7-2016

Appropriability and the retrieval of knowledge after spillovers

Tufool ALNUAIMI
Imperial College London

Gerard GEORGE
Singapore Management University, ggeorge@smu.edu.sg

APPROPRIABILITY AND THE RETRIEVAL OF KNOWLEDGE AFTER SPILLOVERS

TUFOOL ALNUAIMI
Assistant Professor of Data Science and Innovation
Imperial College London
London, UK, SW7 2AZ
+442075945485
t.alnuaimi@imperial.ac.uk

GERARD GEORGE
Dean and Professor of Innovation and Entrepreneurship
Lee Kong Chian School of Business
Singapore Management University
Singapore 178899
+6568085122
ggeorge@smu.edu.sg

Accepted for publication at ***Strategic Management Journal***

Alnuaimi, T., George, G. 2015. Appropriability and the retrieval of knowledge after spillovers, ***Strategic Management Journal***, Advance Online (May 2015).
DOI: 10.1002/smj.2383

APPROPRIABILITY AND THE RETRIEVAL OF KNOWLEDGE AFTER SPILLOVERS

ABSTRACT

Firms create and capture value through innovation. In technology-driven firms, there has been an explicit emphasis on appropriability through imitation deterrence and cumulative inventions that build on prior firm innovation. We introduce systematic empirical evidence for a third mechanism of appropriability namely, knowledge retrieval, which is defined as the re-absorption of previously spilled knowledge. We extend previous studies which consider technological complexity and organizational coupling as predictors of appropriability by examining their impact on knowledge retrieval. We find that technological complexity has a curvilinear relationship with retrieval while organizational coupling has a negative relationship. We discuss the implications of these findings for theories of absorptive capacity, organizational design and appropriability of innovation.

MANAGERIAL SUMMARY

It is a widely held assumption that knowledge should be protected and held tightly within the firm to ensure value creation and value capture. The implicit recognition is that knowledge spillovers, or knowledge leakage, is detrimental to performance. By examining the patterns of citations among patents of 143 semiconductor firms, we study how organizational structure and technological complexity play a role. We find that moderate technological complexity also improves the odds of benefitting from spillovers. If imitation deterrence is paramount, then the optimal structure would be a tightly-coupled organization. In other instances, loosely-coupled organizations may be superior because they foster internal cumulative innovations and, if spillovers were to occur, they also maximize knowledge retrieval. Our

findings suggest that all is not lost when spillovers occur and that firms can continue to benefit in downstream innovations.

Key words: Appropriability, technological complexity, coupling, knowledge retrieval, patents, innovation, patent citations, organization design

RUNNING HEAD: Knowledge retrieval after spillovers

INTRODUCTION

A central focus within strategic management research is how firms capture value from their investments. Within this field, the extensive literature which deals with technology and innovation strategy has enhanced our understanding of the factors that allow firms to maximize profits from innovation (for a recent review, see James, Leiblein and Lu, 2013). The emphasis has been on mechanisms that prevent or deter knowledge spillovers and imitation, which are viewed as key indicators of firms' abilities to capture value from innovation (e.g., Cohen, Nelson and Walsh, 2000; James *et al*., 2013). The broader conclusion from these studies is that the prevention of knowledge spillovers increases returns to an innovation in two ways. First, it increases returns by preventing rivals from creating and profiting from similar innovations. Second, it generates a potential for new returns by creating an avenue for the firm to build on the original innovation (Ahuja, Lampert and Novelli, 2013). The emphasis of these strategies collectively has been on deterrence.

However, spillovers can generate an opportunity for firms to further profit from their innovative endeavors. Outgoing spillovers can enable access to the recipients' capabilities. It can also assist originators in understanding how advances were made using their spilled knowledge (Belenzon, 2012; Yang, Phelps and Steensma, 2010). In turn, this enables the originator to absorb knowledge that has spilled to, and has been leveraged by the recipient firm. We define this phenomenon - specifically, the re-absorption of one's own spilled knowledge that has been leveraged externally - as knowledge retrieval.

Previous empirical evidence suggests that retrieval of knowledge after spillovers is indeed a possibility (Yang *et al*., 2010), and that it often improves a firm's performance (Belenzon, 2012). This implies that the long term effect of outgoing spillovers can potentially outweigh its direct losses. In this study, we examine how mechanisms that are often employed to deter knowledge spillovers relate to the knowledge retrieval process. In doing so, we provide a more holistic view of the variations across firms in their abilities to profit from innovation.

More specifically, we explore the effects of technological complexity and organizational coupling. Previously, complexity and coupling have been shown to affect appropriability for reasons that relate to how the originating firm and rivals are able to comprehend the technology and the organizational process that it yields (e.g., Ethiraj *et al.*, 2008; Sorenson, Rivkin and Fleming, 2006). Technological complexity is a function of the extent to which the elements that make up a technology interact with one another. It determines a firm's success at using existing technologies for its subsequent inventive endeavors while preventing rival firms from doing so (Sorenson *et al.*, 2006). Organizational coupling relates to the interdependencies between the departments of an organization (Sanchez and Mahoney, 1996). It is an organizational design characteristic that could impact knowledge sharing within organizations, and also influence coordination costs in a way that will potentially deter innovative behavior. Cooperation and coordination costs vary with changing levels of organizational coupling, affecting a firm's potential for building on a current innovation (Alcacer and Zhao, 2011). Because coupling also varies the imitation-deterrence potential of firms (Ethiraj *et al*., 2008), it has been described as a key determinant of appropriability. We extend these arguments to examine how technological complexity and organizational coupling affect the retrieval of spilled knowledge.

This study contributes to two important conversations in the strategy literature. First, we add to the traditional discussions of appropriability by opening a window into a third mechanism by which firms can potentially capture value – through knowledge retrieval. Recent theoretical work proffers a complementary perspective to traditional deterrence-based appropriability by arguing that selective revealing of knowledge could offer strategic benefits to innovative firms (Alexy, George and Salter, 2013). This study offers systematic evidence that knowledge retrieval of previously spilled-over knowledge does indeed occur, with meaningful economic significance.

Secondly, previous work on appropriability mainly addresses complexity and coupling as important for capturing value from innovation through either imitation deterrence or through the generation of cumulative innovations. Different causal mechanisms link these factors to knowledge retrieval, which occurs only if the originating firm can evaluate and internalize the advances that a recipient firm has made using the spilled knowledge. By relating these factors to retrieval, we engage in conversations of how attributes of organizational and product design influence innovation and, more broadly, strategic action (e.g., Ethiraj *et al*., 2008; Gulati, Puranam and Tushman, 2012; Singh, 2008). Consequently, our focus on knowledge retrieval provides an additional avenue for scholars to consider when discussing the appropriation of value from innovation.

THEORY DEVELOPMENT

In this study, we conceptualize knowledge retrieval as the re-absorption of knowledge that has previously been spilled. When one firm's knowledge spills over to another firm, all subsequent combinations comprising the spilled knowledge that are generated by the recipient are considered to be external. Knowledge retrieval is then said to occur if the originating firm can absorb knowledge from this subset of external combinations. Prior literature has discussed the relationship between absorptive capacity and the ability of firms to internalize outside knowledge (e.g., Cohen and Levinthal, 1990). We revisit this concept in the context of knowledge retrieval.

There are three components of absorptive capacity: the identification, assimilation and exploitation of potentially useful external knowledge (Cohen and Levinthal, 1990). Of these, identifying the location of useful knowledge should occur readily in the context of knowledge retrieval. The occurrence of spillovers is an event that may naturally draw the attention of the originators and, as a result, facilitate the identification of private knowledge (Ocasio, 1997; Corredoira and Rosenkoft, 2010) and its commercial valuation (Arrow, 1996; Hoetker and Agrawal, 2007). The other two components would depend on how similar the spilled knowledge is to the original and if the originator is equipped to leverage its prior experience for retrieval.

A recent study has suggested that spilled knowledge, even after it is leveraged by a different firm, retains more similarity to the original than other sources of external knowledge (Yang *et al*., 2010). In fact, some firms even deliberately leak their knowledge to stimulate the production of complementary inventions by other firms (Alexy *et al*., 2013). As a result, the assimilation of spilled knowledge occurs more readily than other external knowledge. Still, all spillovers are not equally retrievable. Where spillovers bear more resemblance to the original knowledge, reabsorption is expected to occur more frequently (Yang *et al*., 2010). Hence, any factor which influences the extent of transformation that spilled knowledge undergoes when it is leveraged externally inevitably impacts the propensity for retrieval.

The final component, exploitation, hinges on the firm's ability to leverage prior related knowledge (Cohen and Levinthal, 1990). In the context of retrieval, related knowledge refers to the capabilities that were developed and refined to generate the original invention. The extent to which these are available is partially dependent on the retention of similarity post-spillover. Where the evolution of an invention post-spillover is such that it inherits the underlying knowledge of the original invention, relevant knowledge will be available within the firm. However, availability alone does not suffice, as the originating firm will also need knowledge management systems that facilitate access to and reuse of prior knowledge (Ahuja *et al*., 2013).

Two factors which influence the motivation to retrieve spilled knowledge, the evolution of knowledge, and the propensity to reuse existing capabilities and knowledge bases are technological complexity and organizational coupling. We will elaborate on this in the following section to show how these factors ultimately influence knowledge retrieval. It is worthwhile to note that while these are not the sole determinants of retrieval, we will focus only on them because a rich variety of literature has discussed their influence on appropriability (e.g., e.g., Rivkin, 2000; 2001; Sorenson, Rivkin and Fleming, 2008; Ethiraj *et al*., 2008). While these studies have extended our received wisdom on how firms can generate value from innovation, they consider a short-term view of knowledge spillovers, where spillovers represent value loss. In the longer-term, there is an opportunity for firms to appropriate returns from their original inventions by reabsorbing spilled knowledge. As such, examining the effect of these factors on knowledge retrieval will clarify our understanding of how technological complexity and organizational coupling relate to knowledge retrieval.

Technological Complexity and Knowledge Retrieval

Technological complexity has a multifarious relationship with appropriability. It is beneficial for innovators because it effectively reduces spillovers and precludes imitation (e.g., Rivkin, 2000; 2001). But on the other hand, it makes it problematic for the innovators to build on the invention in a future date; thereby limiting the possibility of deriving further value from it (e.g., Kotha, George and Srikanth, 2013; Sorenson *et al*., 2006; Ahuja *et al.*, 2013). The underlying mechanism linking technological complexity to appropriability relates to the difficulty of transferring complex technological knowledge across time, even within the firm boundaries (Sorenson *et al*., 2006) and the mutation of complex knowledge if and when it is reused (Levinthal and Warglien, 1999). As we mentioned previously, the extent of mutation and the possibility of knowledge reuse both influence retrieval. As such, technological complexity is expected to have an impact on knowledge retrieval.

Borrowing from previous studies (e.g., Kaufmann, 1993; Fleming and Sorenson, 2004; Simon, 1962; Sorenson, Rivkin and Fleming, 2006), we define technological complexity as a function of the level of interdependence between the components which form an invention. To illustrate, consider the following example of word processing software. A possible component of the software is the code which permits users to change the font size of a phrase. A second component could be the code that changes the spacing between lines. In a non-complex (i.e., perfectly modular) invention, components are independent; meaning that existing components can be removed or new components can be added without disrupting the overall system. For instance, if the word-processing software were perfectly modular, one could simply delete the lines of code that perform the line spacing functionality, thereby removing that feature, and still be able to use the remaining features in the software. In contrast, when components are interdependent, a programmer would have to re-write the code in order to remove that same feature while maintaining all the others.

Two other salient points can be deduced from the example of word processing software. When an invention becomes more complex, the interdependence between its components increases. This makes the system as a whole sensitive to even minor changes to the underlying components (Sanchez and Mahoney, 1996). The intricacy of how the components interact with one another is difficult to codify (Salomon and Martin, 2008). As a result, firms are limited in the extent to which they can re-use knowledge from complex inventions, particularly in new contexts. By contrast, the minimal interdependence between the components of modular inventions suggests that these adaptations could take place without loss of functionality or value to the other components (Baldwin and Clark, 1997). Secondly, because varying one component necessitates further variations to be made to other components, complex inventions change more drastically than modular inventions following attempts to modify or alter them (Ethiraj *et al.*, 2008; Levinthal, 1997; Levinthal and Warglien, 1999). As a result, when new inventions are spawned from complex inventions; they bear low resemblance to their predecessor.

The arguments so far suggest that knowledge retrieval favors low complexity. However, in order to derive value from spilled knowledge, the originator will have to use the knowledge towards a new commercial application that differs from advances made by the recipient (or any other competitor). The literature on the evolution of technologically complex systems suggests that the cumulative inventions that spawn from non-complex systems are predictable and can be arrived at independently (Levinthal and Warglien, 1999). In other words, having prior inventive knowledge does not give the originating firm an advantage because other firms will be able to leverage the spilled knowledge in the same way without access to the originator's capabilities.

In comparison to low complexity, the more intricate interactions between components when complexity rises to moderate levels makes prior inventive knowledge useful for future inventions. In comparison to high complexity, access to prior inventive knowledge will be less problematic and the recipient's advances will retain more similarity to the original invention. In creating the original (moderately complex) invention, the originating firm needed to develop a deep understanding of the components and their relationships (Alnuaimi Singh and George, 2012; George, Kotha and Zheng, 2008; Salomon and Martin, 2008). The availability of this prior knowledge makes (re)absorption easier and less costly for the originator than for other firms, who are unfamiliar with the components of the original invention (Cohen and Levinthal, 1990; Lane and Lubatkin, 1998). Therefore, we posit that:

Hypothesis 1: *Conditional on the occurrence of spillovers, the extent of knowledge retrieval by an originating firm will first increase and then decrease as the technological complexity of an invention increases.*

Organizational Coupling and Knowledge Retrieval

The mechanisms that define the relationship between organizational coupling and knowledge retrieval relate to the motivation to reabsorb spilled knowledge and accessibility of relevant inventive knowledge. In a decoupled organization, defined as one whose units or divisions function autonomously (Weick, 1976), knowledge sharing rarely occurs. In this type of organization, the distinct divisions can complete day-to-day activities without needing to coordinate activities or communicate with members from other divisions (Sanchez and Mahoney, 1996). By contrast, organizational coupling refers to the extent to which units depend on one another for the completion of tasks. It follows that organization design governs the flow of knowledge within the boundaries of a firm (Alcacer and Zhao, 2011; Puranam, Raveendran and Kundsen, 2012; Siggelkow and Levinthal, 2003).

There are a number of performance advantages that decoupled organizations enjoy. Autonomous divisions can adapt to a changing business environment without being constrained by the needs of other divisions (Galunic and Eisenhardt, 2001), allowing them to respond to pending needs in the domain which they cater to with more efficiency. In the long run, this would also contribute to the creation of deep pockets of knowledge relevant to each of their domains (Yayavaram and Ahuja, 2008), which would differ from the knowledge base of the other divisions (Fang, Lee and Schilling, 2010). Knowledge depth fosters invention and innovation because these divisions would have a more thorough understanding of the innovation opportunities in that domain and how to best approach them (George *et al*., 2008; Katila and Ahuja, 2002). The downside, however, is that divisions will not be able to access the knowledge generated by other divisions, limiting the range of cumulative innovations the any one division can produce (Katila and Ahuja, 2002; Ahuja *et al.*, 2013).

Tight coupling between distinct divisions raises intra-firm causal ambiguity. Extensive interdependence between the divisions makes it difficult for employees to understand the causes of successes and failures in prior projects (King, 2007; Reed and DeFillippi, 1990). Hence, even if the resources and skills that are needed for subsequent innovations are available within the firm there will be a lack of understanding of how these can be combined successfully, which can impede the re-use of prior knowledge for subsequent innovations (King, 2007). For instance, the inability to appraise the value of existing resources would inevitably decrease the motivation of employees to understand their full scope and how they can be adapted to capture future innovation opportunities. In addition, internal causal ambiguity increases the cost of searching for and mobilizing relevant knowledge such that it can be useful for innovative efforts across the firm. Taken together, these arguments suggest that maintaining moderate levels of organizational coupling would increase knowledge transfer across time and thus, increase the propensity for cumulative innovation (Ahuja *et al*., 2013).

We previously defined knowledge retrieval as a form of cumulative or cumulative innovation, but one which requires the additional effort of re-absorbing spilled knowledge. The division or team which is interested in re-using this knowledge would have to first understand the content and the structure of the external combination, and then incorporate it into a novel and commercially viable application. In a loosely coupled organization, the specialized knowledge of the division in which the original invention was created makes it easier for that division (but not the other divisions) of the originating firm to identify and assimilate the external combination (Cohen and Levinthal, 1990; Zahra and George, 2002). Since the search cost for accessing prior knowledge across divisions will be high, the original division will face difficulties acquiring relevant knowledge from other divisions. Likewise, other divisions wishing to build on the spilled knowledge won't have access to or be aware of the knowledge underlying the original invention. Hence, the number of useful and commercially viable applications that can be generated using the spilled knowledge will be limited. These search costs and knowledge integration challenges can be managed when an organization's divisions are more interdependent – but only to a moderate level.

If the structure of the organization transforms such that divisions are tightly coupled, intra-firm causal ambiguity makes it difficult to re-use previously generated knowledge (Ethiraj and Levinthal, 2004). Additionally, tightly coupled organizations, divisions rarely have the authority to make decisions independently. Thus, if re-absorption is viewed by one division as opportune, it will need to first seek approval from a decision making authority. Since decisions are not made on the basis of what is optimal for each division but rather, based on what is optimal for the organizational as a whole (Sigglekow and Levinthal, 2003) this would reduce motivation for knowledge retrieval. Thus, we hypothesize that:

> ***Hypothesis 2:*** *Conditional on the occurrence of spillovers, the extent of knowledge retrieval by an originating firm will first increase and then decrease as organizational coupling increases.*

DATA AND METHOD

We rely on patent and patent citation data to track the flow of knowledge, and needed an industry that is technology intensive and where firms readily patent their inventions. The US semiconductor industry meets both requirements. Previous studies have noted that US semiconductor firms innovate considerably (Stuart, 2000) and have high patenting propensities (Hall and Ziedonis, 2001). Therefore, in this context, patent data is a more suited proxy for the innovative activities of firms. We populated a list of firms using three different sources. The first was the list of firms that was used by Hall and Ziedonis (2001) to investigate the factors that drive patenting in the US semiconductor industry between 1975 and 1995. This list comprised of 95 publicly traded firms, all of which had a COMPUSTAT record and owned a USPTO patent. We consider firms that have patented even after 1995, and therefore relied on two additional sources. We used the Directory of American Firms Operating in Foreign Countries, which lists 502 US semiconductor firms with substantial investments outside the US. To ensure comprehensiveness in our sample, we used the annual publication by the iSuppli Corporation which ranks US semiconductor firms. Using these three sources, we constructed a list containing 550 unique semiconductor firms headquartered in the US.

A major challenge that confronts research that utilizes patent data is matching each firm to all of the patents it applied for, because there is not a unique assignee identifier in the USPTO database. Instead, a firm's name can appear in full (e.g. International Business Machines), with an alternative spelling (e.g. International Business Machines Corp.), as an acronym (e.g. IBM) or even as the name of one of its foreign subsidiaries. Several steps were taken to ensure that we accurately aggregated each firm's patents. Firstly, data that is made available from the NBER patent project was used to match USPTO assignees with a unique numerical identifier[1]. Secondly, each variation in the names of the subsidiaries of the 502 semiconductor firms, retrieved from the Directory of American Firms Operating in Foreign Countries, was compared against the names of the 247,309 assignees that were granted a USPTO patent during the time-period 1975-2008.

Sample of Patents

Following these above steps, we identified 463 firms that had been granted at least one USPTO patent between 1975 and 2008. Our sample only included patents applied from

[1] This data is available from two sources: https://sites.google.com/site/patentdataproject/Home and http://www.nber.org/patents/. The first source is used for the purpose of this research as it is a more up to date version.

1985, which leaves a 10 year window (1975-1984) to derive the variables that rely on a firm's historical activities (these will be detailed in the next section). Furthermore, to minimize right censoring of the data on knowledge spillovers, the final observations were for patents applied for in 1999, as this allows us to measure forward citations until 2010[2]. Thus our panel spanned a 15 year time period: 1985 to 1999, during which 44,959 patents were applied for by 144 US semiconductor firms.

The two hypotheses propose a relationship between knowledge retrieval and each of technological complexity and organizational coupling conditional on the occurrence of spillovers. Previous research has used external citations, defined as a citation made by an entity other than the originator of a patent, as a proxy for knowledge spillovers (e.g., Jaffe, Trajtenberg and Henderson, 1993). Therefore, to test these hypotheses, we only consider patents which have received at least one external citation, which leaves a sample of 39,538 patents assigned to 142 firms.

Dependent Variable

We use patent data to construct our dependent variables, *Knowledge Retrieval* and *Knowledge Retrieval Frequency*, which are measured at the level of each patent. The two variables capture two related aspects of knowledge retrieval: whether spilled knowledge is retrieved and the number of times spilled knowledge is used in a firm's subsequent patents.

We follow a method similar to Belenzon (2012), who examines the effect of knowledge retrieval on a firm's stock market value. Knowledge retrieval is defined as the extent to which an originating firm is able to build on knowledge that has spilled over and has been leveraged by external firms. The variable, measured at the patent level, is calculated using the following steps. For each focal patent assigned to an original firm, we isolate each forward citation that

[2] We supplemented our core dataset with patent data that was made available by Lai, D'Amour, Yu, Sun and Fleming (2011) which is available at:
http://dvn.iq.harvard.edu/dvn/dv/patent/faces/study/StudyPage.xhtml?studyId=70546&versionNumber=1

is made to that patent by other (external) firms during the time period: *t+1 to t+5*. In Figure 1, these external citations are jointly marked as Group 1.

Next, we examine the forward citations that are made to the citing patent during the five years which follow its application date. Reabsorption is said to occur if any of the citations made to the patents in Group 2 are owned by original firm (Belenzon, 2012). In Figure 1, the patents in Group 2 and Group 3 are citations to the external patents in Group 1. However, only the patents in Group 2 are owned by the originating firm. Knowledge retrieval is measured as the number of patents in Group 1 that are cited by the originating firm (Group 2). In other words, it is the number of spilled patents that are re-absorbed by the firm. Knowledge retrieval frequency is a count of the number of patents in Group 2, which represents the number of times that spilled knowledge (Group 1) is reabsorbed by the firm.

Similar to other studies (e.g. Fleming and Sorenson, 2004), we observe forward citations during the five-year window that follows each patent's application date because during this time period, a patent accumulates the majority of citations. Furthermore, using a fixed window allows us to account for the fact that older patents have a longer exposure time during which they can accumulate citations. By observing citations within a fixed window, we can compare the forward citations received by patents from different years.

------ Insert Figure 1 about here-------

Independent Variables

Technological complexity. We measure the technological complexity of the original patent as we are interested in how technologies developed by originating firms help to maximize retrieval. Thus, it is a variable that is measured at the patent level. Patent documents list multiple subclasses, each of which can be considered as a component in the technology (e.g. Fleming and Sorenson, 2001, 2004, Sorenson *et al.*, 2006). Since technological complexity should capture the difficulty of combining components, a patent can be defined as more complex if its components have not been previously integrated with a wide variety of other components. Alternatively, a technology which contains components which have been readily “mixed and matched” with an array of different components is considered to be non-complex (Sanchez and Mahoney, 1996). We use the equations used in a number of previous studies to calculate the complexity (e.g. Fleming and Sorenson, 2001, 2004, Sorenson *et al.*, 2006), as this measure has been shown to be highly correlated with what inventors perceive to be complex technologies according to interview data (Fleming and Sorenson, 2004).

The variable is measured in two steps. First, we calculate the ease of recombining the subclasses that appear on each focal patent. For each subclass that appears on the focal patent, we first identify all USPTO patents that have that subclass. The ease of recombining that subclass is measured as the number of distinct subclasses that appeared with the focal subclass in all these patents divided by the number of patents featuring that subclass. Next, the focal patent’s Complexity is calculated as the total number of distinct subclasses listed on that patent divided by their cumulative ease of combination. The variable construction is analogous to prior studies (e.g. Fleming and Sorenson, 2004; Sorenson *et al.*, 2006):

$$\text{Ease of recombining subclass } s = E_s = \frac{\text{Number of subclasses combined with subclass } s}{\text{Number of previous patents in subclass } s}$$

$$\text{Complexity}_i = \frac{\text{Number of subclasses in patent i}}{\sum_{s \in i} E_s}$$

Organizational coupling. Organizational coupling is measured at the firm-level as a function of interdependencies between units of an organization (Ethiraj *et al.*, 2008). Because we are limited by the information available in patent data, we define units based on the geographical region if the inventors listed on a patent, where a region is defined as a state if the inventor was from the US and a country otherwise (e.g., Singh, 2008). Interdependencies between units of an organization are said to occur if the inventors listed on a patent are from different regions, as suggested in Alcacer and Zhao (2011).

We closely follow the definition in Yayavaram & Ahuja (2009) to measure organizational coupling, which is a firm-level variable. For each firm *i*, we isolate the patents in a firm's portfolio that were applied for during the 3 years preceding the application date, *t*, of the focal patent. In each firm *i*, coupling between units *j* and *k* is calculated as:

$$L_{i,j-k,t-3tot-1} = a/a+b+c$$

where a is the number of patents assigned to inventors from both units j and k, b is the number of patents that has inventors from unit j but not from unit k, and c is the number of patents that has inventors from unit k but not from unit j. Next, we calculate firm coupling as the average coupling between all pairs of units in the firm.

Our results will depict the effect of Coupling when it is measured during a three-year window preceding the year of patenting. However, in additional analyses that are not presented here, we use a five-year window to calculate coupling, and are results remain largely the same.

Firm-level control variables

Technological opportunity reflects the extent to which an originating firm conducts R&D in technological domains with high patenting activity, which can explain the competition that an originating firm will face when attempting to re-absorb spilled knowledge. We calculate the variable at the firm level during the application year of the focal patent using the following equation from Yang *et al.* (2010):

$$\text{Technological opportunity} = \sum_c p_c \times P_{j,c}$$

$$\text{Technological opportunity} = \sum_c (p_c \times \frac{p_{j,c}}{P_{j,c}})$$

Where p_c refers to the total number of patents in technological class *c* that were applied for in year *t*, $p_{j,c}$ is the total number of patents in technological class *c* and and $P_{j,c}$ refers to the proportion of a firm j's patents that in year t.

Slack is calculated at the firm level as the ratio of current assets to current liabilities of a firm during the application date of the focal patent (Yang *et al.*, 2010)**.** Slack, which increases managerial flexibility (George, 2005), should facilitate knowledge retrieval. The ratio was calculated using COMPUSTAT data, and is divided by 100 to modify its scale.

Absorptive capacity is measured as the originating firm's R&D expenditure (in billions of U.S. dollars) during the application date of the focal patent (e.g., Yang *et al.*, 2010). R&D expenditure is used as a proxy for absorptive capacity because of the correlation between the two variables (Cohen and Levinthal, 1990). Annual R&D expenditures were retrieved from COMPUSTAT. We modified the scale of this variable by dividing by 1,000.

Firm age is calculated as the number of years elapsed between the originating firm's incorporation and the application year of the focal patent. It is included to account for the possibility that older are more experienced at managing spillovers and absorbing external knowledge.

Organizational units counts the total number units within the firm that have applied for a patent during the 3 years preceding the application date, *t*, of the focal patent.

Firm size is the total number of employees in the originating firms during the focal patent's application date. Firms with more employees may have more flexibility when it comes to organizing teams to meet project needs. In terms of knowledge retrieval, this could mean that larger firms may find it easier to deploy teams that are capable of re-absorbing spilled knowledge. We modified the scale of this variable by dividing by 1,000.

Patent-level control variables

Team size, which is measured as the number of inventors on the focal patent, influences the number of citations that a patent receives (Singh and Fleming, 2010). It is included as a control because our dependent variable also depends on a patent's forward citations.

Internal focus is measured by the number of citations that the focal patent makes to prior patents owned by the originating firm as a proportion of the total number of backward citations made in the patent application (Hoetker and Agrawal, 2007). High values for internal focus may reflect that a technology is more related to the originating firm's knowledge base, and these technologies may be easier to re-absorb.

Technological maturity is calculated as ratio of citations that the focal patent makes to prior art to the number of claims that it makes. Patents in technological fields that are more mature typically make more backward citations per claim (Hoetker and Agrawal, 2007; Lanjouw and Schankerman, 2003). Mature technologies are typically easier to understand (Sorenson and Stuart, 2000), which makes retrieval more straightforward from the perspective of the originating firm. However, they may also be less desirable in the marketplace (Hoetker and Agrawal, 2007), which would reduce propensity for retrieval.

Subclasses. The number of subclasses that are listed on a patent are used as a proxy for the number of distinct components that compose the invention (Fleming and Sorenson, 2001; 2004; Sorenson *et al.*, 2006). It is controlled for in our empirical models because of its effect on the complexity of an invention.

External impact is measured as the number of citations (excluding self-citations) that the original patent receives during the 5 year window following its application date. It correlates with a number of other measures that reflect a patent's value, such the patent's contribution to a firm's market value (Hall *et al.*, 2005) and expert evaluation of the patent's value (Albert, Avery and Narin, 1991). External impact is expected to be positively related to Retrieval frequency, and is therefore included as a control variable.

Citation lag is measured as the average difference between application year of the forward citations a patent receives and the application year of the patent. It is included to account for unobserved heterogeneity that may influence the citation rates of patents (Hoetker and Agrawal, 2007). The empirical models include year dummy variables to account for the differences in citation propensities of patents that are applied for in different years (Hall *et al.*, 2001).

Time and technology controls

Although we use a fixed-window during which we observe forward citations, other temporal factors may also influence the extent of forward citations. As a result, comparing forward citations across patents from different years would be inappropriate. To account for this, year dummies are included in all models. In our analyses, year dummies are based on the application year of the focal patent because this more closely resembles the time during which the inventive activities took place. In an analogous manner, patents that fall under different technological categories have different propensities of being cited (Hall *et al.*, 2001). Therefore, technology dummies are also included, which reflect the one-digit technological category as defined by Hall *et al.* (2001).

Empirical model

Both dependent variable, *Retrieval* and *Retrieval Frequency*, are count variables that that are typically over-dispersed. Therefore, the hypotheses were tested using negative binomial regressions. An obvious concern is that the firms in the sample differ systematically in unobserved ways, making a form of omitted variable bias a concern for regression analysis. We account for unobserved firm-level heterogeneity by using negative binomial models with firm fixed effects.

In the first set of regressions, where the dependent variable is Retrieval (i.e., the external patents which the originating firm later builds on), we use *External Impact* as an exposure term. The use of an exposure term allows us to measure the proportion of spilled patents that are retrieved. In the next set of regression, the dependent variable is Retrieval Frequency. Here, we count the number of times the external patents are cited by the originating firm. In this case, the exposure term is *Spillover Pool* which is a count of the total number of citations made to the external patents which build on the originator's knowledge (Groups 2 and 3 in Figure 1). In the latter set of regressions, external impact is included as a control variable because affects the number of the patents in the spillover pool.

RESULTS

Table 1 reports the descriptive statistics and the pairwise correlation coefficients for all the main variables. Tables 2 and 3 report the regression results with Retrieval and Retrieval Frequency as dependent variables, respectively. We ran similar regressions for both dependent variables. In the tables, Model 1 is the baseline model. To test the first hypothesis, Models 2 introduces complexity and its squared term. Model 3 introduces coupling and its squared term to the baseline model in order to test hypothesis 2. Finally, Model 4 is the full model. We use the results of this model for our discussions. It is worthwhile to note that although they are not reported, year and technology dummies are included in all models, and they are jointly significant. Additionally, a Hausman test (1978) was significant, suggesting that fixed effects models were more appropriate than random effects models.

------ Insert Table 1 and Table 2 about here ------

Hypothesis 1 proposes a curvilinear relationship between complexity and knowledge retrieval. In both tables, the coefficient of technological complexity is positive and significant ($b=0.160$, $p < 0.001$) and the coefficient of its squared term is negative and significant (b=-

.024, $p < 0.001$). The relationship is presented graphically in Figure 2; which is plotted by using the values in Table 2 by varying complexity and holding all other variables constant at their mean values. The figure plots the relationship for values of complexity that range from two standard deviations below the mean to two standard deviations above the mean. The vertical lines represent low (one standard deviation below the mean), mean and high (one standard deviation above the mean) values for complexity. Up until a high level of complexity, the relationship shows a non-monotonic increase. For example, the percentage of retrieval is 1.30% higher at mean levels of complexity than at low levels of complexity. By contrast, when complexity increases from mean to high levels, the percentage increase is only 0.58%. Further increases to complexity lead to a reduction in knowledge retrieval. Replicating the figure for the results in Table 3 depict a very similar relationship. Thus, Hypothesis 1 is supported.

----------- Insert Figure 2 ------------

Our second hypothesis proposes a curvilinear relationship between organizational coupling and knowledge retrieval. In Table 2, the coefficient for organizational coupling is negative and significant ($b=-2.52$, $p < 0.001$) and its quadratic term is positive and significant ($b=2.305$, $p < 0.001$). A similar pattern is seen for retrieval frequency (Table 3). These coefficients would suggest a U-shaped relationship between organizational coupling and knowledge retrieval, with an inflection point occurring when coupling takes a value of approximately 0.5. However, 96% of the data lie in the (0, 0.046) region. The positive and significant squared term could be driven by outliers. To examine the functional form more closely, we tested the significance of the slope (Aiken and West, 1991). As shown in Figure 2, the pattern is negative and significant for that range of the data but not beyond. Thus, the relationship is negative and knowledge retrieval is maximized at low levels of organizational coupling rather than the hypothesized moderate levels.

The regression estimates also show how other attributes can improve the propensity for knowledge retrieval and retrieval frequency (Tables 2 and 3). Specifically, larger firms and those with more slack resources are better positioned for knowledge retrieval (b=.02, p<.01). At the level of the technology, retrieval is higher for inventions that build on relatively more internal knowledge and those developed by larger teams (b=.593, p<.01).

Robustness checks

We implemented further analyses to ensure the robustness of the results. First, we repeated the fixed effects negative binomial regression, but only for the subsample of firms which have a patent during the last observation period during which we calculate retrieval, namely 2005 to 2010. In analysing this subsample, we ensure that our analysis is consistent for the firms which did not exit the semiconductor industry. Of the original 144 firms, 102 fit this criterion. Next, we ran an unconditional fixed effects negative binomial regression with robust standard errors clustered at the original firm level. We cluster standard errors because we observe multiple patents per firm, and this could lead to inconsistent standard errors in a standard regression. For this regression, we consider the four firms in our sample which each account for more than 10% of total patents. Collectively, these firms own more than half of the patents in our dataset. We repeat this analysis using zero-inflated negative binomial regressions because both dependent variables take on a value of 0 in 70% of the cases. The outcomes of all these analyses were consistent with our main results.

DISCUSSION AND FUTURE RESEARCH

In a review, Ahuja *et al.*, (2013: 248) defined generative appropriability as firm's ability to maximize profits from an original invention by building on it in the future: "*future inventions could be enhanced or improved versions of the original invention…or derived* *inventions that use the ideas of the original invention in a related but complementary market …or even in unrelated markets*". Previous literature has focused on how the aforementioned methods of maximizing returns to an original invention can be attained in one of two ways: either by erecting mechanisms which limit knowledge spillovers or by generating new inventions that extend the original invention. By preventing spillovers, the firm can capture the largest share of profits from an invention. Furthermore, using the invention in the future generates a new avenue for profiting from the original invention (Ahuja *et al*., 2013). We augment this literature by recognizing that outgoing spillovers can generate a positive return to the original invention if the originator is able to re-absorb knowledge after it spills to, and is leveraged by an external firm (Alexy, *et al*., 2013; Yang, *et al*., 2010).

We find systematic evidence for a third mechanism of value capture, the retrieval of spilled knowledge. We examine the impact of two factors, namely technological complexity and organizational coupling. Previous studies have related these factors to appropriability by examining when they may deter knowledge spillovers and imitation (e.g., Ethiraj *et al.*, 2008; Rivkin, 2000; 2001). We build on these studies by exploring how these factors affect knowledge retrieval and, in doing so, are able to more precisely show how firms can benefit from their inventions. These findings inspire further discussions in theories of absorptive capacity, organizational and product design and the appropriability of innovation.

Technological complexity and knowledge retrieval

We first examined the impact of technological complexity on knowledge retrieval. A number of previous studies have discussed the benefits of moderately complex technologies (e.g., Sorenson *et al.*, 2006). Ethiraj *et al.* (2008) use a simulation to examine the trade-off between the performance benefits of innovations and their susceptibility to imitation for three types of systems: non-modular systems, nearly modular systems and complex systems. Their findings suggest that nearly modular systems provide the better benefits for incremental innovations in comparison to complex systems and better protection against imitation in comparison to non-modular systems. The concept of nearly modular systems is comparable to that of moderate complexity that is used in this study. Likewise, Sorenson *et al.* (2006) find that that, at moderate complexity, the difference between intentional and non-intentional knowledge transfer is maximized.

Our results show that as complexity increases from low to high, the marginal increase in knowledge retrieval diminishes. Thus, if one were to consider knowledge retrieval in isolation, there is a benefit to increasing complexity to high levels. In the broader context of appropriability, moderate technological complexity is ideal. At this level, a firm can deter spillovers and imitation, build on the original invention and, if spillovers were to occur, retrieve spilled knowledge. In so doing, our study contributes to discussions on how technological complexity affects subsequent capability of firms to benefit from their earlier inventions through knowledge retrieval.

Organizational design and knowledge retrieval

We add to the literature on the importance of organizational design for innovation and, more broadly, strategic action (e.g., Gulati *et al*., 2012; Gruber *et al.*, 2015; Singh, 2008). Recent empirical work has shown how the success of radical innovations is influenced by coordination costs and design attributes (e.g., Kotha *et al.*, 2013). Our study extends these findings by showing how organization design influences appropriability through knowledge retrieval. We hypothesized that the relationship between organizational coupling and retrieval would follow an inverted-U shaped curve. Specifically, at this level, divisions that are interested in leveraging the spilled knowledge will have access to complementary knowledge that may exist within other divisions. By contrast, inter-divisional information flow is limited in decoupled organizations, making it difficult to access knowledge that exists elsewhere in the firm (Ahuja *et al.*, 2013; Haas, Criscuolo and George, 2015). In tightly coupled organizations, ambiguities relating to the generation of the original invention would naturally arise. To overcome these uncertainties and learn from original inventive endeavors, a firm would incur high coordination costs. This, in turn, may reduce motivation for embarking on knowledge retrieval.

While we theorized that the relationship between organizational coupling and retrieval would follow an inverted-U shaped curve, we found that the relationship was negative. It is worthwhile to note, however, that only 6.6% of the patents in our data have a value of 0 for coupling. Thus, firms in our dataset tended to have loosely-coupled structures rather than decoupled structures. The broader results of this finding on appropriability are as follows. If imitation deterrence is paramount, then the most optimal structure would be a tightly coupled organization (Ethiraj *et al.*, 2008). In other instances, loosely-coupled organizations may be superior because they foster internal cumulative innovations (Ahuja *et al.*, 2013) and, if spillovers were to occur, they also maximize knowledge retrieval.

Limitations and future research directions

In our study, we do not distinguish between knowledge that spills inadvertently from knowledge that is deliberately revealed. Scholars who have focused on selective revealing (e.g., Alexy *et al.*, 2013; Harhoff, Henkel and von Hippel, 2003), have discussed how organizations selectively reveal information to signal development pathways they will likely pursue. Such selective revealing may help other inventors join the focal firm in creating an innovation ecosystem that makes technical advances. Future research which distinguishes between deliberate and non-deliberate spillovers will be able to enrich our understanding of how knowledge retrieval occurs in these differing contexts.

Our study focused on the originating firm. In doing so, we did not regard inter-organizational relationships, such as strategic alliances. Inter-organizational relationships can facilitate the transfer and integration of knowledge across firm boundaries (e.g., Puranam and

Srikanth, 2007). Future research may wish to revisit the factors that relate to knowledge retrieval by examining the influence of various forms of inter-firm relationships.

We used patent and citation data for our study which is not without its inherent limitations. Firstly, two-thirds of the citations are added by USPTO patent examiners and these may not reflect knowledge flows (Alcacer and Gittelman, 2006). Data on examiner added citations is not available before 2001; rendering our measure for knowledge retrieval correlated but noisy. Secondly, by considering only USPTO patents, we do not consider patents that are granted by other agencies and non-patented inventions. For this reason, our empirical results should be considered a lower bound for the extent of knowledge retrieval.

As a final point, our empirical models included a number of tests to check the robustness of our results. Still, biases could arise due to the endogenous nature of some of the variables. The firm-related factors explored in this study may have been organized specifically to promote knowledge retrieval. A factor that mitigates, but do not eliminate, endogeneity concerns is that at least a portion of the spillovers that we observe are non-deliberate; meaning that at least some firms in our sample did not organize specifically to re-absorb knowledge at a later stage.

Conclusion

Alongside mitigating spillovers and generating new inventions from prior ones, we describe how knowledge retrieval can also explain variations across firms in their abilities to capture value from invention. This study contributes to the emergent conversation on using knowledge retrieval as a strategy to regain benefits from knowledge spillovers. Our study shows how technological and organizational characteristics that have been examined in the context of spillovers and cumulative inventions correspondingly affect knowledge retrieval. In doing so, we provide a more comprehensive account of how these factors relate to appropriability.

REFERENCES

Ahuja, G., Lampert, C., & Novelli, E. 2013. The second face of appropriability: Generative appropriability and its determinants. ***Academy of Management Review***, 38(2): 248–269.

Aiken, L. S., & West, S. G. 1991. ***Multiple regression: Testing and interpreting interactions.*** Thousand Oaks, CA: Sage.

Albert, M. B., Avery, D., Narin, F., & McAllister, P. 1991. Direct validation of citation counts as indicators of industrially important patents. ***Research Policy***, 20(3): 251–259.

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

Alcácer, J., & Zhao, M. 2012. Local R&D strategies and multilocation firms: The role of internal linkages. ***Management Science***, 58(4): 734-753.

Alexy, O., George, G., & Salter, A. J. 2013. Cui Bono? The selective revealing of knowledge and its implications for innovative activity. ***Academy of Management Review***, 38(2): 270–291.

Alnuaimi, T., Singh, J., George, G. 2012. Not with my own: Long term effects of cross-country collaboration on subsidiary innovation, ***Journal of Economic Geography***, 12(5): 943-968.

Arrow, K. 1996. Technical information and industrial structure. ***Industrial and Corporate Change***, 5(2): 645-652.

Baldwin, C. Y., & Clark, K. B. 1997. Managing in an Age of Modularity. ***Harvard Business Review***, 75(5): 84–93.

Belenzon, S. 2012. Cumulative innovation and market value: Evidence from patent citations. ***The Economic Journal***, 122(559): 265–285.

Cohen, W., & Levinthal, D. 1990. Absorptive capacity: A new perspective on learning and innovation. ***Administrative Science Quarterly***, 35: 128–152.

Cohen, W., Nelson, R & Walsh, J. 2000. Protecting Their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (or Not). ***NBER working paper***.

Corredoira, R. A., & Rosenkopf, L. 2010. Should auld acquaintance be forgot? The reverse transfer of knowledge through mobility ties. ***Strategic Management Journal***, 31(2): 159–181.

Ethiraj, S.K. & Levinthal, D., 2004. Modularity and innovation in complex systems. ***Management Science***, 50(2), pp.159-173.

Ethiraj, S. K., Levinthal, D., & Roy, R. R. 2008. The dual role of modularity: Innovation and imitation. ***Management Science***, 54(5): 939–955.

Fang, C., Lee, J., & Schilling, M. A. 2010. Balancing exploration and exploitation through structural design: The isolation of subgroups and organizational learning. ***Organization Science***, 21(3): 625–642.

Fleming, L., & Sorenson, O. 2001. Technology as a complex adaptive system: evidence from patent data. ***Research Policy***, 30(7): 1019–1039.

Fleming, L., & Sorenson, O. 2004. Science as a map in technological search. ***Strategic Management Journal***, 25(8-9): 909–928.

Galunic, D. C., & Eisenhardt, K. M. 2001. Architectural innovation and modular corporate forms. ***Academy of Management Journal***, 6: 1229–1249.

George, G. 2005. Slack resources and the performance of privately held firms. ***Academy of Management Journal***, 48(4): 661-676.

George, G., Kotha, R., & Zheng, Y. 2008. Entry into insular domains: A longitudinal study of knowledge structuration and innovation in biotechnology firms. ***Journal of***

Management Studies, 45(8): 1448–1474.

Gruber, M., Leon, N., George, G., & Thompson, P. 2015. Managing by design. ***Academy of Management Journal***, 58(1): 1-7.

Gulati, R., Puranam, P., Tushman, M. 2012. Meta-organization design: Rethinking design in interorganizational and community contexts. ***Strategic Management Journal***, 33: 571-586.

Haas, M.R., Criscuolo, P., & George, G. 2015. Which problems to solve? Online knowledge sharing and attention allocation in organizations. ***Academy of Management Journal***, 58(3): in press.

Hall, B. H., Jaffe, A. B., & Trajtenberg, M. 2001. ***Market Value and Patent Citations: A First Look***. Department of Economics, Institute for Business and Economic Research: University of California at Berkeley.

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

Hall, B. H., & Ziedonis, R. H. 2001. The patent paradox revisited: An empirical study of patenting in the U.S. semiconductor industry, 1979-1995. ***The RAND Journal of Economics***, 32(1): 101–128.

Hausman, J.A., Specification tests in econometrics. ***Econometrica***, 46(6):1251– 1272.

Harhoff, D., Henkel, J., & von Hippel, E. 2003. Profiting from voluntary information spillovers: how users benefit by freely revealing their innovations. ***Research Policy***, 32(10): 1753–1769.

Hoetker, G., & Agarwal, R. 2007. Death hurts, but it isn’t fatal: the postexit diffusion of knowledge created by innovative companies. ***Academy of Management Journal***, 50(2): 446–467.

Jaffe, A. B., Trajtenberg, M., & Henderson, R. 1993. Geographic Localization of Knowledge

Spillovers as Evidenced by Patent Citations. ***Quarterly Journal of Economics***, 108(3): 577–598.

James, S., Leiblein, M., Lu, S. 2013. How firms capture value from their innovations ***Journal of Management***, 39(5): 1123-1155.

Katila, R., & Ahuja, G. 2002. Something old, something new: A longitudinal study of search behavior and new product introduction. ***Academy of Management Journal,*** 45(6): 1183-1194.

Kauffman, S. 1993. ***The Origins of Order***. New York: Oxford University Press.

King, A.W. 2007. Disentangling interfirm and intrafirm causal ambiguity: A conceptual. model of causal ambiguity and sustainable competitive advantage. ***Academy of Management Review*** 32(1): 156-178.

Kotha, R., George, G., & Srikanth, K. 2013. Bridging the mutual knowledge gap: Coordination and the commercialization of science. ***Academy of Management Journal***, 56: 498-524.

Lane, P. J., & Lubatkin, M. 1998. Relative absorptive capacity and interorganizational learning. ***Strategic Management Journal***, 19(5): 461–477.

Lanjouw, J. O., & Schankerman, M. 2003. Enforcement of patent rights in the United States. In W. M. Cohen & S. A. Merill (Eds.), ***Patents in the knowledge-based economy***: 145–179. Washington, DC: National Academies Press.

Levinthal, D. A. 1997. Adaptation on rugged landscapes. ***Management Science***, 43(7): 934–950.

Levinthal, D. A., & Warglien, M. 1999. Landscape design: designing for local action in complex worlds. ***Organization Science***, 10(3): 342–357.

Ocasio, W. 1997. Towards an attention-based view of the firm. ***Strategic Management Journal***, 18(S1): 187–206.

Puranam, P., Raveendran, M., & Knudsen, T. 2012. Organization Design: The Epistemic Interdependence Perspective. ***Academy of Management Review***, 37(3): 419–440.

Puranam, P., & Srikanth, K. 2007. What they know vs. what they do: how acquirers leverage technology acquisitions. ***Strategic Management Journal***, 28(8): 805–825.

Reed, R. & DeFillippi, R. 1990. *Causal ambiguity, barriers to imitation* and sustainable competitive advantage. ***Academy of Management Review***, 15(1): 88-102.

Rivkin, J. W. 2000. Imitation of complex strategies. ***Management Science***, 46(6): 824–844.

Rivkin, J. W. 2001. Reproducing knowledge: Replication without imitation at moderate complexity. ***Organization Science***, 12(3): 274–293.

Salomon, R., & Martin, X. 2008. Learning, knowledge transfer, and technology implementation performance: A study of time-to-build in the global semiconductor industry. ***Management Science***, 54(7): 1266 –1280.

Sanchez, R., & Mahoney, J. T. 1996. Modularity, flexibility, and knowledge management in product and organization design. ***Strategic Management Journal***, 17(Winter): 63–76.

Siggelkow, N., & Levinthal, D. A. 2003. Temporarily divide to conquer: Centralized, decentralized, and reintegrated organizational approaches to exploration and adaptation. ***Organization Science***, 14(6): 650–669.

Simon, H. A. 1962. The architecture of complexity. ***Proceedings of the American Philosophical Society***, 106(6): 467–482.

Singh, J. 2008. Distributed R&D, cross-regional knowledge integration and quality of innovative output. ***Research Policy***, 37(1): 77–96.

Singh, J., & Fleming, L. 2010. Lone inventors as sources of breakthroughs: myth or reality? ***Management Science***, 56(1): 41–56.

Sorenson, J. B., & Stuart, T. E. 2000. Aging, obsolescence, and organizational innovation. ***Administrative Science Quarterly***, 45(1): 81–112.

Sorenson, O., Rivkin, J. W., & Fleming, L. 2006. Complexity, networks and knowledge flow. ***Research Policy***, 35(7): 994–1017.

Stuart, T. 2000. Inter-organizational alliances and the performance of firms: A study of growth and innovation rates in a high-technology industry. ***Strategic Management Journal***, 21(8): 791–811.

Weick, K. E. 1976. Educational organizations as loosely coupled systems. ***Administrative Science Quarterly***, 21(1): 1–19.

Yang, H., Phelps, C., & Steensma, K. 2010. Learning from what others have learned from you: The effects of knowledge spillover on originating firms. ***Academy of Management Journal***, 54(2): 371 – 389.

Yayavaram, S., Ahuja, G. 2008. Decomposability in knowledge structures and its impact on the usefulness of inventions and knowledge-base malleability. ***Administrative Science Quarterly***, 53: 333-362

Zahra, S., & George, G. 2002. Absorptive capacity: A review, reconceptualization and extension. ***Academy of Management Review***, 27(2): 185–203.

TABLE 1

Summary statistics and correlations

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	Retrieval	1														
2	Retrieval frequency	0.36	1													
3	Technological opportunity	0	-0.01	1												
4	Slack	0	-0.01	0.02	1											
5	Absorptive capacity	0.01	-0.01	-0.02	0	1										
6	Firm age	-0.06	-0.03	-0.06	0.09	-0.08	1									
7	Organizational units	-0.07	-0.03	0.14	0	-0.11	0.74	1								
8	Firm size	-0.01	-0.02	0.18	0.01	-0.08	0.73	0.73	1							
9	Team size	0.06	0.01	0.02	0.02	0.01	0	0.01	0.01	1						
10	Internal focus	-0.01	-0.01	-0.06	-0.03	0.02	-0.06	-0.02	-0.08	0.05	1					
11	Technological maturity	0	0.01	-0.02	0.01	-0.01	-0.03	-0.04	-0.05	0.04	0	1				
12	Subclasses	0.03	0.02	-0.03	-0.12	0	0.03	0.06	0.02	0.05	0.05	0.02	1			
13	Citation lag	-0.11	-0.03	0.12	-0.03	0.03	-0.03	0.11	0.07	0.05	-0.07	0.04	0.02	1		
14	Technological complexity	0.02	-0.02	0.07	0.08	0.01	-0.07	-0.04	-0.03	-0.04	0.02	-0.02	-0.39	-0.01	1	
15	Organizational coupling	-0.02	-0.01	-0.04	-0.01	0.04	-0.15	-0.21	-0.13	0.03	-0.02	0	0.01	0.01	0	1
	Mean	0.85	2.96	0.44	0.63	0.11	26.75	24.34	23.6	2.14	0.04	1.01	4.8	2.42	2.14	0.01
	S.D.	2.34	36.07	1.47	1.3	0.14	18.48	14.62	23.7	1.4	0.13	2.06	3.08	1.33	0.97	0.04
	Min	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0
	Max	47	2927	58.67	8.74	9.19	95	55	86.56	20	1	86	46	12	9.41	1

n = 44,959, correlation coefficients that are greater than |0.063| are significant at $p < 0.05$

TABLE 2

Fixed effects negative binomial regression of Knowledge Retrieval

	(1)	(2)	(3)	(4)
Firm-level controls				
Technological opportunity	-0.028**	-0.029**	-0.029**	-0.030**
	(0.009)	(0.009)	(0.009)	(0.009)
Slack	0.021**	0.021**	0.019**	0.020**
	(0.007)	(0.007)	(0.007)	(0.007)
Absorptive capacity	-0.079	-0.078	-0.075	-0.074
	(0.098)	(0.096)	(0.094)	(0.093)
Firm age	0.000	0.000	0.000	0.001
	(0.002)	(0.002)	(0.002)	(0.002)
Organizational units	-0.006*	-0.006**	-0.007**	-0.007**
	(0.002)	(0.002)	(0.002)	(0.002)
Firm size	0.004**	0.004**	0.004**	0.004**
	(0.001)	(0.001)	(0.001)	(0.001)
Patent-level controls				
Team size	0.036**	0.037**	0.036**	0.037**
	(0.004)	(0.004)	(0.004)	(0.004)
Internal focus	0.602**	0.593**	0.602**	0.593**
	(0.065)	(0.065)	(0.065)	(0.065)
Technological maturity	-0.004	-0.005	-0.004	-0.005
	(0.004)	(0.003)	(0.004)	(0.003)
Subclasses	-0.009**	-0.004	-0.009**	-0.004
	(0.002)	(0.002)	(0.002)	(0.002)
Citation lag	-0.088**	-0.088**	-0.088**	-0.087**
	(0.007)	(0.007)	(0.007)	(0.007)
Explanatory variables				
Technological complexity		0.160**		0.160**
		(0.033)		(0.033)
(Technological complexity)^2		-0.024**		-0.024**
		(0.007)		(0.007)
Organizational coupling			-2.567**	-2.520**
			(0.693)	(0.690)
(Organizational coupling)^2			2.316**	2.305**
			(0.892)	(0.888)
Constant	-1.615**	-1.840**	-1.566**	-1.792**
	(0.093)	(0.100)	(0.094)	(0.101)
Chi^2	2684.444	2742.210	2702.143	2759.692
Log-likelihood	-39197.07	-39174.25	-39189.69	-39167.09

Notes. Standard errors in parentheses. * p<0.05, ** p<0.01. Firms = 87, Observations = 38, 976 Dummy variables for the application year and the technological category of each patent is included in all models, but they are not reported in the table. Additionally, all models include *External Impact* as an exposure term.

TABLE 3

Fixed effects negative binomial regression of Knowledge Retrieval Frequency

	(1)	(2)	(3)	(4)
Firm-level controls				
Technological opportunity	-0.015	-0.015	-0.017	-0.018
	(0.009)	(0.009)	(0.009)	(0.009)
Slack	0.047**	0.047**	0.047**	0.046**
	(0.007)	(0.007)	(0.007)	(0.007)
Absorptive capacity	-0.248*	-0.241*	-0.220	-0.214
	(0.118)	(0.118)	(0.115)	(0.114)
Firm age	-0.003	-0.002	-0.002	-0.002
	(0.002)	(0.002)	(0.002)	(0.002)
Organizational units	0.006**	0.006**	0.005*	0.004*
	(0.002)	(0.002)	(0.002)	(0.002)
Firm size	0.006**	0.006**	0.006**	0.006**
	(0.001)	(0.001)	(0.001)	(0.001)
Patent-level controls				
Team size	0.046**	0.046**	0.046**	0.047**
	(0.005)	(0.005)	(0.005)	(0.005)
Internal focus	0.696**	0.688**	0.692**	0.684**
	(0.068)	(0.068)	(0.068)	(0.068)
Technological maturity	-0.006	-0.007	-0.006	-0.007
	(0.004)	(0.004)	(0.004)	(0.004)
Subclasses	-0.005*	-0.003	-0.005*	-0.002
	(0.003)	(0.003)	(0.003)	(0.003)
Citation lag	-0.035**	-0.034**	-0.034**	-0.033**
	(0.008)	(0.008)	(0.008)	(0.008)
External impact	-0.022**	-0.022**	-0.022**	-0.022**
	(0.001)	(0.001)	(0.001)	(0.001)
Explanatory variables				
Technological complexity		0.142**		0.141**
		(0.037)		(0.037)
(Technological complexity)^2		-0.025**		-0.025**
		(0.008)		(0.008)
Organizational coupling			-3.256**	-3.230**
			(0.688)	(0.688)
(Organizational coupling)^2			3.081**	3.064**
			(0.923)	(0.922)
Constant	-4.836**	-5.013**	-4.786**	-4.962**
	(0.092)	(0.101)	(0.092)	(0.101)
Chi^2	2443.181	2463.535	2467.378	2487.297
Log-likelihood	-52579.13	-52569.83	-52566.84	-52557.72

Notes. Standard errors in parentheses. * p<0.05, ** p<0.01. Firms = 87, Observations = 38, 976 Dummy variables for the application year and the technological category of each patent is included in all models, but they are not reported in the table. Additionally, all models include *Spillover Pool* as an exposure term.

FIGURE 1
Illustration of knowledge retrieval

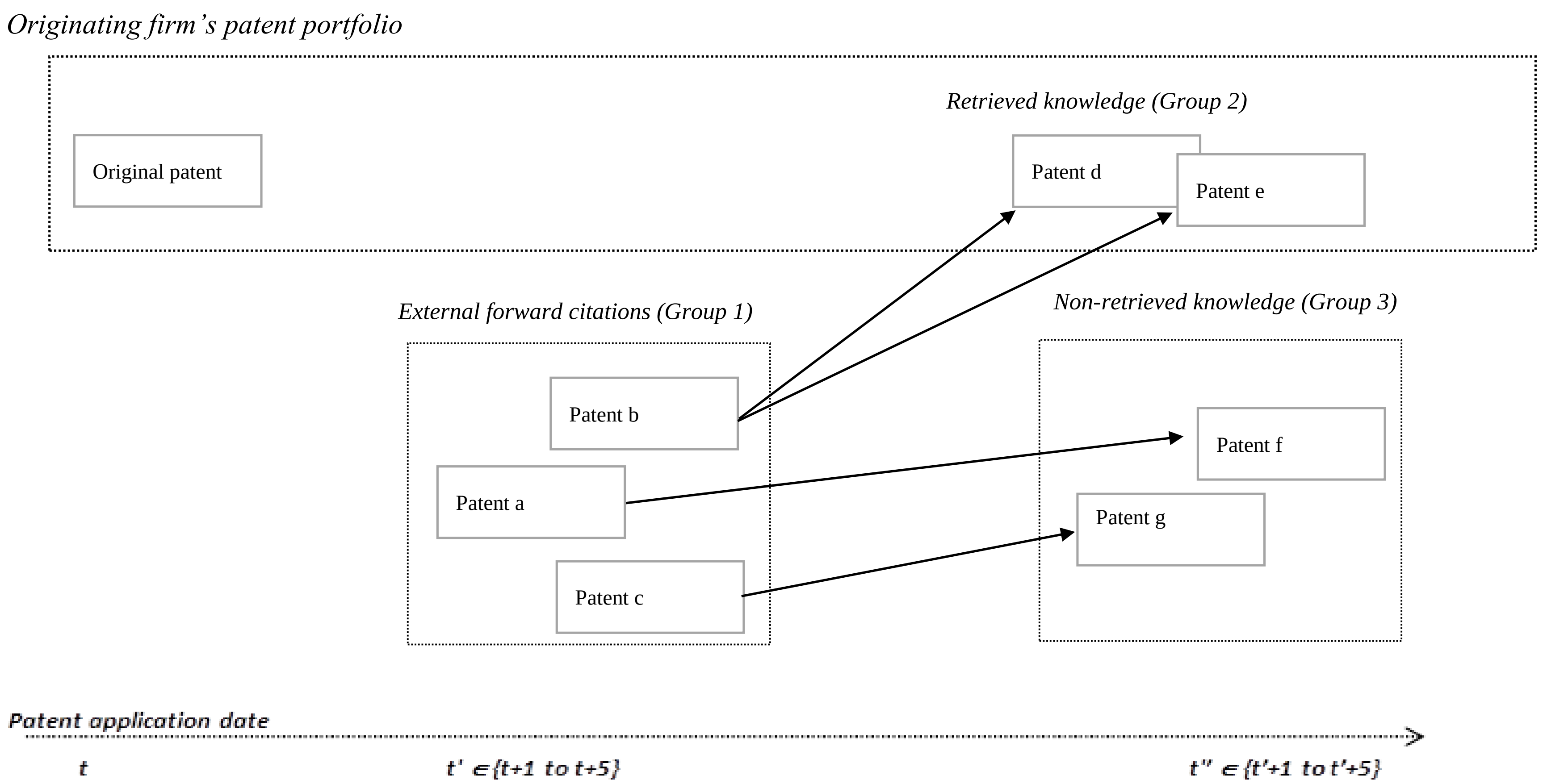

FIGURE 2
Technological complexity, organizational coupling and the extent of knowledge retrieval