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

Hongyan Yang
The Hong Kong Polytechnic University
Department of Management and Marketing
Hung Hom, Kowloon, Hong Kong
Tel: +852 2766 4021
Fax: +852 2765 0611
E-mail: msyhy@polyu.edu.hk

H. Kevin Steensma
University of Washington
Mackenzie Hall Box 353200
Seattle, WA 98195
Tel: (206) 543-4897
Fax: (206) 685-9392
E-mail: steensma@u.washington.edu

Corey Phelps
University of Washington
Mackenzie Hall Box 353200
Seattle, WA 98195
Tel: (206) 543-6579
Fax: (206) 685-3232
E-mail: cphelps@u.washington.edu

We acknowledge use of the NUS Patent Database.
LEARNING FROM WHAT OTHERS HAVE LEARNED FROM YOU: THE EFFECTS OF KNOWLEDGE SPILLOVERS ON ORIGINATING FIRMS

Abstract

Although research suggests knowledge spillovers benefit imitators often at the expense of originators, we investigate how originating firms may benefit from their own spillovers. When an originating firm’s spillovers are recombined with complementary knowledge by recipient firms, a pool of external knowledge is formed that is inherently related to the originating firm’s knowledge base. This relevant knowledge pool contains valuable opportunities for the originator to learn vicariously from recipients’ recombinatorial activities. In a longitudinal study of 87 telecommunications equipment manufacturers, we find that a firm’s rate of innovation and the extent to which these innovations build on and integrate knowledge from the relevant knowledge pool is greater when its relevant knowledge pool is larger in size and similar to the firm’s existing knowledge base.
In the early 1980s, scientists at Eastman Kodak began their pioneering work on a light-emitting 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 pioneering efforts by combining Kodak’s core OLED 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 build on the advances made by those firms that built on Kodak’s original core OLED technology. In this study, 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 other 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) 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). Knowledge produced by an originating firm may be used by recipient firms for little incremental cost, effectively reducing recipients’ costs to innovate (Arrow, 1962; Spence, 1984). 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; Jaffe, 1986; Jaffe, Trajtenberg, & Henderson, 1993), there is little discussion of the potential benefits of spillovers for originating firms. With the exception of the intent of creating an industry standard (Spencer,
2003), researchers generally assume that originating firms have nothing to gain from knowledge spillovers (Kogut & Zander, 1992). Indeed, spillovers facilitate firm entry into a technological area and can reduce the originator’s ability to profit from innovation (Jaffe, 1986; Stuart, 1999).

In contrast to the conventional view that spillovers benefit recipient firms at the expense of originating firm profitability, we consider conditions where knowledge spillovers benefit originating firms by enhancing their ability to innovate. Drawing on research that characterizes knowledge creation as a recombinatorial search process (e.g., Fleming, 2001), we argue that when knowledge spills over from an originating firm and is recombined with complementary knowledge by recipient firms, a pool of knowledge is amassed that is inherently related to the originator’s knowledge base. Through the spillover process, recipients connect an originator’s knowledge with external knowledge. Because the construction of such a knowledge pool is beyond the control of an originating firm and reflects the choices and innovative efforts of other firms, its influence on the originator’s subsequent learning and innovativeness is exogenous.

We refer to this firm-specific pool of related external knowledge as the originating firm’s relevant knowledge pool. Although originating firms have little control over the spillover process and the development of such knowledge pools (Nelson, 1959), we argue that relevant knowledge pools provide viable opportunities for originating firms to learn vicariously from the recombinatorial activities of recipient firms. By observing how recipients exploit the originating firm’s knowledge, the originating firm can refine its search behavior and more effectively exploit recombinatorial opportunities in the future. We argue that originating firms can learn more effectively and efficiently from the knowledge produced by other firms when this knowledge is directly related, via the spillover process, to their existing knowledge.

To better understand our phenomenon of interest, we interviewed research scientists and R&D managers in corporate, governmental, and academic sectors. Analysis of these interviews provided grounding for our theory development. Our theory draws on research in organizational
learning and recombinatorial search in firm innovation. Based on our interview data and
deductive theory, we argue that two aspects of a firm’s relevant knowledge pool will influence
the firm’s innovativeness. We test these predictions using longitudinal data on a panel of 87
telecommunications equipment manufacturers over a 10-year period. We find that a firm’s rate
of innovation and the extent to which these innovations build on knowledge from the relevant
knowledge pool is greater when its relevant knowledge pool is larger in size and more similar to
the firm’s existing knowledge base. Although spillovers may introduce competition and hamper
an originating firm’s ability to profit from innovation, our results suggest an originating firm can
also benefit from the spillover process by learning from how recipient firms exploit the
originating firm’s knowledge.

While research indicates that knowledge spillovers enhance the innovation performance
of recipient firms, this study is the first of which we are aware to show that originating firms can
benefit from their own spillovers. Our results suggest that originating firms can learn from what
other firms have learned from them. We conclude that managers should view spillovers as
potentially valuable learning opportunities and implement mechanisms to enhance their firms’
abilities to exploit these opportunities in future innovation efforts. As we discuss below, our
results have substantive implications for research on absorptive capacity, technological
opportunity, innovation management and sources of economic growth.

THEORY AND HYPOTHESES

Search, Organizational Learning, and Innovation

We adopt a recombinatorial search perspective in explaining the process of innovation
(Fleming, 2001). Innovation is a problem-solving process in which solutions to economically
valuable problems are discovered via search (Dosi 1988). Search processes leading to the
creation of new knowledge most often involve the novel recombination of existing elements of
knowledge, problems, or solutions (Fleming, 2001), or the reconfiguration of the ways in which
knowledge elements are linked (Henderson & Clark, 1990). The search for new recombinations of knowledge elements is uncertain, costly, and guided by prior experience (Dosi, 1988). How search is conducted has important implications for the efficacy and efficiency of innovation (Simon, 1962). Both experiential and heuristic search guide a firm’s quest for valuable solutions.

Experiential search is based on a firm’s direct experience with its own prior solution attempts. In experiential search, innovators draw on their existing knowledge or refine prior solutions (Gavetti & Levinthal, 2000). New recombination attempts involve altering one element of a known solution at a time and then observing the resulting change in performance. The process continues until the problem is adequately solved (Gavetti & Levinthal, 2000). Firms learn from their own experience, myopic to others’ search efforts and outcomes (Levinthal & March, 1993). Prior successful combinations of knowledge provide firms with a template for identifying valuable solutions to new problems, while failed attempts provide insight into which combinations and components to avoid (Nelson & Winter, 1982).

Exploiting existing knowledge through experiential search has certain advantages. Learning from past outcomes leads to a more focused search space (Fleming, 2001). Over time, feedback from past search efforts becomes embodied in organizational routines that efficiently guide the innovation search activities of organizational members (Nelson & Winter, 1982). Recombining the knowledge in one’s existing knowledge base leads to positive, timely, and predictable returns (March, 1991). Because bounded rationality biases individuals toward searching salient areas of their existing knowledge, firms generally exploit their existing

---

1 The concept of recombination is an evolutionary metaphor (Nelson & Winter, 1982). The use of this metaphor has a well-established tradition in studies of technological innovation (Basalla, 1988; Fleming, 2001; Gilfillan, 1935; Nelson & Winter, 1982; Penrose, 1959; Schumpeter, 1934; Usher, 1954). In biology, recombination refers to new combinations of genes in children that were not present in their parents. In the long line of studies on technological innovation that have employed this metaphor, recombination involves novel combinations of discrete elements of knowledge (e.g., Basalla, 1988; Fleming, 2001). This literature uniformly and abstractly characterizes knowledge as discrete elements or components, which 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.
knowledge in their innovation efforts (Helfat, 1994; Stuart & Podolny, 1996).

Although exploiting existing knowledge through experiential search is an efficient means for discovering valuable solutions, it has several limitations. Because search is conducted within the realm of the firm’s established competence, advancements tend to be incremental (Dosi, 1988). Continually exploiting existing competencies can eventually exhaust their potential as sources of novel solutions (Fleming, 2001) or lead to sub-optimal solutions (Levinthal, 1997).

Integrating external knowledge into the search process can overcome the limitations of relying on one’s existing knowledge base (Rosenkopf & Nerkar, 2001). Increasing the number and variety of knowledge components available for recombination increases the number of combinatorial possibilities and potential solutions (Fleming, 2001). Exploring external knowledge sources can challenge existing beliefs about cause-effect relationships and lead to novel insights and solutions (Simonton, 1999). By searching external knowledge, firms can develop multiple conceptualizations of problems and solutions and can potentially apply solutions from one area to problems in another (Hargadon & Sutton, 1997). Exploring external knowledge increases the variance in the set of possible solutions, increasing both the number of failures and the number of highly novel solutions (Levinthal & March, 1981). Despite the benefits of exploring external knowledge, doing so is more uncertain, more costly, and less successful on average than deriving solutions from one’s existing knowledge base (March, 1991).

Heuristic search can reduce the risks of exploration and enhance a firm’s innovation performance. As opposed to relying on feedback from experiential search, heuristic search allows members of a firm to evaluate alternative knowledge components and combinations cognitively and assess their implications for solution performance. An innovator’s cognitive representation of the solution space is used to identify potentially valuable combinations quickly, and the combinations are then investigated via experiential search (Gavetti & Levinthal, 2000). Compared to experiential search, heuristic search enables consideration of a more extensive set
of alternative combinations. Because a firm can evaluate a solution without directly implementing it, heuristic search is cheaper than experiential search, reduces the risks of experimentation with alternative components and combinations that differ from established competences, 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; Manz & Sims, 1981, Levitt & March, 1988). 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 template or recipe to guide 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; Winter, 1987). Thus, a firm’s innovation efforts, including its routines and the outcomes of these routines, can serve as templates for other firms in their innovative pursuits (Hoetker & Agarwal, 2007).

In the next section, we integrate insights from heuristic and experiential search to develop the concept of a relevant knowledge pool. We suggest that an originating firm’s relevant knowledge pool, developed through the spillover process, represents viable opportunities for the originating firm to learn vicariously from the innovation efforts of recipient firms.

**Knowledge Spillovers, Knowledge Pools, and the Relevant Knowledge Pool**

A knowledge spillover is a flow of knowledge from an originating firm to a recipient firm. Knowledge is deemed to have spilled over from an originating firm only if it is used by
recipients in their subsequent innovation pursuits (Grossman & Helpman, 1992). Thus, spillovers reflect the choices and innovative efforts of recipient firms and are therefore largely beyond the control of an originating firm. Although firms originate and receive knowledge flows concurrently, originators and recipients are conceptually distinct in a spillover process. Knowledge flows from originators to recipients, regardless of whether recipients actively search for or passively absorb the knowledge (Kogut & Zander, 1992; Nelson & Winter, 1982).

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 within 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 that are located in the same geographic region as the firm (Jaffe et al., 1993). We build on this notion of knowledge pools that can enhance firms’ innovative performance. However, in contrast to prior research that conceives of knowledge pools that are equally accessible by all firms within the particular boundary (i.e., industry, technological domain, geography), we envision an external knowledge pool that is unique and specific to each originating firm. We conceptualize a pool that is bounded by the recombinatorial activities of those firms that exploit spillovers from the originating firm.

When knowledge spills over from originating firms to recipient firms, the spillover typically does not flow as a complete and well-packaged gift (Sorenson et al., 2006). For recipient firms to exploit the knowledge of others, recipients often need to combine this knowledge with additional knowledge from their own idiosyncratic knowledge context. Because each firm’s knowledge context is unique, the manner in which recipient firms exploit knowledge spillovers varies across recipients and differs from the originating firm. Through the recombinatorial activity of recipient firms, the knowledge produced by an originating firm is
linked directly to external knowledge. In essence, the flow of knowledge from an originator to
recipients amasses a pool of external knowledge that is related to the originating firm’s
knowledge base via the recombinatorial activities of recipient firms. An originator’s relevant
knowledge pool represents all external knowledge that has been linked directly to the
originator’s knowledge by recipient firms through the spillover process.

Figure 1 provides a simple stylized example of the formation of such a knowledge pool.
Consistent with prior research on technological innovation and recombinatorial search, we
characterize knowledge spillovers as discrete knowledge components (see footnote 1 above). The
originating firm’s (Firm 1) knowledge component \(a\) spills over to recipient firms 2 and 3. Firm 2
combines knowledge component \(a\) with knowledge components \(b\) and \(c\). Firm 3 combines
knowledge component \(a\) with knowledge components \(d\) and \(e\). By recombining component \(a\)
with components \(b\), \(c\), \(d\), and \(e\), new knowledge combinations \(g\) and \(f\) are created. The boundary
of Firm 1’s relevant knowledge pool is distinct. Knowledge components \(b\), \(c\), \(d\), \(e\), \(g\), and \(f\)
constitute the knowledge within the originating firm’s relevant knowledge pool. Knowledge
component \(a\) is not part of the relevant knowledge pool because it was produced by the
originating firm and is thus part of its existing knowledge base. Knowledge components \(h\), \(i\), and
\(j\) are also outside of the relevant knowledge pool. Although \(h\), \(i\), and \(j\) may have been created by
firms in a similar technological domain or within the same geographic region as the originating
firm, this knowledge has not been connected to the originating firm’s knowledge by recipient
firms through the spillover process. Thus, in contrast to knowledge pools based on industry,
technological domain, or geographic location, the relevant knowledge pool is unique to each firm.

*** Insert Figure 1 About Here ***

Returning to 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 is 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 the firm (Ginter & White, 1982; Ingram, 2002; Manz & Sims, 1981). As compared to knowledge outside of the relevant knowledge pool, the combinations and knowledge components contained within an originating firm’s relevant knowledge 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; Couzin, 2005). Scientific and industrial associations, employee mobility among firms, and individual social networks further facilitate information flow (von Hippel, 1994). These information channels facilitate not only the flow of knowledge from originating firms to recipient firms, but also knowledge flowing back to originating firms once recipients have recombined the spillover with other knowledge.

For example, a research scientist in a large electronics company conveyed 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 use such resources as the *Science Citation Index*, conference presentations, journal publications, patent applications, and informal conversations to track how their knowledge is 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 Relevant Knowledge Pools and Their Effect on Originating Firms**

Our core proposition is that a firm’s relevant knowledge pool will influence its search processes and innovativeness by highlighting novel combinations of knowledge and the organizational routines of the recipient firms that generated the innovations. These novel combinations and routines represent templates, which can be learned vicariously and incorporated into an originating firm’s heuristic search. Because the relevant knowledge pool is the direct extension of the originating firm’s knowledge, the externally developed templates encompassed within the relevant knowledge pool will be salient, easily understood, and readily applied to future solution search. The originating firm can more easily understand and exploit knowledge in the relevant knowledge pool than knowledge outside of the pool, 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 relative to unrelated knowledge (Lane & Lubatkin, 1998). Two characteristics of a firm’s relevant knowledge pool – size and similarity – will influence a firm’s innovativeness in terms of its rate of innovation and the extent to which these innovations build on the firm’s relevant knowledge pool.

**Pool size.** The relevant knowledge pool highlights new recombinatorial opportunities
associated with an originating firm’s existing knowledge base. As opposed to exhausting recombinatorial opportunities, recombinations spawn even more recombinatorial opportunities (Fleming & Sorenson, 2001). As the relevant knowledge pool of an originating firm grows larger in terms of knowledge components, the number of future recombinatorial opportunities associated with the originating firm’s knowledge base increases (Fleming & Sorenson, 2001).

In general, larger relevant knowledge pools will provide greater potential for the originating firm to learn vicariously from the recombinatorial activity of other firms. By observing how recipient firms combine the originating firm’s knowledge with other knowledge components, the originating firm’s search routines are enhanced, enabling it to search more efficiently and effectively both internal and external knowledge for possible solutions.

Learning vicariously from recipient firms will improve the originating firm’s ability to exploit existing competencies. When recipient firms combine knowledge from the originating firm with other knowledge components, they may do so in ways currently unexplored by the originating firm. Observing how recipient firms exploit the knowledge of the originating firm can assist the originating firm in identifying novel combinations of well-understood components or combinations to avoid. As the originating firm’s search routines become refined, the firm understands its existing competencies more deeply and thus can exploit these competencies more fully. 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.

Learning vicariously from recipient firms can be especially valuable when recipients combine relatively novel knowledge from technology areas where the originating firm has limited experience. 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
originating firm 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 technology and also provided insight to Kodak as to how to further advance its original discovery.

The extent to which a firm learns from its relevant knowledge pool will be evident in the knowledge that is embodied in its subsequent innovations. Firms are biased toward exploiting knowledge that is familiar and accessible because it is generally more efficient to do so (Helfat, 1994; March, 1991). Although knowledge within the firm’s existing knowledge base is the most accessible and easiest to exploit, relying solely on internal knowledge is limiting (Fleming, 2001; Levinthal, 1997). A firm’s relevant knowledge pool is a source of external knowledge that is relatively more accessible and easier to exploit than external knowledge not in the pool. Larger relevant knowledge pools increase the efficiency of the originating firm’s innovative pursuits because the originating firm can draw upon a larger body of relatively accessible external knowledge from which to build its innovations. Firms with larger relevant knowledge pools can rely to a greater extent on the knowledge that resides within their relevant 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 relevant knowledge pools will be compelled to search less accessible external knowledge sources beyond their somewhat limited relevant knowledge pools. Because of the general bias to exploit knowledge sources that are relatively accessible and efficient to search, the larger a firm’s relevant knowledge pool, the greater the extent to which the originating firm’s subsequent innovations will build on knowledge from its relevant
knowledge pool.

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

**Hypothesis 1a:** The larger an originating firm’s relevant knowledge pool, the greater the extent to which the subsequent innovations of the originating firm integrate knowledge from the relevant knowledge pool.

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

**Pool similarity.** Beyond size, relevant knowledge pools will also vary in terms of the similarity between the knowledge in the pool and the originating firm’s existing knowledge base. Similarity refers to the extent to which the knowledge in a firm’s relevant knowledge pool is located in the same domains of knowledge in which the originating firm has expertise. For some originating firms, knowledge within their relevant knowledge pools will be relatively different (i.e., distant) from their existing knowledge base. For other originating firms, the knowledge within their relevant knowledge pools will be more closely related (i.e., similar) to their existing knowledge base. The similarity of the knowledge within the relevant knowledge pool to the originating firm’s existing knowledge will influence the extent to which the vicarious learning opportunities associated with the knowledge pool are valuable and accessible to the originating firm. 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 art innovations in these technological areas.*
When the knowledge within the relevant knowledge pool is similar to the originating firm’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 and easily understood by the originating firm. By observing the recombinatorial activities of recipient firms in domains of knowledge similar to its own, an originating firm will deepen its understanding of its existing competencies and identify new ways to exploit those competencies.

Many of our interview subjects explained how the innovation of recipient firms in domains similar to their own enhanced their innovation efforts. The head of R&D for a large 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 a drug developed and commercialized by his company was shown, by another organization, to be effective against a similar disease and how this enhanced his firm’s understanding of the compound:

*...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 deleted] acts.*

In contrast, absorbing external knowledge that is dissimilar to existing competencies can be quite challenging and costly (Cohen & Levinthal, 1990; Lane & Lubatkin, 1998). Comprehending interactions among unfamiliar knowledge is difficult due to limited experience and cognitive capacity (Fleming & Sorenson, 2001). Absorbing similar knowledge is more efficient than absorbing dissimilar knowledge (Cohen & Levinthal, 1990). Thus, firms must expend greater effort and more resources to integrate dissimilar knowledge, often resulting in diseconomies of scale in their innovation efforts (Ahuja & Lampert, 2001). Because originators do not have direct access to the recipient firms’ routines that helped produce the innovations, they will 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 extend 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 relevant knowledge pool to the originating firm’s knowledge base will influence both its rate of innovation and the extent to which these innovations integrate knowledge from the relevant knowledge pool. The more dissimilar the external knowledge is to the existing knowledge, the greater the recombinatorial uncertainty and the less likely are successful recombinations (Fleming & Sorenson, 2001). Thus, relevant knowledge pools that are more similar to the existing knowledge of the originating firm will 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, the originating firm’s innovations will build on knowledge from its relevant knowledge pool to a greater extent when the knowledge within the pool is similar to the originating firm’s existing knowledge base.

**Hypothesis 2a:** The more similar the knowledge in an originating firm’s relevant knowledge pool to its existing knowledge, the greater the extent to which subsequent innovations of the originating firm integrate knowledge from the relevant knowledge pool.

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

**METHODS**

Our study combines qualitative field work with hypothesis testing using quantitative panel data. To improve our understanding of the phenomenon and inform our theory development, we interviewed 10 senior research scientists and R&D managers in corporate, governmental, and academic sectors. Using a semi-structured protocol, we explored: (1) the extent to which the subject and his/her colleagues monitor how their research ideas and outputs are used by others inside and outside the organization, (2) what influences the intensity and focus
of this monitoring activity, and (3) the ways in which the results of these monitoring efforts influence their subsequent research efforts. The interviews lasted, on average, nearly 45 minutes and generated a total of 162 pages of printed transcripts. We use this evidence to inform, illustrate, and validate our theoretical arguments.

To test our hypotheses, we identify and examine the knowledge spillovers and innovations generated by a sample of firms. We observe such knowledge creation by the production of novel technological inventions using patents. We use patent citations to the originating firm’s patents to assess knowledge spillovers. We build on a large body of research that has used patents to proxy for firm innovation (e.g., Griliches, 1990; Hagedoorn & Cloodt, 2003) and numerous studies that have employed the citations a patent receives as an indicator of the extent to which the knowledge embodied in the patent has diffused or spilled over to other organizations (e.g., Almeida, 1996; Hoetker & Agarwal, 2007; Jaffe et al., 1993; Peri, 2005). 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 an observable paper trail of knowledge flows, regardless of the diffusion mechanism (Jaffe et al., 1993; Jaffe et al, 2000).

Patent data have a number of advantages for our study. First, knowledge creation is instantiated in the form of inventions (Schmookler, 1966), 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 a positive expectation by the inventor of the economic significance of the invention (Griliches, 1990). Although patents reflect a codifiable portion of a firm’s technological knowledge, they positively correlate with measures that incorporate tacit knowledge, such as ratings by technical experts on the technical competencies of firms (Narin, Noma, & Perry,
1987), the introduction of new products (Brouwer & Kleinknecht, 1999), and innovation counts (Basberg, 1982). Trajtenberg (1987) 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 (the “prior art”), which represent valid and reliable indicators of knowledge diffusion. These prior art citations represent the pre-existing technological components that have been 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 list of cited patents contained in the granted patent and ensure that all relevant prior art has been cited. Examiners often add citations to patent applications (Alcacer & Gittelman, 2006), suggesting that applicant firms are not necessarily aware of all cited patents. While examiner-added citations may add noise to measuring knowledge spillovers, many studies have demonstrated that patent citations are valid indicators of actual knowledge flows (cf. Jaffe, Trajtenberg, & Fogarty, 2002; Duguet & MacGarvie, 2005).

**Empirical context and sample**

Both theoretical and practical considerations influenced our choice of empirical setting (i.e., industry). 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 advance in order to observe sufficient knowledge spillovers. Third, because we use patent data for multiple measures and systematic differences exist in the use of patents across industries (Levin et al., 1987), we needed to study an industry in which firms actively patent their inventions (Griliches, 1990). 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 enables 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 is 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 numerous Bureau of Labor Statistics’ studies (e.g., Hecker 1999; Luker & Lyons 1997). Second, technological knowledge has diffused quickly and advances have accumulated over time: 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 relatively large number of citations to other patents (Hall, Jaffe, & Trajtenberg 2001). Finally, with respect to patenting propensity, research shows that telecom equipment firms routinely patent their inventions (Griliches, 1990; Levin et al., 1987; OECD, 1999). Hagedoorn and Cloodt (2003) found that patents are a particularly good measure of innovation in this industry.

A number of practical considerations guided the construction of our sample. Because we need to control for unobserved sources of firm differences in innovativeness in order to accurately assess the effect of a firm’s relevant knowledge pool on its innovativeness, we required sufficient time varying data on the same set of firms. To minimize the influence of left and right censoring regarding the collection of patent data and to ensure access to firm financial data, we chose to limit the sample period to 1987-1997. As we explain below, our independent variables require a ten-year window of patent data prior to the focal firm-year observation, requiring the collection of patents applied for in the mid-1970s (toward the beginning of our patent data sources). Furthermore, collecting financial data on many international firms prior to 1987 proved quite difficult. Given the lag between the application date of a patent and its
eventual granting, we chose to end the sample in 1997. This allowed eight years to elapse between the end of the sample and the end of our patent data collection. Nearly 100% of patent applications are decided upon by the USPTO within seven years of application (Hall, Jaffe, & Trajtenberg, 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. Patent data for each firm were collected for the period 1977-2005. Using patents from a single country maintains 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.’s reputation for providing effective IP protection (Pavitt 1988, Rivette 1993).

Because patents are often assigned to subsidiaries, which may change their names or merge, we took significant care in aggregating patents 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.

We collected financial data from Compustat, annual reports, SEC filings for U.S firms and from The Japan Company Handbook, Worldscope, and Global Vantage for non-U.S. firms.

**Operationalizing the Relevant Knowledge Pool**
We constructed a firm’s relevant 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 10 years prior to, but not including year $t-1$ (i.e., $t-2$ to $t-11$). This 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 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 relevant knowledge pool for firm $i$ in year $t-1$, we joined the lists of forward citing patents and backward citing patents and removed all redundant patent numbers. All patents contained in a firm’s relevant knowledge pool are unique patents from firms other than firm $i$. These patents reflect the technological components that firm $i$’s patents are related to as a result of recipient firms’ recombinatorial efforts.

To illustrate our operationalization of the relevant knowledge pool, consider the following example as shown in Figure 2. 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 relevant knowledge pool of firm $i$ at year $t-1$.

**** Insert Figure 2 About Here ****

Dependent Variables

$Innovative\ output_{it}$: We operationalized $Innovative\ output$ as the number of successful
patent applications for firm $i$ in year $t$. We used the Delphion database to collect yearly patent counts for each firm. While only patents that were ultimately granted were counted, patents were counted in the year of application to capture more precisely the timing of knowledge creation (Griliches, 1990).

**Knowledge integration**: We operationalized the extent to which the innovative output of the originating firm builds on the knowledge from its relevant knowledge pool as the proportion of prior art patents contained in firm $i$’s patents of year $t$ that belong to its relevant 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 relevant knowledge pool. The extent to which a firm uses elements of knowledge (e.g., patents) contained in its relevant knowledge pool reflects that it is searching and exploiting this knowledge pool.

### Independent Variables

After constructing the relevant knowledge pool for each firm-year, we measured two aspects of the pool as our independent variables. We use the subscript $it-1$ below to indicate that the independent and control variables are lagged one year relative to the dependent variable.

**Pool size**$_{it-1}$ was operationalized as the total number of unique patents in firm $i$’s relevant knowledge pool at year $t-1$. This variable was log-transformed because of skewness.

**Pool similarity**$_{it-1}$: To measure the technological similarity between a firm’s knowledge pool and its extant knowledge base, we use an index developed by Jaffe (1986). For each firm-year in the sample, we constructed an index that measures the distribution of a firm’s patents across primary patent classes and the distribution of a firm’s relevant 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 $j$. This approach assumes 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 relevant 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 relevant knowledge pool that are in patent class $j$. The similarity of firm $i$’s relevant knowledge pool in year $t-1$ was calculated as:

$$\text{Pool Similarity}_{it-1} = \left[ \frac{1}{J} \sum_{j=1}^{J} d_j e_{ij} / \left( \sum_{j=1}^{J} d_j^2 \right)^{1/2} \left( \sum_{j=1}^{J} e_{ij}^2 \right)^{1/2} \right]$$

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

**Control variables**

*Technological opportunity*$_{it-1}$: “Technological opportunity” refers to differences in the set of possibilities for technological advance that exist within technologies and industries at different points in time (Klevorick et al., 1995). Thus, some firms may be active in relatively richer technological areas than other firms. Based on Patel and Pavitt (1997), we control for firm-specific differences in technological opportunity in year $t-1$ as follows:

$$\text{Technological opportunity}_{it-1} = \sum_{j=1}^{J} \left[ \text{Patents}_{jt-1} \ast P_{jit-1} \right],$$

where $\text{Patents}_{jt-1}$ refers to the number of patents granted in the U.S. in patent class $j$ in year $t-1$, and $P_{jit-1}$ is the proportion of firm $i$’s patents applied for in class $j$ in year $t-1$ (and subsequently granted). The total number of patents granted in each patent class in a given year proxies for the underlying rate of technical change in that area of technology (Patel & Pavitt, 1997). We divided this variable by 1000 to reduce its scale and ease its interpretation.

*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). To control for this potential confound, we include a variable that indexes the number of unique companies whose
patents, applied for in year $t-1$ (and subsequently granted), cite firm $i$’s 10-year stock of patents.

**Patent stock quality$_{t-1}$**: The number of (forward) citations a patent receives is a good indicator of the value or quality of the patented invention (Harhoff et al., 1999; Trajtenberg, 1990). Firms that produce valuable patents are at greater risk of having larger relevant knowledge pools. To control for a firm’s patent stock quality, we first identify the patents it applied for and was granted in the 10 years prior to and including year $t-1$. We then identify all citations this stock of patents received by the end of 2005. We compute firm $i$’s Patent Stock Quality in year $t-1$ as the number of forward citations (to the patent stock).

**Firm size$_{t-1}$**: Prior research has proved inconclusive in determining whether small or large firms are more innovative (Cohen & Levin, 1989). This may be due to the presence of both negative and positive effects of size on firm innovation (Teece, 1992). We control for the influence of firm size using firm $i$’s sales in year $t-1$ (in billion $US$).

**R&D$_{t-1}$**: The annual R&D expenditures of a firm represent formal investments in knowledge creation (Griliches, 1990) and contribute to the firm’s ability to absorb knowledge from external sources (Cohen & Levinthal, 1990). We control for the influence of R&D expenditure for firm $i$ in year $t-1$(in billion $US$).

**Current ratio$_{t-1}$**: Organizations having significant slack resources participate more in exploratory search than their slack-deprived counterparts (Singh, 1986), which may result in enhanced innovative performance (Nohria & Gulati, 1996). We control for the slack resources of firm $i$ in year $t-1$ using its current ratio (current assets/current liabilities, measured at year end).

**Firm technological diversity$_{t-1}$**: Increasing technological diversity can increase firm innovation due to internal spillovers (Garcia-Vega 2006) and can increase a firm’s ability to absorb knowledge from external sources (Cohen & Levinthal 1990). We measure firm $i$’s technological diversity in year $t-1$ using Hall’s (2002) adjusted Herfindahl index:

$$\text{Firm Technological Diversity}_{t-1} = \left[1 - \sum_{j=1}^{J} \left( \frac{N_{ji-1}}{N_{i-1}} \right)^2 \right] \cdot \frac{N_{i-1} - 1}{N_{i-1} - 1},$$
where $N_{it,j}$ is the total 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 between 0 (no diversity) to 1 (maximum diversity).

**Estimation**

We estimate two variants of Griliches’ (1979) firm-level knowledge production function. In the first, we model the production of knowledge outputs in the form of patents as a function of a set of knowledge inputs. In the second, we model the production of knowledge outputs in the form of patent citations. Our knowledge production functions include measures of a firm-specific external knowledge pool to explicitly capture the spillback effects of a firm’s spillovers. Our models also include internal firm knowledge stock and flow variables and other covariates. All explanatory and control variables are lagged one year, which accounts for the delay in converting innovation inputs into outputs, reduces concerns of reverse causality and avoids simultaneity. We estimated all models using panel regression techniques. The panel is unbalanced as some firms were acquired or restructured, making within-firm comparisons difficult.

We employ two dependent variables in this study (i.e., **Innovative output** and **Knowledge integration**). The numerical values each variable can take on and their distributions are quite different. Accordingly, we estimate our models using regression methods appropriate to each dependent variable.

Our first dependent variable, **Innovative output**, is a count variable that can take on only non-negative integer values. The use of linear regression to model such data can result in inefficient, inconsistent, and biased coefficient estimates (Long, 1997). While Poisson regression is appropriate to model count data, our data were significantly overdispersed, violating a basic assumption of the Poisson estimator (Hausman, Hall and Griliches, 1984). As a more appropriate measure, we use 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). The negative binomial panel estimator accommodates explicit control of persistent individual unobserved effects through both fixed and random effects. We include year dummies to control for unobserved systematic period effects. We also employ firm fixed effects to control for unobserved, temporally stable firm differences in patenting. We use 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 differ substantively across these two estimation approaches. We use fixed rather than random effects because the results of Hausman tests for the models estimated below indicated rejection of the random effects specification. Allison and Waterman (2002) showed that conventional standard error estimates are biased downward when using unconditional fixed effects, and that this downward bias is effectively corrected by multiplying standard errors by the square root of the deviance statistic divided by its degrees of freedom. We implemented that correction in all reported models.

The second dependent variable, Knowledge integration, is a proportion. Estimation involving a proportional dependent variable presents several challenges to linear regression (Gujarati, 1995). Following standard econometric practice (Greene, 1997), we transform this variable using a logit (i.e., log odds) transformation. We estimate our models using panel linear regression with firm fixed effects and year dummies and employ robust standard errors.

RESULTS

Allison and Waterman (2002) criticized Hausman Hall and Griliches’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 with linear fixed effects estimators. Allison and Waterman (2002) developed an unconditional (maximum likelihood) negative binomial model that uses dummy variables to represent fixed effects, which effectively controls for all stable individual effects. According to Allison and Waterman, Hausman et al. did not formulate a true fixed effects model in the mean of the random (dependent) variable. Their formulation layers the fixed effect into the heterogeneity portion of the model and not the conditional mean. This portion is then conditioned out of the distribution to produce the Hausman et al. model that is estimated.

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.
Table 1 shows descriptive statistics and correlations for all variables. Table 2 reports the regression results using firm and year fixed effects. We ran similar models for both of our 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 include the squared terms of Pool size and Pool similarity in Models 3 and 6. Models 2 and 5 are the full model, while Models 3 and 6 are for a post hoc test. 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 is more appropriate than random effects. Given that our estimations include 95 dummy variables (86 firm and nine year dummies) and many time-varying controls, we believe our results represent conservative tests of our hypotheses.

**** Insert Tables 1 and 2 About Here ****

Hypothesis 1a and 1b proposed positive relationships between pool size and our two dependent variables: innovative output and knowledge integration. The effect of pool size is significant and positive for both innovative output ($p<0.01$) and knowledge integration ($p<.01$). Thus, Hypotheses 1a and 1b are supported.

Hypothesis 2a and 2b proposed that the similarity between the relevant knowledge pool and the originating firm’s existing knowledge base would have a positive influence on both innovative output and knowledge integration. The effect of pool similarity is significant ($p<0.01$) in all models for both dependent variables. Thus, hypotheses 2a and 2b are supported.

Control variables. Six of the eight control variables exhibited consistently significant effects on Innovative output (Models 1-3). Technological opportunity and Firm technological diversity were not significant. Finally, counter to our expectation, Number of recipients, a measure of crowding, had a negative and significant effect on Innovative output. While crowded areas of technological activity may lead to increased rates of firm innovation due to increased

---

4 The average Variance Inflation Factor for models 3 and 6 is 5.14, indicating that multicollinearity is not an issue.
competitive pressures to innovate (Stuart, 1999), our finding suggests crowding may lead to the preemption or foreclosure of innovative opportunities. In contrast to these results, in models with *Knowledge integration* as the dependent variable (Models 4-6), only *Firm size* was significant. **Robustness checks.** We tested for possible 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<0.1) in Model 6, suggesting an inverted U-shaped effect of *Pool size* on *Knowledge integration*. Based on the coefficients in Model 6, 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 squared term of *Pool similarity* has a positive and significant influence on both *Innovative output* and *Technology integration* (p<0.01 and p<0.05 respectively). Given the positive and significant effect of both first and second order terms, pool similarity appears to have an exponentially increasing influence on a firm’s innovation for our sample.

To test the sensitivity of our results to our choice of a 10-year window for the construction of a firm’s relevant knowledge pool, we reconstructed each knowledge pool variable using a five-year window. The results using this alternative window were consistent with those reported in Table 2. **Supplementary analysis**

The preceding firm-level analysis provides support for our basic premise that originating firms can more easily understand and exploit knowledge from their relevant knowledge pool as compared to knowledge outside of the pool. As such, relevant knowledge pools influence the innovation of originating firms. To further investigate our basic premise, we used an experimental design at the patent level of analysis by employing the patent case-control method (Agarwal, Cockburn & McHale, 2006; Almeida 1996; Almeida & Kogut, 1999; Furman & Stern,
Empirically, we want to examine whether patents that are part of a firm’s relevant knowledge pool are more often used by the firm in its subsequent innovation efforts (i.e., cited more frequently by the firm) than patents with similar observable characteristics that are not contained in the relevant knowledge pool. If originating firms cite randomly-sampled patents that have been “treated” with knowledge from their relevant knowledge pools more frequently than comparable untreated control patents, then we will have additional evidence that firms tend to exploit knowledge contained in their respective relevant knowledge pools more so than knowledge from outside the pool.

To construct the sample for this analysis, we began by randomly sampling 1000 patents from the relevant knowledge pools of our sample firms measured in 1987. We chose to sample patents from the first year of our primary sample (described above) to allow for the maximum time to observe subsequent citations to these patents. We stratified this sample based on the size of each firm's relevant knowledge pool in 1987. The number of patents sampled for each firm was computed as: $1000 \times \left( \frac{RKP_i}{\sum RKP_j} \right), \ i \in j$, where $RKP_i$ is the size of firm $i$'s relevant knowledge pool in 1987 and the denominator is the sum of the sizes of all sample firms knowledge pools in 1987. We refer to this set of patents as our treatment patents.

Following Jaffe et al. (1993), we matched each of these treatment patents to a similar control patent based on the primary three-digit technology class and year of issuance of the treatment patent. Thus, a control patent is issued in the same year and assigned to the same primary technology class as its counterpart treatment patent, but it is not contained in the firm's relevant knowledge pool or existing knowledge base during the period of observation. When multiple patents met these criteria as a control patent, we randomly selected one for inclusion in the sample. By matching control patents to treatment patents in this way, we attempt 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 consists of 1000 treatment and 1000 control patents (sample size = 2000).

The dependent variable for this analysis, \( \text{Forward Cites}_{ijt} \), is the count of forward patent citations received by focal patent \( i \) from focal firm \( j \) in year \( t \). Forward citations are observed for the years 1988-1997. The explanatory variable of interest is a time-invariant dummy variable, \( RKP_i \), which takes on the value of "1" when the focal patent was randomly selected from the focal firm's relevant 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 include three control variables that have been shown to be good proxies for invention quality (Lanjouw & Schankerman, 2004). The claims section of a patent delineates the scope or breadth of property rights protected by the patent (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 control for the number of claims contained in patent \( i \) (\( \text{Claims}_i \)). Patents that cite a larger number of prior art patents may contain a lower degree of novelty and be less valuable (Harhoff, Scherer & Vopel, 2003). We control 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 control for the cumulative number of forward citations patent \( i \) received as of year \( t \), excluding those made by focal firm \( j \) in that year (\( \text{Cumulative Forward Cites}_{it} \)). Patents that build on a wide range of technologies may be more novel and therefore more highly cited (Trajtenberg et al., 1997; Singh, 2008). We control for the originality of patent \( i \) (\( \text{Originality}_i \)) using Trajtenberg et al.'s (1997) originality measure: \( 1 - \sum s_{ij}^2 \), 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 has been at risk of forward citation, we control for the age of patent \( i \) in year \( t \) (\( \text{Age}_{it} \)).
The unit of analysis is the patent. Because we have a balanced panel and the dependent variable is a count, we use panel negative binomial regression. Our data are also hierarchically nested. In addition to the fact that multiple observations are nested within the same patent (i.e. the same patent is observed over time), multiple patents are also nested within a single firm. To accommodate the multilevel nature of our data, we estimate a mixed level negative binomial regression model in which patents are assigned random effects and firms are assigned fixed effects. This allows us to control for unobserved heterogeneity at both the patent and firm level and compute standard errors for patent-level coefficients that are free from the influence of spatial autocorrelation caused by patents being clustered within firms. We also account for the possibility that a patent's value and subsequent citation differs based on its area of technology by including fixed effects for each patent's technological category as defined by the NBER patent database (Hall, Jaffe & Trajtenberg, 2001). Finally, we include year dummies in our estimations.

--- Insert Tables 3 and 4 here ---

Table 3 shows the descriptive statistics and correlations for the variables involved in the supplementary analysis. Table 4 provides the results of the negative binomial regression analysis. The sole explanatory variable, $RKP_i$, is positive and significant ($p<0.01$). The incident-rate ratio $(\exp\beta_i)$ for this variable is 2.05, which indicates that a one unit change in $RKP_i$ (i.e., a patent changes from being outside of a firm’s relevant knowledge pool to being part of the pool) yields a 105% increase in the rate at which an average firm will cite the patent. Thus, we find 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 relevant knowledge pool relative to knowledge outside of the pool.

--- Insert Tables 3 and 4 here ---

Table 3 shows the descriptive statistics and correlations for the variables involved in the supplementary analysis. Table 4 provides the results of the negative binomial regression analysis. The sole explanatory variable, $RKP_i$, is positive and significant ($p<0.01$). The incident-rate ratio $(\exp\beta_i)$ for this variable is 2.05, which indicates that a one unit change in $RKP_i$ (i.e., a patent changes from being outside of a firm’s relevant knowledge pool to being part of the pool) yields a 105% increase in the rate at which an average firm will cite the patent. Thus, we find 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 relevant knowledge pool relative to knowledge outside of the pool.

--- Insert Tables 3 and 4 here ---

Table 3 shows the descriptive statistics and correlations for the variables involved in the supplementary analysis. Table 4 provides the results of the negative binomial regression analysis. The sole explanatory variable, $RKP_i$, is positive and significant ($p<0.01$). The incident-rate ratio $(\exp\beta_i)$ for this variable is 2.05, which indicates that a one unit change in $RKP_i$ (i.e., a patent changes from being outside of a firm’s relevant knowledge pool to being part of the pool) yields a 105% increase in the rate at which an average firm will cite the patent. Thus, we find 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 relevant knowledge pool relative to knowledge outside of the pool.

--- Insert Tables 3 and 4 here ---

Table 3 shows the descriptive statistics and correlations for the variables involved in the supplementary analysis. Table 4 provides the results of the negative binomial regression analysis. The sole explanatory variable, $RKP_i$, is positive and significant ($p<0.01$). The incident-rate ratio $(\exp\beta_i)$ for this variable is 2.05, which indicates that a one unit change in $RKP_i$ (i.e., a patent changes from being outside of a firm’s relevant knowledge pool to being part of the pool) yields a 105% increase in the rate at which an average firm will cite the patent. Thus, we find 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 relevant knowledge pool relative to knowledge outside of the pool.

--- Insert Tables 3 and 4 here ---

Table 3 shows the descriptive statistics and correlations for the variables involved in the supplementary analysis. Table 4 provides the results of the negative binomial regression analysis. The sole explanatory variable, $RKP_i$, is positive and significant ($p<0.01$). The incident-rate ratio $(\exp\beta_i)$ for this variable is 2.05, which indicates that a one unit change in $RKP_i$ (i.e., a patent changes from being outside of a firm’s relevant knowledge pool to being part of the pool) yields a 105% increase in the rate at which an average firm will cite the patent. Thus, we find 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 relevant knowledge pool relative to knowledge outside of the pool.

--- Insert Tables 3 and 4 here ---

Table 3 shows the descriptive statistics and correlations for the variables involved in the supplementary analysis. Table 4 provides the results of the negative binomial regression analysis. The sole explanatory variable, $RKP_i$, is positive and significant ($p<0.01$). The incident-rate ratio $(\exp\beta_i)$ for this variable is 2.05, which indicates that a one unit change in $RKP_i$ (i.e., a patent changes from being outside of a firm’s relevant knowledge pool to being part of the pool) yields a 105% increase in the rate at which an average firm will cite the patent. Thus, we find 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 relevant knowledge pool relative to knowledge outside of the pool.

--- Insert Tables 3 and 4 here ---

Table 3 shows the descriptive statistics and correlations for the variables involved in the supplementary analysis. Table 4 provides the results of the negative binomial regression analysis. The sole explanatory variable, $RKP_i$, is positive and significant ($p<0.01$). The incident-rate ratio $(\exp\beta_i)$ for this variable is 2.05, which indicates that a one unit change in $RKP_i$ (i.e., a patent changes from being outside of a firm’s relevant knowledge pool to being part of the pool) yields a 105% increase in the rate at which an average firm will cite the patent. Thus, we find 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 relevant knowledge pool relative to knowledge outside of the pool.

--- Insert Tables 3 and 4 here ---

Table 3 shows the descriptive statistics and correlations for the variables involved in the supplementary analysis. Table 4 provides the results of the negative binomial regression analysis. The sole explanatory variable, $RKP_i$, is positive and significant ($p<0.01$). The incident-rate ratio $(\exp\beta_i)$ for this variable is 2.05, which indicates that a one unit change in $RKP_i$ (i.e., a patent changes from being outside of a firm’s relevant knowledge pool to being part of the pool) yields a 105% increase in the rate at which an average firm will cite the patent. Thus, we find 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 relevant knowledge pool relative to knowledge outside of the pool.

--- Insert Tables 3 and 4 here ---

Table 3 shows the descriptive statistics and correlations for the variables involved in the supplementary analysis. Table 4 provides the results of the negative binomial regression analysis. The sole explanatory variable, $RKP_i$, is positive and significant ($p<0.01$). The incident-rate ratio $(\exp\beta_i)$ for this variable is 2.05, which indicates that a one unit change in $RKP_i$ (i.e., a patent changes from being outside of a firm’s relevant knowledge pool to being part of the pool) yields a 105% increase in the rate at which an average firm will cite the patent. Thus, we find 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 relevant knowledge pool relative to knowledge outside of the pool.

--- Insert Tables 3 and 4 here ---

Table 3 shows the descriptive statistics and correlations for the variables involved in the supplementary analysis. Table 4 provides the results of the negative binomial regression analysis. The sole explanatory variable, $RKP_i$, is positive and significant ($p<0.01$). The incident-rate ratio $(\exp\beta_i)$ for this variable is 2.05, which indicates that a one unit change in $RKP_i$ (i.e., a patent changes from being outside of a firm’s relevant knowledge pool to being part of the pool) yields a 105% increase in the rate at which an average firm will cite the patent. Thus, we find 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 relevant knowledge pool relative to knowledge outside of the pool.

--- Insert Tables 3 and 4 here ---

Table 3 shows the descriptive statistics and correlations for the variables involved in the supplementary analysis. Table 4 provides the results of the negative binomial regression analysis. The sole explanatory variable, $RKP_i$, is positive and significant ($p<0.01$). The incident-rate ratio $(\exp\beta_i)$ for this variable is 2.05, which indicates that a one unit change in $RKP_i$ (i.e., a patent changes from being outside of a firm’s relevant knowledge pool to being part of the pool) yields a 105% increase in the rate at which an average firm will cite the patent. Thus, we find 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 relevant knowledge pool relative to knowledge outside of the pool.

--- Insert Tables 3 and 4 here ---

Table 3 shows the descriptive statistics and correlations for the variables involved in the supplementary analysis. Table 4 provides the results of the negative binomial regression analysis. The sole explanatory variable, $RKP_i$, is positive and significant ($p<0.01$). The incident-rate ratio $(\exp\beta_i)$ for this variable is 2.05, which indicates that a one unit change in $RKP_i$ (i.e., a patent changes from being outside of a firm’s relevant knowledge pool to being part of the pool) yields a 105% increase in the rate at which an average firm will cite the patent. Thus, we find 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 relevant knowledge pool relative to knowledge outside of the pool.
DISCUSSION

An extensive body of research has investigated the existence of knowledge spillovers and their effect on the innovative performance of recipient firms. Yet, the possibility that knowledge spillovers may benefit the firms from 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. In this study, we investigated the conditions under which knowledge spillovers enhance the originating firm’s innovativeness. This study is the first of which we are aware to investigate this question.

We argued that when an originating firm’s knowledge spillovers are recombined by recipient firms, a pool of knowledge is formed that is different from, yet linked to the originator’s knowledge base through the spillover process. We developed the concept of the relevant knowledge pool to denote this firm-specific pool of related knowledge. The relevant knowledge pool represents viable opportunities for the originating firm 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 the relevant knowledge pool as compared to knowledge outside of the pool.

We found that a firm’s rate of innovation and the extent to which these innovations build on and integrate knowledge from the relevant knowledge pool are greater when this relevant knowledge pool is larger in size and similar to the firm’s existing knowledge base. For those originating firms that have substantial familiarity with the areas of knowledge contained in their relevant knowledge pools, such knowledge will be particularly salient and accessible, reducing the uncertainty of integrating such knowledge in subsequent innovative efforts. In general, larger relevant knowledge pools provide greater opportunity for the originating firm 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 relevant knowledge pool on the extent to which it builds on knowledge from the pool. This suggests there may be diminishing and ultimately negative returns to pool size for knowledge integration. Our theory may offer an explanation for this unanticipated result. Although larger pools provide the originating firm 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; Levinthal & March, 1993). Consequently, as the size of a firm’s relevant 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.

In sum, our results suggest that originating firms can benefit from knowledge spillovers by learning from what others have learned from them. As knowledge spills forward and is exploited by recipient firms, relevant knowledge can also spill back to the originating firm. In effect, we find evidence of a circle of knowledge flow from originator to recipients, and back again to the originator. As such, our findings are consistent with Klevorick et al.’s argument that, “…what a firm learns from its own R&D may be augmented by feedback from other firms that make or use the new product or process” (1995: 192).

Implications

This study makes several contributions to our understanding of firm innovation and its implications for economic growth. The relative 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). Each of these approaches assumes the availability of spillovers is homogenous across firms within a particular boundary. In contrast, our conceptualization of the relevant knowledge pool and the results of our study suggest that the pool of available spillovers is firm-specific.

Our study also contributes to research on absorptive capacity. The concept of a relevant knowledge pool exploits the notion that firms do not possess absorptive capacity in an absolute sense but rather in 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 (Cohen & Levinthal, 1990; Lane & Lubatkin, 1998). While distinguishing between what is and is not related is often undefined in absorptive capacity research (Lane, Koka & Pathak, 2006), our analysis suggests that the domain in which a firm has absorptive capacity is partly circumscribed by its relevant knowledge pool.

The results of our study also speak directly to the economics literature on endogenous growth. Knowledge spillovers play a central role in endogenous growth theory by providing for increasing returns to knowledge production (Romer, 1994). Our results suggest that economic growth may be enhanced not just by the benefits of spillovers for recipient firms but also from their innovation-enhancing effects for originating firms.

We provide a novel explanation as to why firms innovate despite the disincentive associated with spillovers. Existing research suggests that spillovers reduce a firm’s ability to appropriate the returns from investments in knowledge production (Arrow, 1962; Levin & Reiss, 1988). However, this static efficiency argument ignores the dynamic efficiency effect of spillovers that we identify in this paper. Spillovers result in the creation of a firm-specific pool of related extramural knowledge, which the 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 advancements made by recipient firms, will reduce its total costs of innovation and increase the private returns from the original spillovers. Thus, from a dynamic perspective, knowledge that spills over may generate private returns to the originating firm in the future. On the margin, this benefit may mitigate the disincentive effect of knowledge spillovers for originating firms and increase their incentives to invest in innovation. Belenzon (2006) provides initial empirical support for this proposition.

This dynamic incentive effect of knowledge spillovers addresses a fundamental tension in the economics of innovation and endogenous growth literatures. 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 relevant knowledge pool and the results of this study potentially help to reconcile this conflict between the incentive and efficiency effects of appropriability. A relevant 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 the originating firm with useful vicarious learning opportunities and thereby increases the technological opportunities available to the originating firm. As the level of technological opportunity increases in an industry or domain of technology, the rate of innovation in such areas increase (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 and incompatible.

Our analysis suggests that managers should not view knowledge spillovers 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 the originating firm has a comprehensive view of their relevant 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:

_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 use patent citations to assess knowledge spillovers, we do not capture all knowledge flows and may count as evidence of spillovers citations that do not correspond to actual spillovers. Because citations added by patent examiners typically reflect ignorance of these cited patents by inventors (Jaffe et al., 2002), a more conservative measure of knowledge flows would be to define spillovers as only citations included by the applicant (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 patents are used to measure innovation, we may not observe all innovative activity of the sample firms. Our findings may also be unique to the time period studied, the sampled firms, and/or the industry context. Further corroboratory evidence using data from different time periods, samples, and industries is needed to further validate this study’s findings.

The results of our study and its limitations suggest several interesting opportunities for future research. Our theoretical development and analysis provides 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 originating firms’ reduced profit potential due to knowledge spillovers. To what extent and under what conditions might the potential to learn vicariously from recipient firms effectively compensate for economic losses experienced by originating firms due to knowledge spillovers? We believe this is a worthy question that logically extends from our work.

Further, given our focus on different aspects of the relevant knowledge pool, we did not investigate characteristics of recipient firms that may influence an originating firm’s ability to exploit its relevant 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 relevant knowledge pools play in a firm’s innovative output. For example, a relevant knowledge pool may provide greater learning benefits when the recipient firms contributing to such a pool are in the same industry or geographic proximity as the originating firm.

Likewise, aspects of the originating firm may enhance its ability to learn vicariously from its relevant 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. Our results suggest that firms that use formal and informal information systems that monitor others who are building on their research may be more innovative than firms that do not have such systems. This represents an interesting proposition for future research.
REFERENCES


Nelson, R. 1959. The simple economics of basic scientific research. *Journal of Political Economy*, 67: 297-


TABLE 1
Descriptive statistics and correlations

<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></td>
<td></td>
<td>0.37**</td>
<td></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>0.54**</td>
<td>0.60**</td>
<td></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>0.38**</td>
<td>0.23**</td>
<td>0.38**</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>0.32**</td>
<td>0.14**</td>
<td>0.34**</td>
<td>0.23**</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>0.42**</td>
<td>0.89**</td>
<td>0.65**</td>
<td>0.23**</td>
<td>0.17**</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>0.40**</td>
<td>0.92**</td>
<td>0.59**</td>
<td>0.21**</td>
<td>0.14**</td>
<td>0.95**</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>0.36**</td>
<td>0.78**</td>
<td>0.60**</td>
<td>0.17**</td>
<td>0.13</td>
<td>0.86**</td>
<td>0.82**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9 R&amp;D</td>
<td>0.47</td>
<td>0.97</td>
<td>0.35**</td>
<td>0.85</td>
<td>0.60**</td>
<td>0.21**</td>
<td>0.12**</td>
<td>0.87**</td>
<td>0.83**</td>
<td>0.93**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 Current ratio</td>
<td>2.33</td>
<td>1.59</td>
<td>-0.08*</td>
<td>-0.23**</td>
<td>-0.30**</td>
<td>0.06</td>
<td>0.04</td>
<td>-0.23**</td>
<td>-0.22**</td>
<td>-0.25**</td>
<td>-0.21**</td>
<td></td>
</tr>
<tr>
<td>11 Firm technological diversity</td>
<td>0.76</td>
<td>0.31</td>
<td>0.11**</td>
<td>0.30**</td>
<td>0.54**</td>
<td>-0.06</td>
<td>-0.09*</td>
<td>0.32**</td>
<td>0.29**</td>
<td>0.32**</td>
<td>0.32**</td>
<td>-0.38**</td>
</tr>
</tbody>
</table>

N= 724  * p < .05  ** p < .01
### TABLE 2
Negative binomial regression models

<table>
<thead>
<tr>
<th>Dependent 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>Pool size</td>
<td>0.203**</td>
<td>0.215**</td>
</tr>
<tr>
<td></td>
<td>(0.040)</td>
<td>(0.041)</td>
</tr>
<tr>
<td>Pool similarity</td>
<td>0.697**</td>
<td>1.036**</td>
</tr>
<tr>
<td></td>
<td>(0.212)</td>
<td>(0.240)</td>
</tr>
<tr>
<td>Pool size squared</td>
<td>-0.005</td>
<td>-0.115+</td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.068)</td>
</tr>
<tr>
<td>Pool similarity squared</td>
<td>2.186**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.729)</td>
<td></td>
</tr>
<tr>
<td>Control Variables</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technological</td>
<td>0.191</td>
<td>0.058</td>
</tr>
<tr>
<td>Opportunity</td>
<td>(0.127)</td>
<td>(0.123)</td>
</tr>
<tr>
<td>Number of recipients</td>
<td>-0.270**</td>
<td>-0.221**</td>
</tr>
<tr>
<td></td>
<td>(0.049)</td>
<td>(0.046)</td>
</tr>
<tr>
<td>Patent stock quality</td>
<td>0.038**</td>
<td>0.030**</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.006)</td>
</tr>
<tr>
<td>Firm size</td>
<td>-0.021*</td>
<td>-0.015+</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>R&amp;D</td>
<td>0.608**</td>
<td>0.564**</td>
</tr>
<tr>
<td></td>
<td>(0.111)</td>
<td>(0.103)</td>
</tr>
<tr>
<td>Current ratio</td>
<td>0.061+</td>
<td>0.038</td>
</tr>
<tr>
<td></td>
<td>(0.035)</td>
<td>(0.033)</td>
</tr>
<tr>
<td>Firm technological</td>
<td>1.094**</td>
<td>0.708+</td>
</tr>
<tr>
<td>diversity</td>
<td>(0.368)</td>
<td>(0.366)</td>
</tr>
<tr>
<td>Constant</td>
<td>4.078**</td>
<td>3.785**</td>
</tr>
<tr>
<td></td>
<td>(0.482)</td>
<td>(0.467)</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>Degree of freedom</td>
<td>636</td>
<td>634</td>
</tr>
<tr>
<td>R sq</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scaled chi-sq</td>
<td>577.99**</td>
<td>598.19**</td>
</tr>
<tr>
<td>Log likelihood</td>
<td>516632.38</td>
<td>497901.65</td>
</tr>
</tbody>
</table>

- N=739;  +p<.1  *p<.05  **p<.01
- Standard errors in parentheses; adjusted by square root of deviance statistic (Allison & Waterman, 2002)
- Two-tailed test
### TABLE 3
Descriptive Statistics and Correlations for Supplementary Patent-level Analysis

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>S.D.</th>
<th>Min.</th>
<th>Max.</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>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. RKP</td>
<td>0.50</td>
<td>0.50</td>
<td>0</td>
<td>1</td>
<td>0.05</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>0.02</td>
<td>0.05</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>0.03</td>
<td>0.19</td>
<td>0.17</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Cumulative</td>
<td>10.61</td>
<td>15.45</td>
<td>0</td>
<td>390</td>
<td>0.11</td>
<td>0.25</td>
<td>0.18</td>
<td>0.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Forward Cites</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></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>0.01</td>
<td>0.11</td>
<td>0.10</td>
<td>0.35</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td>7. Age</td>
<td>8.29</td>
<td>4.58</td>
<td>0</td>
<td>22</td>
<td>-0.03</td>
<td>0.00</td>
<td>-0.04</td>
<td>-0.10</td>
<td>0.07</td>
<td>-0.03</td>
</tr>
</tbody>
</table>

### TABLE 4
Negative Binomial Regression Model of Forward Citations Made by Focal Firm

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>RKP</td>
<td>0.715**</td>
<td>(0.180)</td>
<td></td>
</tr>
<tr>
<td>Claims</td>
<td>-0.003</td>
<td>(0.008)</td>
<td></td>
</tr>
<tr>
<td>Backward Cites</td>
<td>0.003</td>
<td>(0.010)</td>
<td></td>
</tr>
<tr>
<td>Cumulative Forward Cites</td>
<td>0.048**</td>
<td>(0.006)</td>
<td></td>
</tr>
<tr>
<td>Originality</td>
<td>0.009</td>
<td>(0.310)</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>-0.124**</td>
<td>(0.028)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-3.193**</td>
<td>(0.692)</td>
<td></td>
</tr>
<tr>
<td>Firm dummies</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technology category dummies</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year dummies</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>20,000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log Likelihood</td>
<td>-1283.324</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wald Chi Squared (dof)</td>
<td>196.93***</td>
<td>(80)</td>
<td></td>
</tr>
</tbody>
</table>

Standard errors in parentheses

** p<0.01, * p<0.05, + p<0.1; two-tailed tests
FIGURE 1
The relevant knowledge pool

\[ \bigcirc + \bigtriangleup + \bigtriangleup = f \]

\[ a \quad b \quad c \quad f \]

\[ a \quad d \quad e \quad g \]

\[ \bigtriangleup b \bigtriangleup c \]

\[ \bigboxtimes d \bigboxtimes e \]

\[ f \]

\[ g \]

\[ h \]

\[ i \]

\[ j \]
FIGURE 2
Temporal development of the relevant knowledge pool and its influence on learning and innovation

10-year knowledge base of originating firm

Relevant knowledge pool

Innovation of originating firm

- Innovative output
- Integration of knowledge from the pool

$t-11$ to $t-2$

$t-1$

$t$