it does not control for all time-invariant sources of heterogeneity. As a result, it is possible to estimate coefficients for time-invariant variables when using the Hausman et al. fixed effects estimator. This is not possible
differ substantively for these two estimation approaches. We used fixed rather than random effects because Hausman tests indicated rejection of the random effects specification for the models estimated below. Although conventional standard errors are biased downward when using unconditional fixed effects, this bias is effectively corrected by multiplying standard errors by the square root of the deviance statistic divided by its degrees of freedom (Allison & Waterman, 2002). We implemented that correction in all reported models.

The second dependent variable, knowledge integration, was a proportion. Estimation involving a proportional dependent variable presents several challenges to linear regression (Greene, 1997). Following standard econometric practice (Greene, 1997), we transformed this variable using a logit (i.e., log odds) transformation.4 We estimated our models using panel linear regression with firm fixed effects and year dummies and employed robust standard errors.

RESULTS

Table 1 provides descriptive statistics and correlations for all variables. Table 2 reports the regression results.5 We ran similar models for both dependent variables. Models 1 and 4 include only control variables. Models 2 and 5 introduce pool size and pool similarity. To explore the possibility of diminishing returns, we included the squared terms of pool size and pool similarity in models 3 and 6. Although they are not reported, year and firm dummies are included in all models. Hausman tests (1978) for all reported models were significant, suggesting that the fixed effects estimator was more appropriate than random effects.

Hypotheses 1a and 1b propose positive relationships between pool size and our two dependent variables. The effect of pool size is significant and positive for both innovative output (p < .01) and knowledge integration (p < .01). Thus, Hypotheses 1a and 1b are supported.

Hypotheses 2a and 2b propose that the similarity between a spillover knowledge pool and an originating firm’s existing knowledge base has a positive influence on both innovative output and knowledge integration. The effect of pool similarity is significant (p < .01) in all models for both dependent variables. Thus, Hypotheses 2a and 2b are supported.

To test the robustness of our findings, we tested for diminishing or negative returns to pool size and pool similarity by including squared terms of these variables in models 3 and 6. The squared term for pool size is not significant in model 3 in its effect on innovative output. The squared term for pool size is negative and marginally significant (p < .1) in model 6, suggesting an inverted U-shaped effect of pool size on knowledge integration. The coefficient

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>s.d.</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Knowledge integration</td>
<td>0.25</td>
<td>0.28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Innovative output</td>
<td>143.61</td>
<td>311.04</td>
<td>.37**</td>
<td>.57**</td>
<td>.60**</td>
<td>.62**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Pool size</td>
<td>5.87</td>
<td>3.02</td>
<td>.38**</td>
<td>.23**</td>
<td>.38**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Pool similarity</td>
<td>0.71</td>
<td>0.21</td>
<td>.38**</td>
<td>.23**</td>
<td>.38**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Technological opportunity</td>
<td>0.82</td>
<td>0.45</td>
<td>.38**</td>
<td>.14**</td>
<td>.34**</td>
<td>.23**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Number of recipients</td>
<td>0.19</td>
<td>0.38</td>
<td>.42**</td>
<td>.89**</td>
<td>.65**</td>
<td>.23**</td>
<td>.17**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Value of patent stock</td>
<td>6.22</td>
<td>14.00</td>
<td>.40**</td>
<td>.92**</td>
<td>.59**</td>
<td>.21**</td>
<td>.14**</td>
<td>.95**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Firm size</td>
<td>7.40</td>
<td>15.07</td>
<td>.36**</td>
<td>.78**</td>
<td>.60**</td>
<td>.17**</td>
<td>.13</td>
<td>.86**</td>
<td>.82**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. R&amp;D</td>
<td>0.47</td>
<td>0.97</td>
<td>.35**</td>
<td>.85</td>
<td>.60**</td>
<td>.21**</td>
<td>.12**</td>
<td>.87**</td>
<td>.83**</td>
<td>.93**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Current ratio</td>
<td>2.33</td>
<td>1.50</td>
<td>.08**</td>
<td>.23**</td>
<td>.30**</td>
<td>.06</td>
<td>.04</td>
<td>.23**</td>
<td>.22**</td>
<td>.25**</td>
<td>.21**</td>
<td></td>
</tr>
<tr>
<td>11. Firm technological diversity</td>
<td>0.76</td>
<td>0.31</td>
<td>.11**</td>
<td>.30**</td>
<td>.54**</td>
<td>.06</td>
<td>.09**</td>
<td>.32**</td>
<td>.29**</td>
<td>.32**</td>
<td>.32**</td>
<td>.38**</td>
</tr>
</tbody>
</table>

4 The transformed variable is \( \ln(knowledge\ integration / 1 - knowledge\ integration) \). Because the transformation is undefined when knowledge integration is equal to 0 or 1, we recoded these values as 0 = 0.0001 and 1 = 0.9999.

5 The average variance inflation factor for models 3 and 6 was 5.14, indicating that multicollinearity was not an issue.
The coefficients in model 6 show that the point at which knowledge integration is maximized is when pool size is 4.53, which is within the observed range of this variable. The square of pool similarity has a positive and significant influence on both innovative output and knowledge integration ($p < .01$ and $p < .05$, respectively). Given the positive and significant effect of both the first- and second-order terms, spillover knowledge pool similarity appears to have an exponentially increasing influence on a firm’s innovation.

To test the sensitivity of our results to our choice of using a ten-year window to construct a firm’s spillover knowledge pool, we reconstructed each knowledge pool variable using a five-year window. The results using this alternative window were consistent with those in Table 2.

**Supplementary Analysis**

The firm-level analysis provided evidence consistent with our basic premise that originating firms can understand and exploit knowledge from their spillover knowledge pools more easily than they can understand and exploit knowledge from outside those pools. To further investigate our basic premise, we used an experimental design at the patent level of analysis by employing the patent case control method (Jaffe et al., 1993). Empirically, we wanted to examine whether a firm more often used patents that were part of its spillover knowledge pool in its subsequent innovation efforts (i.e., whether it cited these patents more frequently) than patents with similar observable characteristics that were not contained in the spillover knowledge pool. If originating firms cited randomly sampled patents in the treatment group that were from their spillover knowledge pools more frequently than comparable untreated control patents, then we would have additional evidence that firms tend to exploit knowledge contained in their spillover knowledge pools more than knowledge from outside the pools.

To construct the sample for this analysis, we randomly sampled 1,000 patents from the spillover knowledge pools of our sample firms measured in 1987. Choosing this year provided the maximum time within which to observe subsequent citations to these patents. We stratified this sample on the basis of the

**TABLE 2**

Results of Negative Binomial Regression Analyses*

<table>
<thead>
<tr>
<th>Variables</th>
<th>Innovative Output</th>
<th>Knowledge Integration</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model 1</td>
<td>Model 2</td>
</tr>
<tr>
<td><strong>Independent</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pool size</td>
<td>0.20** (0.04)</td>
<td>0.23** (0.04)</td>
</tr>
<tr>
<td>Pool similarity</td>
<td>0.70** (0.21)</td>
<td>1.04** (0.24)</td>
</tr>
<tr>
<td>Pool size squared</td>
<td></td>
<td>−0.01 (0.01)</td>
</tr>
<tr>
<td>Pool similarity squared</td>
<td>2.19** (0.73)</td>
<td></td>
</tr>
<tr>
<td><strong>Control</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technological opportunity</td>
<td>0.19 (0.13)</td>
<td>0.06 (0.12)</td>
</tr>
<tr>
<td>Number of recipients</td>
<td>−0.27** (0.05)</td>
<td>−0.22** (0.05)</td>
</tr>
<tr>
<td>Patent stock quality</td>
<td>0.04** (0.01)</td>
<td>0.03** (0.01)</td>
</tr>
<tr>
<td>Firm size</td>
<td>−0.02† (0.01)</td>
<td>−0.02† (0.01)</td>
</tr>
<tr>
<td>R&amp;D</td>
<td>0.61** (0.11)</td>
<td>0.56** (0.10)</td>
</tr>
<tr>
<td>Current ratio</td>
<td>0.06† (0.04)</td>
<td>0.04 (0.03)</td>
</tr>
<tr>
<td>Firm technological diversity</td>
<td>1.09** (0.37)</td>
<td>0.71† (0.7)</td>
</tr>
<tr>
<td>Constant</td>
<td>4.08** (0.48)</td>
<td>3.79** (0.47)</td>
</tr>
<tr>
<td>Year dummies</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Firm dummies</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>df</td>
<td>636</td>
<td>634</td>
</tr>
<tr>
<td>$R^2$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scaled $\chi^2$</td>
<td>577.99**</td>
<td>598.19**</td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>516,632.38</td>
<td>497,901.65</td>
</tr>
</tbody>
</table>

* $n = 739$. Standard errors are in parentheses. They are adjusted by the square root of the deviance statistic (Allison & Waterman, 2002).
† $p < .10$
* $p < .05$
** $p < .01$
Two-tailed test.
size of each firm’s spillover knowledge pool in 1987. We refer to these patents as “treatment patents.”

Following Jaffe et al. (1993), we matched each of these treatment patents to a similar control patent according to the primary three-digit technology class and year of issuance of the treatment patent. Thus, a control patent was issued in the same year and assigned to the same primary technology class as its counterpart treatment patent, but it was not contained in the firm’s spillover knowledge pool or existing knowledge base during the period of observation. When multiple patents met these criteria for a control patent, we randomly selected one for inclusion in the sample. By matching control patents to treatment patents in this way, we attempted to minimize unobserved heterogeneity associated with the patenting process and control for any other bias that might arise from differences in technological domain or vintage. The final sample for our supplementary analysis consisted of 1,000 treatment and 1,000 control patents.

The dependent variable for this analysis, forward cites, was the count of forward patent citations received by focal patent \(i\) from focal firm \(j\) in year \(t\). Forward citations were observed for the years 1988−97. The explanatory variable of interest was a time-invariant dummy variable, \(SKP_i\), which took on the value 1 when the focal patent was randomly selected from the focal firm’s spillover knowledge pool and 0 if it was part of the control sample. To account for differences in the underlying quality of each patented invention in the sample, we included three control variables that have been shown to be valid proxies for invention quality (Lanjouw & Schankerman, 2004). The claims section of a patent delineates the scope or breadth of property rights it protects (Lanjouw & Schankerman, 2004). Patents with broader claims provide patent owners with more general protection in the use of their inventions, leading to enhanced patent value (Reitzig, 2003). We controlled for the number of claims contained in patent \(i\) (\(\text{claims}_i\)). Patents that cite a larger number of prior art patents may contain less novelty and be less valuable (Harhoff, Scherer, & Vopel, 2003). We controlled for the number of backward patent citations contained in patent \(i\) (\(\text{backward cites}_i\)). Finally, the number of forward citations a patent receives is a good indicator of the value or quality of the patented invention (Trajtenberg, 1990). We controlled for the cumulative number of forward citations patent \(i\) had received as of year \(t\), excluding those made by focal firm \(j\) in that year (\(\text{cumulative forward cites}_{ji}\)). Patents that build on a wide range of technologies may be more novel and therefore more highly cited (Trajtenberg et al., 1997). We controlled for the originality of patent \(i\) (\(\text{originality}_i\)) using Trajtenberg et al.’s (1997) originality measure, which is calculated as \(1 - \sum S_{ij}^2\), where \(s_{ij}\) refers to the fraction of patents cited by patent \(i\) that belong to technology class \(j\). To account for differences in the time in which a patent had been at risk of forward citation, we controlled for the age of patent \(i\) in year \(t\) (\(\text{age}_{it}\)).

The unit of analysis was the patent. Because we had a balanced panel and the dependent variable was a count, we used panel negative binomial regression. Our data were also hierarchically nested. In addition to the fact that multiple observations were nested within the same patent (i.e., the same patent was observed over time), multiple patents were also nested within single firms. To accommodate the multilevel nature of our data, we estimated a mixed-level negative binomial regression model in which patents were assigned random effects and firms were assigned fixed effects. This procedure allowed us to control for unobserved heterogeneity at both the patent and firm levels and to compute standard errors for patent-level coefficients that were free from the influence of spatial autocorrelation caused by patents being clustered within firms. We also accounted for the possibility that a patent’s value and subsequent citation differed according to its area of technology by including fixed effects for each patent’s technological category as defined by the

### Table 3

<table>
<thead>
<tr>
<th>Variables</th>
<th>Mean</th>
<th>s.d.</th>
<th>Minimum</th>
<th>Maximum</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Forward cites</td>
<td>0.02</td>
<td>0.14</td>
<td>0</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Spillover knowledge pool</td>
<td>0.50</td>
<td>0.50</td>
<td>0</td>
<td>1</td>
<td>.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Claims</td>
<td>12.10</td>
<td>9.92</td>
<td>0</td>
<td>94</td>
<td>.02</td>
<td>.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Backward cites</td>
<td>7.31</td>
<td>7.32</td>
<td>0</td>
<td>145</td>
<td>.03</td>
<td>.19</td>
<td>.17</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Cumulative forward cites</td>
<td>10.61</td>
<td>15.45</td>
<td>0</td>
<td>390</td>
<td>.11</td>
<td>.25</td>
<td>.18</td>
<td>.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Originality</td>
<td>0.37</td>
<td>0.28</td>
<td>0</td>
<td>0.92</td>
<td>.01</td>
<td>.11</td>
<td>.10</td>
<td>.35</td>
<td>.10</td>
<td></td>
</tr>
<tr>
<td>7. Age</td>
<td>8.29</td>
<td>4.58</td>
<td>0</td>
<td>22</td>
<td>.03</td>
<td>.00</td>
<td>.04</td>
<td>.10</td>
<td>.07</td>
<td>.03</td>
</tr>
</tbody>
</table>
NBER patent database (Hall, Jaffe, & Trajtenberg, 2001). Finally, we included year dummies in our estimations.

Table 3 presents descriptive statistics and correlations. Table 4 provides the results of the negative binomial regression analysis. The sole explanatory variable, spillover knowledge pool $i$ (SKP$_i$), is positive and significant ($p < .01$). The incident-rate ratio (exp$\beta$) for this variable is 2.05, which indicates that a one-unit change in SKP$_i$ (i.e., a patent changes from being outside of a firm’s spillover knowledge pool to being part of the pool) yields a 105 percent increase in the rate at which an average firm will cite the patent. Thus, we found strong support for our core theoretical claim that an originating firm has an efficiency advantage in that it can more easily understand and exploit knowledge components in its spillover knowledge pool than knowledge outside of the pool.

### DISCUSSION

The possibility that knowledge spillovers provide some benefit to the firms with which such spillovers originate has been largely unexplored. Indeed, when considered from the perspective of an originating firm, knowledge spillovers are typically viewed in a negative light. We investigated the conditions under which knowledge spillovers enhance an originating firm’s innovativeness. This study is the first of which we are aware to investigate this question.

We argued that when recipient firms recombine an originating firm’s knowledge with external knowledge, a pool of knowledge is formed that is different from the originator’s knowledge base, yet linked to it through the spillover process. We developed the concept of the spillover knowledge pool to denote this firm-specific pool of related knowledge. The spillover knowledge pool represents viable opportunities for originating firms to learn vicariously from recipient firms. By observing the recombinatorial actions of recipient firms and the outcomes of those actions, originating firms are able to refine and expand their search routines. We maintain that originating firms can more easily understand and exploit knowledge derived from spillover knowledge pools as compared to knowledge that is outside of these pools.

We found that the size of a firm’s spillover knowledge pool positively influences its innovativeness and the extent to which its innovations integrate knowledge from the pool. In general, larger spillover knowledge pools provide greater opportunity for originating firms to learn vicariously from the recombinatorial activity of recipient firms. Larger pools contribute a greater number of relatively accessible external knowledge components that can be used as input in the innovation process. However, we also found marginally significant evidence of an inverted U-shaped effect of the size of a firm’s spillover knowledge pool on the extent to which it builds on knowledge from the pool. This finding suggests there may be diminishing and ultimately negative returns to pool size for knowledge integration. Although larger pools provide originating firms with a large number of novel templates and learning opportunities, larger pools also increase the number and complexity of interdependencies and interactions among components. Attending to and understanding this increasing complexity may exceed the limited cognitive resources firms can dedicate to heuristic search (March, 1991). Consequently, as the size of a firm’s spillover knowledge pool begins to exhaust the firm’s cognitive resources, it may shift its focus to exploiting its own knowledge base because doing so is more efficient.

We also found that a firm’s rate of innovation and the extent to which these innovations build on and
integrate knowledge from the spillover knowledge pool are greater when this pool is similar to the firm’s existing knowledge base. For those originating firms that have substantial familiarity with the areas of knowledge contained in their spillover knowledge pools, such knowledge is particularly salient and accessible, reducing the uncertainty of integrating it into subsequent innovative efforts. Indeed, we found evidence of an exponentially increasing effect of similarity on our two measures of originating firm innovation. In other words, as dissimilarity increases, its negative effect on originating firm innovation and knowledge integration accelerates. We believe these effects are consistent with our theory. Because firms must expend greater effort and resources to utilize dissimilar knowledge, and because originators do not necessarily have direct access to recipient firms’ routines and the individuals who helped produce the innovations, originators should find it increasingly difficult and costly to learn vicariously from combinations that incorporate dissimilar knowledge. Thus, an originating firm should find it increasingly efficient to exploit its spillover knowledge pool as the pool becomes more similar. Taken together, our theory and results suggest that the costs to an originating firm of exploiting its spillover knowledge pool accelerate with the dissimilarity of the pool, while the originator’s innovative productivity decelerates as dissimilarity increases.

Our results suggest originating firms can benefit from knowledge spillovers by learning from what others have learned from them. As knowledge spills forward and is used by recipient firms, useful knowledge can also spill back to the originator.

Implications

This study makes several contributions to understanding of firm innovation and its implications for economic growth. The spillover knowledge pool represents a novel means of demarcating a stock of knowledge. This conception of a knowledge pool is more refined than those based on industry, technology, or geography and more tightly linked to a focal firm. Prior research suggests the extent to which a firm has access to and benefits from knowledge spillovers is limited to the knowledge produced by other organizations that compete in the same industry (Henderson & Cockburn, 1996), pursue similar technologies (Jaffe, 1986), or are located in the same geographic region (Jaffe et al., 1993). Researchers using these approaches have assumed the availability of spillovers is homogeneous for firms within a particular boundary. In contrast, our study suggests that the pool of available spillovers is firm-specific.

Our study also contributes to research on absorptive capacity. The concept of a spillover knowledge pool exploits the notion that firms possess absorptive capacity in an absolute, not a relative, sense (Lane & Lubatkin, 1998). A firm’s absorptive capacity is a function of the “relatedness” between its knowledge stock and external knowledge sources (Lane & Lubatkin, 1998). Although the distinction between what is and is not related is often undefined in absorptive capacity research (Lane, Koka, & Pathak, 2006), we conceptualize one means by which external knowledge can be viewed as being “related” to a firm’s existing knowledge stock. Our analysis suggests that the domain in which a firm has absorptive capacity is partly circumscribed by its spillover knowledge pool.

Knowledge spillovers also play a central role in endogenous growth theory by providing for increasing returns to knowledge production (Romer, 1990). Our results suggest that economic growth may be enhanced not just by the benefits of spillovers for recipient firms but also by their innovation-enhancing effects for originating firms.

We provide a novel explanation as to why firms innovate despite the disincentive associated with spillovers. Research suggests that spillovers reduce a firm’s ability to appropriate the returns from investments in knowledge production (Levin & Reiss, 1988). However, this static efficiency argument ignores the dynamic efficiency effect of spillovers that we identified in this study. Spillovers result in the creation of a firm-specific pool of related external knowledge, which an originating firm can more easily absorb than knowledge not in the pool. The extent to which the originating firm can reabsorb its spilled knowledge, including the advances made by recipient firms, reduces its total costs of innovation and increases the private returns from the original spillovers. Thus, from a dynamic perspective, knowledge that spills over may generate future private returns to the originating firm. On the margin, this benefit may mitigate the disincentive effect of spillovers for originating firms and increase their incentives to invest in innovation. Belenzon (2006) provided empirical support for this idea.

This dynamic incentive effect of knowledge spillovers addresses a fundamental tension in research on the economics of innovation and endogenous growth. On the one hand, stronger appropriability conditions increase a firm’s private incentives to invest in innovation production. On
the other hand, weaker appropriability conditions increase the volume of spillovers, reducing the costs and increasing the efficiency of innovation for recipient firms (Klevorick et al., 1995). The concept of the spillover knowledge pool and the results of this study potentially help to reconcile the conflict between these incentive and efficiency effects of appropriability. A spillover knowledge pool emerges as the result of an originating firm’s knowledge spilling over to recipient firms. The extent to which recipients recombine these knowledge components in novel ways and with other novel components provides originating firms with useful vicarious learning opportunities and thereby increases the technological opportunities available to them. As the level of technological opportunity increases in an industry or a domain of technology, the rate of innovation in such areas increases (Klevorick et al., 1995). Our study suggests that an originating firm’s own spillovers can dynamically increase its technological opportunities and innovative efficiency and also increase its incentives to invest in innovation. Thus, the efficiency and incentive effects of appropriability are not necessarily mutually exclusive.

Our results should not be interpreted as a prescription for encouraging spillovers. The losses that occur from spillovers because competition increases may outweigh the benefits of learning from the resulting spillover knowledge pool. However, our analysis does suggest that managers should not view the knowledge that does spill over simply as a loss, but rather as an opportunity. Managers should consider systems and processes to enhance awareness of this developing knowledge pool and to assimilate and utilize such knowledge in future innovation efforts. The monitoring of recipient firms’ use of technology can be conducted through patent documents, published and working papers, conference papers, and informal communications among individuals. Only if an originating firm has a comprehensive view of its spillover knowledge pool can it fully exploit the embedded recombinatorial opportunities and adapt its search routines accordingly. For example, the head of R&D for a leading biotech firm described his firm’s use of formal information management tools as follows:

We have, of course, a whole group in Library Science that does nothing but review the adequacy of our current [information search] tools and similarly our scientists are always trying to push for . . . and some are inventing other ways to search for information.

Limitations and Future Research

Although promising, our study has limitations. Because we used patent citations to assess knowledge spillovers, we did not capture all knowledge flows and may have counted citations that did not correspond to actual spillovers as relevant evidence. Because citations added by patent examiners typically reflect inventors’ ignorance of these cited patents (Jaffe et al., 2002), a more conservative measure of knowledge flows would be to define spillovers as only citations included by applicants (Alcacer & Gittelman, 2006). We were unable to use such a measurement strategy because data identifying whether citations were made by applicants or examiners only became available for U.S. patent applications as of 2001.

Because we used patents to measure innovation, we may not have observed all the innovative activity of the sample firms. Although we relied on patent data to measure innovation and knowledge flows, our theory is not necessarily limited to knowledge that is patentable. An originating firm can learn by observing how recipient firms recombine its ideas with other ideas regardless of whether such ideas are patented or not. Further corroborative evidence using different methods, samples, and industries is needed to further validate this study’s findings.

The results of our study and its limitations suggest interesting opportunities for future research. Our theoretical development and analysis provide no guidance as to the net economic benefit of spillovers for originating firms. Although our results suggest that knowledge spillovers may provide some benefit to originating firms in terms of learning opportunities and subsequent innovation, we do not account for reduction of originating firms’ profit potential that is due to knowledge spillovers. To what extent and under what conditions does the potential to learn vicariously from recipient firms compensate for economic losses experienced by originating firms because of knowledge spillovers? We believe this is a worthy question that logically extends from our work.

Further, given our focus on different aspects of spillover knowledge pools, we did not investigate characteristics of recipient firms that may influence an originating firm’s ability to exploit its spillover knowledge pool. Such moderating influences may include the status or reputation of recipient firms (Podolny & Stuart, 1995), their geographic proximity to the originator (Furman et al., 2007), whether or not they are product-market competitors (Stuart, 1999), and whether they have formal knowledge-sharing relationships (e.g., technology alliances).
with the originating firm (Gomes-Casseres, Hagedoorn, & Jaffe, 2006). Researchers should investigate the characteristics of recipient firms that influence the role that spillover knowledge pools play in a firm’s innovative output. For example, a spillover knowledge pool may provide greater learning benefits when the recipient firms contributing to such a pool are in the same industry or geographic region as the originating firm.

Likewise, aspects of an originating firm may enhance its ability to learn from its spillover knowledge pool. Understanding internal systems that firms use to exploit the opportunity to learn from what others have learned from them would be worthy of future study. Firms that use formal and informal information systems that monitor others who are building on their research may be more innovative than firms that lack such systems. A longitudinal qualitative study using a sample of innovations identified in real time may be ideal for delving deeper into the phenomenon of learning from what others have learned from you. These innovations and their originating firms could be tracked over time to establish how recipient firms used them and to examine the process by which the originating firms learned from how the recipient firms integrated their innovations. By doing so, researchers could gain a richer understanding of this feedback loop back to originating firms, the conditions under which such a phenomenon occurs, and the systems that are needed to benefit from it.

REFERENCES


Hall, B. 2005. Note on the bias in the Herfindahl based on...


---

Hongyan Yang (hyyang2006@gmail.com) is an assistant professor of management in the Hong Kong Polytechnic University. She received her Ph.D. in strategic management from the University of Washington. Her research interests focus on the dynamics of knowledge spillovers, firm technology strategy, and entrepreneurial firm growth.

Corey Phelps (phelps@hec.fr) is an associate professor of strategy and business policy at HEC Paris. He completed his Ph.D. in strategic management at the Stern School of Business, New York University. Dr. Phelps’s current research is focused on understanding how companies learn from extramural sources of knowledge.

H. Kevin Steensma (steensma@u.washington.edu) is a professor of management and McCabe Fellow at the Foster School of Business, University of Washington, Seattle. He received his Ph.D. from the Kelley School of Business at Indiana University. His research interests include knowledge transfer across organizations, alliances, and social identity.