Available online at www.sciencedirect.com

ScienceDirect

Research Policy 35 (2006) 994–1017

research policy

www.elsevier.com/locate/respol

# Complexity, networks and knowledge flow

Olav Sorenson[a,*], Jan W. Rivkin[b], Lee Fleming[c]

[a] London Business School, Sussex Place, Regent's Park, London NW1 4SA, United Kingdom
[b] Morgan Hall 239, Harvard Business School, Boston, MA 02163, USA
[c] Morgan Hall T95, Harvard Business School, Boston, MA 02163, USA

Received 13 April 2005; received in revised form 29 April 2006; accepted 3 May 2006
Available online 15 June 2006

**Abstract**

Because knowledge plays an important role in the creation of wealth, economic actors often wish to skew the flow of knowledge in their favor. We ask, when will an actor socially close to the source of some knowledge have the greatest advantage over distant actors in receiving and building on the knowledge? Marrying a social network perspective with a view of knowledge transfer as a search process, we argue that the value of social proximity to the knowledge source depends crucially on the nature of the knowledge at hand. Simple knowledge diffuses equally to close and distant actors because distant recipients with poor connections to the source of the knowledge can compensate for their limited access by means of unaided local search. Complex knowledge resists diffusion even within the social circles in which it originated. With knowledge of moderate complexity, however, high-fidelity transmission along social networks combined with local search allows socially proximate recipients to receive and extend knowledge generated elsewhere, while interdependencies stymie more distant recipients who rely heavily on unaided search. To test this hypothesis, we examine patent data and compare citation rates across proximate and distant actors on three dimensions: (1) the inventor collaboration network; (2) firm membership; and (3) geography. We find robust support for the proposition that socially proximate actors have the greatest advantage over distant actors for knowledge of moderate complexity. We discuss the implications of our findings for the distribution of intra-industry profits, the geographic agglomeration of industries, the design of social networks within firms, and the modularization of technologies.

*Keywords:* Diffusion; Information; Knowledge; Social networks; Competitive advantage

The flow of knowledge plays a central role in a wide variety of fields (for a review, see Rogers, 1995). Sociologists began investigating diffusion processes – and the importance of social structure to those processes – to understand the adoption patterns of agricultural and medical innovations (Ryan and Gross, 1943; Coleman et al., 1957). To students of technology management, knowledge flow first arises as an important issue in the context of technology transfers within the firm (Allen, 1977; Teece, 1977), but questions of diffusion also arise when technology scholars ask whether incumbent firms or upstarts first develop and commercialize new inventions (Reinganum, 1981; Tushman and Anderson, 1986). Both students of organizational learning (for a review, see Argote, 1999) and industrial economists (Griliches, 1957; Zimmerman, 1982; Irwin and Klenow, 1994) study how knowledge moves through firms and how it spills over to other firms. In short, a diverse array of scholars shares an interest in knowledge diffusion processes.

The normative interpretation given to diffusion, however, differs dramatically across fields. Economists and

* Corresponding author.
*E-mail addresses:* osorenson@london.edu (O. Sorenson), jrivkin@hbs.edu (J.W. Rivkin), lfleming@hbs.edu (L. Fleming).

0048-7333/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.respol.2006.05.002

sociologists tend to focus on the societal benefits of spillovers (i.e. the flow of knowledge across actors, usually firms). The generation of new knowledge often requires substantial investment in research and development, but the repeated application of this knowledge, once produced, entails little if any incremental cost (Arrow, 1962). Knowledge diffusion, therefore, engenders scale economies and stimulates economic development by allowing several firms to benefit from the R&D activities undertaken by a single firm (Marshall, 1890; Scherer, 1984; Romer, 1987). Management scholars, by contrast, note that when knowledge escapes to competing firms the returns to innovation become fleeting at best. As rivals imitate new products and processes, the degree of differentiation or cost advantage accruing to the innovator erodes. The business literature thus urges managers to defend against spillovers (Lippman and Rumelt, 1982; Kogut and Zander, 1992).

Though their prescriptions differ, economists, sociologists, strategists, and students of technology management all seek a better understanding of why some knowledge disperses widely while other knowledge does not. In this quest, some scholars have focused on the attributes of the knowledge itself. For example, highly specific knowledge may flow slowly because few parties other than the initial innovator either have the baseline knowledge and skills necessary to absorb it (Cohen and Levinthal, 1990) or can benefit from its application (Henderson and Cockburn, 1996; McEvily and Chakravarthy, 2002). Other studies focus on how social networks structure the flow of knowledge (e.g., Coleman et al., 1957; Hansen, 1999; Singh, 2005), implicitly attributing the rate of diffusion to the locus of innovation in the network.

This paper seeks to augment our understanding of knowledge flow by examining the interplay between two features: social proximity and the complexity of the underlying knowledge.[1] Social proximity here refers to the distance between two parties in a social network; for example, one would consider those who have a direct relationship to each other to be closer than those who have a mutual acquaintance but have never met. We meanwhile define complexity in terms of the level of interdependence inherent in the subcomponents of a piece of knowledge (Simon, 1962; Kauffman, 1993; cf. Zander and Kogut, 1995). Interdependence arises when a subcomponent significantly affects the contribution of one or more other subcomponents to the functionality of a piece of knowledge. When subcomponents are interdependent, a change in one may require the adjustment, inclusion or replacement of others for a piece of knowledge to remain effective.

Consider then an actor who is a source of knowledge and two potential recipients of that knowledge—one socially close to the source and one further away. When does the proximate actor have the greatest advantage over the distant in receiving and building on the knowledge? We argue that *the advantage should peak when the underlying knowledge is of moderate complexity*. Our expectation emerges from the recognition that receiving and building on knowledge frequently requires the recipient to engage in search to fill in gaps and correct transmission errors in the knowledge conveyed—the cost and difficulty of which increase with knowledge complexity. Social proximity reduces the need for search by facilitating high-fidelity transmission (i.e., complete information with negligible noise). On the other hand, as the social distance separating the source and the would-be receiver grows, unaided search plays an increasingly important role in diffusion. Under such conditions, simple knowledge should flow universally – to actors near and far – because search can easily substitute for high-fidelity transmission. Highly interdependent knowledge meanwhile defies diffusion, regardless of whether one relies on search or social proximity. For knowledge of moderate complexity, however, a gap emerges between the ability of close actors, relative to that of distant actors, to receive and build on knowledge. High-fidelity transmission gives proximate actors sufficient insight that they can succeed in receiving and building on knowledge, even where more distant actors, who rely more heavily on search, fail.

We analyze patent data to test our thesis empirically. Citation patterns across patents offer something of a fossil record for the flow of knowledge—providing a lasting reflection of ephemeral interactions. Using this record, we estimate the effect of knowledge complexity on the likelihood of future citations as a function of the social proximity of future inventors to the inventor of the original piece of knowledge, comparing those socially close to and far from the source. To assess social proximity, we calculate the geodesic length between patents' inventors in a collaboration network. We also supplement this metric with indicators of geographic proximity and employment within the same organization. To gauge complexity, we develop a measure that reflects

[1] Hansen (1999) also focuses on the interplay between social relations and knowledge flow. His research differs from ours in three respects: (1) it does not explore the issues related to recipient search as a mechanism for the interplay; (2) it focuses on the strength of the connection between inventors rather than social proximity in a network; and (3) it analyzes the effects of a *portfolio* of relations rather than the characteristics of a connection in a dyad.

the historical interdependence of a patent's subcomponents with other subcomponents. The findings provide strong support for our core hypothesis: the higher likelihood of citation among proximate inventors peaks for knowledge of an intermediate level of complexity (interdependence).

This work contributes to the literature in several ways. First, from the perspective of social networks, it identifies one condition under which social proximity should prove especially important to knowledge flow: for knowledge of intermediate complexity. Though social scientists have usefully demonstrated that networks matter for the diffusion of knowledge, relatively little research considers precisely when those networks should matter most (Strang and Soule, 1998; Baker and Faulker, 2004). By synthesizing the social network perspective with work on conceptions of knowledge receipt as search, we identify scope conditions on the relevance of social connections to the diffusion process. Second, with respect to evolutionary economics, our work highlights social connections as an important channel through which "insiders" gain superior access to knowledge. Extant work asserts that insiders – defined usually as those within the same firm as the source – have better access to an original success, which serves as a template in efforts to transfer and extend that knowledge (Nelson and Winter, 1982: 119; Rivkin, 2001). Yet this work fails to establish the source of this preferential access. Does it come from incentives that reward transfer, from the confidentiality agreements that employees sign, or from some other source? Our research points to direct social connections as a critical factor differentiating these internal parties from those outside the firm.

## 1. The flow of complex knowledge

Our discussion begins with the most common finding of classic diffusion studies: the S-shaped cumulative adoption curve (Ryan and Gross, 1943; Griliches, 1957; Rogers, 1995, provides an excellent review). Researchers consistently find that the adoption of an innovation over time follows a common pattern: growing slowly at first, then accelerating rapidly, and finally slowing to reach some asymptotic saturation level. These dynamics resemble that of an epidemic spreading through a population; the innovation first 'infects' those most at risk of exposure – actors closest to the original source (Hägerstrand, 1953) – and those most susceptible to infection – those most prepared to accept the uncertainty associated with an untested technology (Mansfield, 1968) or whose idiosyncratic characteristics make the innovation appear most attractive (Griliches, 1957). Over time, awareness of the innovation spreads, uncertainty ebbs, and the economics of the invention become favorable to a larger share of the population. Diffusion then takes off. In this classic perspective, new knowledge resembles a stone thrown into a calm pond, its ripples moving steadily across the entire surface.

Though this pattern accurately describes the diffusion of a wide variety of innovations and knowledge, critics have faulted this focus on the S-curve for several reasons (cf. Mahajan et al., 1990; Hargadon, 1998). Two of these critiques have particular relevance here. First, the classic diffusion literature typically depicts knowledge as moving unaltered as it passes from one actor to the next. Contrary to this depiction, in reality transmission rarely occurs with perfect fidelity. Both gaps in the information sent and errors in its interpretation typically require the receiver to reconstruct portions of the original knowledge. This process occurs so commonly that it even forms the basis of amusement in the children's game of telephone.[2] Most knowledge, therefore, requires effort to acquire and transmutes to some extent as actors strive to receive and build upon it; recipients assimilating new knowledge must actively process it by experimenting with its application to new problem domains and environmental contexts. Witness, for instance, the efforts of American automakers as they struggled to digest the knowledge embodied in Japanese lean production techniques (Womack et al., 1990) or the labors of computer makers as they sought to imitate Dell's direct distribution model (Porter and Rivkin, 1999). In both cases, the receipt of knowledge required years of trial, error, reflection, and adjustment and, arguably, remains incomplete.

Even within the supportive infrastructure of an organization, receiving and building on new knowledge can prove difficult. Teece (1977), for example, reports that the transmission and assimilation of technical know-how accounted for 19% of project costs, on average – running as high as 59% in one case – in 26 international technology transfer projects. Chew et al. (1990) find the internal transfer of best practices so incomplete in multiplant commercial food operations that, within a firm, the best plants produce twice as efficiently as the worst, even after controlling for differences in processing technology, location, and plant size (Szulanski, 1996, offers additional evidence). Hence, we regard the act of receiving and building on knowledge not as the acceptance of a

[2] In this game, one child whispers a message into the ear of another, who then whispers what she heard into the ear of a third child and so forth. At the end, the final person announces the message he heard and the first person reveals the message that she originally whispered; the two usually differ dramatically.

complete, well-packaged gift, but rather as the beginning of a trial-and-error process.

Our second concern regarding the simple S-curve characterization of diffusion arises from its inattention to the crucial role that social networks play in diffusion. Several studies, largely out of sociology, demonstrate that knowledge spreads from its source not in concentric circles, but along conduits defined by social connections (Lazarsfeld et al., 1944; Coleman et al., 1966; Burt, 1987; see Marsden and Friedkin, 1993, for a review). Consider some of the relevant findings: Hedström (1994) discovered that network density and geographic proximity can explain most of the spread of the idea of unionization in Sweden. In an analysis of adoption patterns for "poison pills" and "golden parachutes," Davis and Greve (1997) offered strong evidence that information about these policies travelled through corporate board interlocks. And Hansen (1999) found that strong ties best conveyed complex knowledge across product development teams within a firm. A growing literature thus points to the importance of social networks as pathways that channel the flow of knowledge among actors.

We synthesize these two perspectives – knowledge receipt as an active process of experimentation and search, and an appreciation for the role of social networks – into a model of knowledge flow. The model offers unique predictions regarding how knowledge complexity influences patterns of success among efforts to receive and extend knowledge.

### *1.1. Knowledge receipt as search*

Building on the intellectual scaffolding of evolutionary economics, our perspective conceptualizes a piece of knowledge as a recipe (Nelson and Winter, 1982).[3] The list of potential ingredients encompasses both physical components and processes. The recipe details how to combine these ingredients – in which proportions, in what order, under what circumstances – to achieve a desired end. For instance, a recipe for a McDonald's outlet might read something like: "When a customer places a special order, the counter clerk keys the order into the register, which causes the order to show up on the computer screen in the kitchen, which induces the cook to put a raw hamburger on the grill..." or "when opening a new outlet, a manager in the real estate department secures a site while the franchising office identifies a franchisee. Next, the franchisee contacts construction contractors while hiring shift managers...." Though these recipes may appear in writing, they more commonly reside in the form of behavioral routines, individual memory, or technology (March and Simon, 1958).

The conceptualization of knowledge as a recipe leads naturally to thinking of innovation as a process of searching for new recipes. Following a long tradition (Schumpeter, 1939; Gilfillan, 1935; Usher, 1954), Nelson and Winter (1982) explicitly treat innovation as a search process; inventors explore the space of possible combinations of ingredients, or recipes, for new and better alternatives. This exploration involves not just the search for the best combinations of ingredients but also the quest for the most effective methods of integrating them. Researchers who conceptualize innovation as search frequently exploit a landscape metaphor as a means of providing an intuitive understanding of the search process (Levinthal, 1997; Rivkin, 2000; Fleming and Sorenson, 2001). Innovators – depicted as myopic in their awareness of the terrain – search these landscapes for peaks, which represent good recipes or useful inventions.

Once a useful innovation has been located, transferring its recipe, even between cooperative actors, can fail for two reasons. First, the recipient rarely grasps the original recipe completely, due to imperfections in the transfer process. Gaps emerge in what the sender conveys – perhaps the chef forgets an ingredient or skips a step – and the receiver may misinterpret some of the information that is transmitted. And, unless the recipient understands *perfectly* the recipe that generated the success – an unlikely situation – she must engage in search to fill the gaps and correct the errors in her version of the recipe. Any attempt to receive and extend a recipe in new settings will likewise require the recipient to rediscover the original combination, or some variant of it better suited to the new context.

Second, the local ingredients and cooking experience of the receiving chef rarely match identically those of the sender. Research on absorptive capacity (Cohen and Levinthal, 1990) emphasizes that successful knowledge diffusion requires the receiver to possess a base of knowledge and skills to assimilate new information. Without this baseline, the transmission of new discoveries would often entail the communication of exorbitant amounts of data; imagine how long a recipe would become if one needed to detail every step of the process—how to chop vegetables, how to boil water, etc. These two factors imply that knowledge recipients rarely, if ever,

[3] This assumption limits the applicability of our theory to innovations that involve multiple components. This restriction should not severely constrain its scope, however; few innovations do not involve the combination of multiple physical components or processes. For example, even the synthesis of nylon, a polymer, involved the integration of several distinct processes (Smith and Hounshell, 1985).

act merely as passive beneficiaries; they actively search, recreate, and build upon the original recipes.

In this process, certain types of recipes prove particularly tricky to transfer because the sender finds it difficult to specify and communicate precisely where the original combination resides in the combinatorial space of ingredients; on the figurative treasure map, it is hard to place the "X" that "marks the spot." This communication difficulty could arise as a result of causal ambiguity (Lippman and Rumelt, 1982; Reed and DeFillippi, 1990): the innovator might not fully understand the connection between actions and outcomes so the roots of the original success remain unclear. It could also occur because the production process calls on tacit personal skills or connections among individuals that the involved parties themselves do not consciously understand (Polanyi, 1966; von Hippel, 1988), or that eludes codification (Zander and Kogut, 1995). These factors essentially increase the likelihood that the knowledge transmitted has gaps. The complexity of the recipe itself can also impair knowledge flow by increasing the difficulty for the recipient of filling these gaps and correcting transmission errors.

As noted above, complexity refers to the degree to which the components in a recipe interact sensitively in producing the desired outcome. Our definition here closely follows Simon (1962), who classifies a piece of knowledge as complex if it comprises many elements that interact richly (see also Kauffman, 1993; cf. Zander and Kogut, 1995). We adopt Simon's definition, but pay particular attention to the intensity of interdependence among the ingredients in the recipe. A high degree of interdependence indicates that many ingredients influence the effectiveness of others so that a change in one may dramatically reduce the usefulness of the recipe. Replicating the functionality of the original recipe often requires adjustments in the set of other ingredients or the processes for combining them. Low interdependence implies small cross-component effects and a corresponding opportunity to adapt and change ingredients independently.

Discovering, or rediscovering, a complex piece of knowledge poses a stiff challenge. Interdependence produces two effects that undermine the recipient's attempts to receive and build on the original. First, small errors in reproduction cause large problems when ingredients cross-couple in a rich manner. In highly interdependent systems, implementers often realize no value from adopting a set of practices unless each-and-every component fits into place perfectly; a single error threatens the effectiveness of the entire system. An American automaker that attempts to adopt lean production techniques, for instance, may alter its human resource practices and inventory policies, yet see no benefit because it failed to invest appropriately in flexible production equipment. The fragility of such tightly coupled systems has been well documented (Weick, 1976; Perrow, 1984). Second, interdependence leads to a proliferation of "local peaks." These internally consistent – though not necessarily optimal – ways of combining ingredients elude improvement through incremental search because altering any single element degrades the quality of the outcome (Kauffman, 1993). Such local peaks would pose no problem to omniscient actors, who could assess the entire space of possibilities, but for individuals with finite cognitive abilities and a limited purview of the landscape, such search proves difficult; in the face of high interdependence, searchers frequently find themselves trapped on local peaks. Moreover, these local peaks tend to correspond to poor recipes precisely when interdependence creates a thick web of potentially conflicting constraints.

### 1.2. Complexity and access to a template

Success in receiving and building on complex knowledge depends crucially on access to the original recipe, which serves as a *template* (Nelson and Winter, 1982: 119–120; Winter, 1995). For reasons explored below, individuals differ in their access to the template. Superior access facilitates the knowledge recipient's search in at least two ways. First, the recipient begins searching in closer proximity to the ultimate target—as a result of either fewer errors in the interpretation of the transmission or smaller gaps in the information sent. Second, superior access allows the recipient to solicit advice when problems arise, helping the recipient to home in on the desired knowledge more efficiently.

Consider two actors both trying to receive and build on a valuable piece of knowledge but who differ in their access to the template. The first has superior, though admittedly still imperfect, access to and understanding of the original, successful recipe. The second has far poorer access. To what degree does the first actor's superior but imperfect access to the template have value, in the sense that it enables the actor to receive and build upon the original recipe more effectively? We contend that the value of this access depends on the complexity of the underlying knowledge in an inverted U-shaped relationship; that is, intermediate levels of interdependence maximize the value of preferential access.

Suppose first that the ingredients of the knowledge do not interact; getting one element in the recipe wrong diminishes that component's contribution to the whole, but it does not undermine the other components. In this situation, the first actor's access to the template does not educe a persistent advantage. Through routine, incremental search efforts, the second actor can reconstruct the recipe. Few local peaks threaten to trap the poorly informed recipient. As a result, both actors eventually fare equally well; search on the part of a recipient can easily substitute for high-fidelity transmission.

Next consider knowledge with an intermediate degree of interdependence. Local peaks now appear, but they remain relatively few in number. The well-informed actor begins its search near, but not precisely at, the original combination of ingredients. Through incremental search, and with recourse to the template, it can assemble the proper combination of ingredients. The second actor, who likely begins search farther from the target and receives less guidance about the direction in which to explore, more likely becomes ensnared on some local peak, away from and inferior to the original success. Here superior access to the template gives the first actor an advantage that the second cannot recreate through search.

Finally, imagine a piece of maximally interdependent knowledge: ingredients depend on one another in an extremely delicate way, and none produces much benefit unless all align perfectly. Local peaks now pervade the landscape and neither actor's incremental search will likely reproduce or build upon the original knowledge with any success. The first actor's superior access to the template thus has little value beyond the second's highly imperfect access.

Taken together, these arguments imply that the advantage of superior but imperfect access to the template reaches its peak at moderate levels of interdependence between knowledge components. With moderate interdependence, the smoothness of the landscape allows a party that begins its search near the desired peak to rediscover it through local search. Yet the landscape also has sufficient ruggedness that an actor that begins search far from the target likely finds itself trapped on a lower peak. In contrast, the single-peaked landscape that comes with independent components allows both parties to succeed in receiving and building on the source knowledge through local search. The highly rugged landscape produced by extreme interdependence meanwhile stymies both parties thoroughly. (For a more formal treatment, see the simulation in the Appendix A.)

### *1.3. Social networks and template access*

The quality of an actor's access to the template may depend on many factors. One crucial factor is the nature of the actor's social relations, which provide conduits through which valuable information travels (Homans, 1950; Hägerstrand, 1953). In particular, we claim that the quality of an actor's access to a template declines with social distance—that is, the number of nodes that separate the actor from the source of the knowledge in a social network. Direct, single-step connections provide the most obvious and valuable links between inventors and those attempting to receive and build on knowledge because they permit two-way communication. The recipient can therefore interactively query the original source of the knowledge to correct errors or to fill gaps in the original transmission.[4]

Short, indirect paths – for example with one or two intervening steps – can also provide beneficial access to the template, as even second-hand information provides important clues about how to reconstruct and build on new knowledge. Mutual acquaintances may also allow for direct communication with the source if they will introduce and vouch for a potential knowledge recipient (Burt, 1992). Moreover, actors removed by only a few steps from the knowledge source will share more background knowledge, a larger proportion of specialized language, and a wider range of beliefs with the source (for a review, see McPherson et al., 2001). All of these facilitate high-fidelity transmission (Durkheim, 1912; Arrow, 1974; Cohen and Levinthal, 1990). The quality of template access, however, undoubtedly declines rapidly as the number of actors between the innovator and the would-be receiver increases; as in the game of telephone, each step in the path between the two parties offers an opportunity for errors and omissions to creep into the transmission.

The previous subsection argued that superior access to the template creates the greatest advantage in knowledge diffusion with knowledge of intermediate complexity (interdependence). Combining that idea with the notion that social proximity provides superior access to the template, we arrive at the central proposition of our paper:

**Hypothesis.** In attempts to receive and build on knowledge, actors who are socially close to the source of the knowledge have the greatest advantage over distant actors when the knowledge is of intermediate interdependence.

In sum, we view knowledge diffusion as a search to receive and build on an effective recipe. Recipients

[4] Though not considered here, one might also consider the importance of tie "strength." Weak ties have long reach but low bandwidth; thus, they operate most prominently in the diffusion process when transferring only short, simple messages (Hansen, 1999).

socially proximate to the source of the knowledge have superior, though still imperfect, access to the original recipe. This advantage in access translates into higher fidelity reproduction that benefits the actor most significantly when the ingredients of the recipe display moderate interdependence. Simple recipes spread through the social network thoroughly, placing recipients both near and far on equal footing. Highly intricate recipes resist diffusion to even nearby actors. But for recipes of intermediate interdependence, nearby actors receive enough guidance from the template that local search delivers them an effective replica of the original knowledge on which they can build, while distant actors begin their search processes from such flawed starting points that subsequent efforts to receive and build on the interdependent recipe tend to fail.

## 2. Empirical corroboration

To test our hypothesis, we analyzed prior art citations to all U.S. utility patents granted in May and June of 1990 ($n = 17{,}264$).[5] The data came from the Micro Patent database and NBER public access data on patents (Hall et al., 2001). Following much previous research, we view a prior art citation as evidence of knowledge diffusion: the applicant has successfully assimilated the knowledge underlying the original patent to a new setting and built upon it. Our statistical approach is to estimate the likelihood that a focal patent receives a citation from a future patent as a function of several factors: the interdependence of the knowledge underlying the focal patent, the proximity of the inventors of the focal and citing patent in a social network, the interaction of interdependence and social proximity, and a set of control variables. The results of the estimation allow us to examine how the likelihood of citation by a socially proximate inventor compares to the likelihood of citation by a distant inventor as a function of knowledge interdependence. The crucial test of our hypothesis is whether the gap between the two probabilities peaks when the focal patent embodies moderately interdependent knowledge.

### 2.1. Patents and the meaning of citations

Patents and their citation patterns provide an attractive test bed for our hypothesis for several reasons. First, these citations have been carefully assigned. The U.S. Patent Office requires all applicants to demonstrate awareness of their invention's precedents by citing similar "prior art" patents. Patent examiners in each technological domain review and supplement the prior art references to ensure accurate and comprehensive citations. Second, consistent with our ontology of knowledge, technology historians have demonstrated that one can conceptualize patented inventions as combinations of pre-existing technological components (Basalla, 1988). The process of invention therefore involves both the replication of prior discoveries and the extension of those discoveries to new applications and in new combinations. When a citation to prior art emerges on a new patent, it suggests that the inventor has both successfully received and built upon the knowledge underlying the earlier patent. Third, Fleming and Sorenson (2001) have developed a technique for measuring the interdependence among the components of an invention. The technique draws on information uniquely available for patents and potentially difficult to duplicate in other settings.

This setting nevertheless also has its limitations. First, our analysis rests on the assumption that some potential knowledge recipients have better access to the template than others. If every patent fully revealed the inventor's underlying knowledge of the invention, this assumption would not hold. Inventor's incentives, however, minimize the likelihood of this problem. Patent applicants prefer to disclose as little as possible to limit their competitors' ability to benefit from their disclosure (Lim, 2001). Indeed, conversations with the U.S. Patent Office indicate that applicants often intentionally obfuscate their descriptions to diminish the value of the knowledge revealed (Stern, 2001).

Second, the use of citations as an indicator of knowledge flows has been cast into doubt recently by the work of Alcacer and Gittelman (in press), who find that examiners add 40% of the citations found on U.S. patents. On the one hand, this finding is comforting as it suggests that examiners actively work to prevent applicants from excluding citations to relevant prior art for strategic reasons, such as those mentioned above. It is nonetheless potentially problematic for our study to the extent that examiners most frequently insert socially proximate citations to patents of intermediate interdependence. The few studies that analyze the characteristics of examiner-added citations, however, show no evidence of such a bias (Alcacer and Gittelman, in press; Sampat, 2004). Indeed, *self*-citations – which almost certainly reflect true knowledge flows – as frequently come from examiners as from inventors. This suggests to us that, on balance, examiner intervention *improves* the quality of patent data for our purposes and cannot account for our results.

[5] We constructed this dataset in the course of prior research. For details on its construction, see Fleming and Sorenson (2001).

Consistent with this conclusion, Duguet and MacGarvie (2005) find that firms' patent citation patterns match their survey responses regarding technology acquisition and dispersion. At worst, if examiners add citations that do not reflect true knowledge flows and do so in an unbiased way, this should only add noise, increasing the difficulty of finding statistical support for our hypothesis.

Third, patents admittedly offer imperfect measures of invention. Inventors may limit their patent applications to a subset of their discoveries, and one must ask whether this selection process biases our results. Inventors most likely seek legal protection when a patent raises a meaningful barrier to imitation (e.g., when inventing around the patent proves difficult), when the invention will not quickly become obsolete, and when few alternative "natural" defenses protect the knowledge (Levin et al., 1987). Of these conditions, the last seems most germane to our study. It implies that our sample may under-represent inventions that involve highly tacit, causally ambiguous and complex knowledge. Empirical research, however, suggests that this selection bias may not exist: Cohen et al. (2000), for example, find that firms in industries with complex products disproportionately choose to patent.

Finally, we recognize that patents represent but one embodiment of knowledge. Though we have no reasons to expect *a priori* that they should differ from other pieces of knowledge, they may. Despite this potential limitation on the scope of the applicability of our results, patents offer an excellent first test bed for our ideas for the reasons noted above.

### 2.2. *Case-control design*

Our unit of analysis is a patent dyad, one patent issued in May or June of 1990 and one issued later that may or may not cite the first. Hence our approach conceptually follows that of other studies of the likelihood of tie formation—in this case, the likelihood that a future patent builds on the knowledge embodied in one of our focal patents. These studies have typically estimated tie formation on the entire matrix of possible relations (e.g., Podolny, 1994; Gulati, 1995). This approach has two disadvantages. With large numbers of nodes, in this case patents, it can generate enormous, sparse matrices, increasing the difficulty of estimation and variable construction. In our situation, this method would generate nearly 20 billion dyads with only around 60,000 realized citations. In addition, this approach raises questions regarding network autocorrelation and the non-independence of repeated observations on the same patents across multiple observations in the error structure.

Instead, our analysis follows Sorenson and Stuart (2001) in adopting a case-control approach to analyzing the formation of ties (see Sorenson and Fleming, 2004, for an earlier application to patents). The case-control sampling procedure works as follows. We begin by including all cases of future patents, from July 1990 to June 1996, that cite any of our 17,268 focal patents: 60,999 in total. Since these citations occur, the dependent variable $\text{Cite}_{ij}$ takes a value of "1" for these cases to denote a realized citation. In addition, we pair each focal patent with four future patents that do not cite it (but that could have).[6] We set $\text{Cite}_{ij}$ to zero for these control cases. Though this generates a data set of 130,055 dyads, our analysis restricts the sample used for estimation to the 72,801 cases where both inventors reside in the U.S.[7] To address the fact that focal patents enter the data more than once, we report robust standard errors estimated without the assumption of independence across repeated observations of the same focal patent.

The use of a matched sample introduces one new problem. Logistic regression can yield biased estimates when the proportion of positive outcomes in the sample does not match the proportion of citations in the population (Prentice and Pyke, 1979; Scott and Wild, 1997). In particular, uncorrected logistic regression using a matched sample tends to produce underestimates of the factors that predict a positive outcome (King and Zeng, 2001). Large samples do not necessarily alleviate this problem.

We adjust the coefficient estimates using the method proposed by King and Zeng (2001) for the logistic regression of rare events (cf. Manski and Lerman, 1977). The traditional logistic regression model considers the dichotomous outcome variable a Bernoulli probability

[6] We chose four patents for the "control" group so that the sample would have a roughly equal proportion of realized and unrealized dyads. Although some feel that conditioning on important factors improves the statistical power of a case-control sample (e.g., Jaffe et al., 1993, implicitly make such an argument in drawing controls from the same classes as the citing patents), the ideal method of selecting controls remains an open debate. Matching controls to cases on one or more dimensions can lead to two problems in particular that concern us. First, correcting the logit for over-sampling on the dependent variable requires that one knows the sampling probabilities (King and Zeng, 2001); matching controls to cases precludes the possibility of calculating this information. Second, matching on an endogenously determined factor risks generating biased results (e.g., when investigating diffusion processes, one would not want to consider the geographic distribution of activity exogenous). Given these concerns, we sample future patents at random and control for heterogeneity in the estimation.

[7] Including the foreign inventors does not change the results qualitatively.

function that takes a value 1 with the probability $\pi$:

$$\pi_i = \frac{1}{1+\mathrm{e}^{-X_i\beta}},$$

where $X$ represents a vector of covariates and $\beta$ denotes a vector of parameters. Researchers typically use maximum likelihood methods to estimate $\beta$. King and Zeng (2001) prove that the following weighted least squares expression estimates the bias in $\beta$ generated by oversampling rare events:

$$\text{bias}(\hat{\beta}) = (X'WX)^{-1}X'W\xi,$$

where $\xi = 0.5Q_{ii}[(1+w_1)\hat{\pi}_i - w_1]$, the $Q$ are the diagonal elements of $Q = X(X'WX)^{-1}X'$, $W = \text{diag}\{\hat{\pi}_i(1-\hat{\pi}_i)w_i\}$, and $w_1$ represents the fraction of ones (citations) in the sample relative to the fraction in the population. At an intuitive level, one regresses the independent variables on the residuals using $W$ as the weighting factor. Tomz (1999) implements this method in the relogit Stata procedure.

This case-control approach offers two principal advantages over the count models employed in most patent research. First, this method permits far more fine-grained controls for heterogeneity in citing patents. Count models preclude the possibility of controlling for detailed features of a citing patent. The ability to account for the attributes of the potential citing patents proves critical, however, to testing our hypotheses, which suggest that the ability of future inventors to receive and build on the original knowledge varies as a function of their social proximity. Second, analyzing citations at the level of the citing-patent/cited-patent dyad avoids the potential for aggregation bias inherent in count models.

### 2.3. Interdependence

Following Fleming and Sorenson (2001), we measure the complexity of the knowledge in a patent by observing the historical difficulty of recombining the elements that constitute it. Though it involves intensive calculation, the intuition behind the metric is straightforward: a technology whose components have, in the past, been mixed and matched readily with a wide variety of other components has exhibited few sensitive interdependencies. The measure considers the subclasses identified in a patent as proxies for the underlying components. Though in many cases subclasses correspond to identifiable physical components (such as in the example below), they do not always align so closely. Our measure, however, requires only that these subclasses define pieces of knowledge rather than physical components. Combining some pieces that interact sensitively to each other proves more difficult than connecting relatively independent chunks of knowledge.

We calculate the measure of interdependence, **k**, in two stages.[8] Eq. (1) details our measurement of the ease of recombination – the inverse of interdependence – for subclass $i$ used in patent $j$. We first identified every use of the subclass $i$ in previous patents from 1980 to 1990.[9] The sum of the number of prior uses provided the denominator. For the numerator, we counted the number of different subclasses appearing with subclass $i$ on previous patents. Hence, our measure increases as a particular subclass combines with a wider variety of technologies, controlling for the total number of applications, and captures the ease of combining a particular technology. To create our measure of interdependence for an entire patent, we averaged the inverted ease of recombination scores for the subclasses to which it belongs (Eq. (2)):

$$\text{Ease of recombination of subclass } i \equiv E_i = \frac{\text{Count of subclasses previously combined with subclass } i}{\text{Count of previous patents in subclass } i} \qquad (1)$$

$$\text{Interdependence of patent } j \equiv \mathbf{k}_j = \frac{\text{Count of subclasses on patent } j}{\sum_{i\in j} E_i}. \qquad (2)$$

Intuitively, the measure operates as follows. Suppose a patent embodies subclasses that have been combined with a wide variety of subclasses, even in a handful of previous patents. This indicates that the patent's components do not have delicate interdependencies that prevent widespread recombination and the components can

[8] Our measure **k** is related to but distinct from the parameter $K$ in the NK simulation models that have become popular in theoretical work on complex systems (Kauffman, 1993). In NK simulations, the contribution of each element in a system to overall system fitness depends on the states of $K$ other elements. $K$ is set by the modeler and, like our empirically measured **k**, reflects the degree of interdependence among components in a system. Despite the conceptual linkage between our measure **k** and Kauffman's $K$, we do not purport to have measured his $K$ in a literal sense. For instance, our **k** does not equal the number of elements that affect the contribution of each focal element.

[9] Some might worry about the stability of this measure over time. To test its robustness, we constructed a second **k** measure using data from 1790 to 1990. That measure yielded a qualitatively identical set of results.

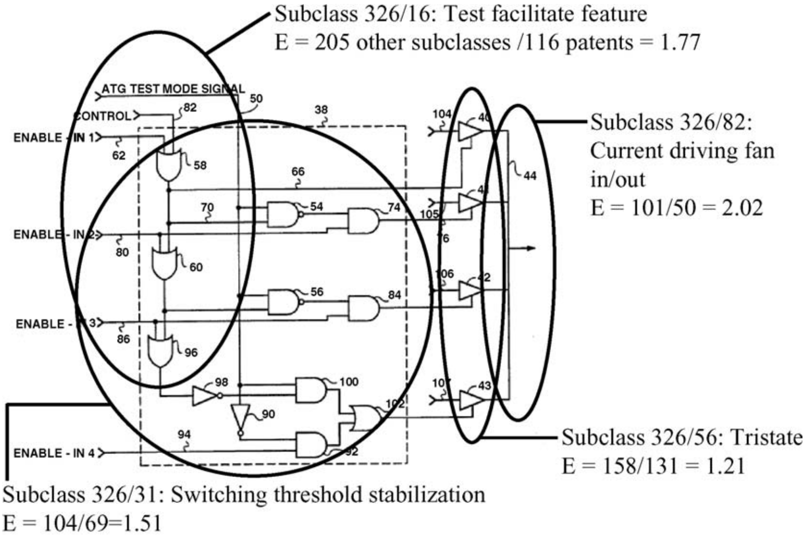
be mixed and matched independently. Such a patent receives a low value of **k**. Suppose instead that a patent embodies subclasses that have been combined, again and again, with the same small set of other subclasses. We presume those subclasses to be highly interdependent; their repeated joint appearance in patents suggests that the presence of one requires the appearance of the others. Hence the patent's **k** is high.

In addition to the measure's face validity, it has been validated externally via a survey of inventors. Fleming and Sorenson (2004) asked a sample of patent holders the following question, based on Ulrich's (1995) definition of interdependence: "Modules are said to be coupled when a change made to one module requires a change to the other module(s) in order for the overall invention to work correctly. How coupled were the modules of your invention?" They then compared survey responses to calculated **k** for the corresponding patents and found a significant correlation between inventors' perceptions of coupling and the calculated degree of interdependence.

Concrete examples may clarify the metric further and help to link it to our core hypothesis. Consider a digital technology patent, #5,136,185, filed by the third author of this paper. Fig. 1 outlines the calculation of **k** for this patent and the mapping of the USPTO classification scheme to the components used. 326/16 identifies the "Test facilitate feature" subclass, which implements a testing mode within a semiconductor chip. Prior to its appearance here, this subclass had been recombined 116 times with 205 other components, implying an observed ease of recombination score of 205/116 = 1.77. 326/56 indicates the "Tristate" subclass, and 326/82 points to "Current driving fan in/out" subclass. 326/31 meanwhile identifies the "Switching threshold stabilization" subclass (essentially a priority encoder). Fig. 1 illustrates the location of these components on the circuit, the calculation of their ease of recombination scores, and the calculation of the patent's interdependence, **k** (=0.61)—a level slightly above the mean **k** for our sample.

The invention described above assists engineers in testing the logic gates on new chips—a difficult task when chips can contain hundreds of thousands or even millions of such gates. Even though the patent appears to disclose much of the important information, it does not reveal the proprietary test generation algorithm, and how that algorithm manipulated the components (in particular, the "test facilitate feature"). Without access to, or an understanding of, that algorithm, rivals could see the components of the knowledge in the patent but not how the components worked together. As a result, competitors faced an uphill battle in exploiting the knowledge. Even within the firm, effective transmission required the inventor to travel around the country to teach others how to use the technology. Similarly, competitors found it difficult to reproduce IBM's copper interconnect technology – another invention of intermediate complexity – until enough engineers defected to rivals to diffuse the relevant knowledge of how to fabricate the copper interconnect without contaminating the wafer's other materials (Lim, 2001).

By comparison, inventions involving extremely high levels of interdependence defy diffusion even within


Subclass 326/16: Test facilitate feature
E = 205 other subclasses /116 patents = 1.77
ATG TEST MODE SIGNAL
CONTROL
ENABLE - IN 1
ENABLE - IN 2
ENABLE - IN 3
ENABLE - IN 4
Subclass 326/82: Current driving fan in/out
E = 101/50 = 2.02
Subclass 326/56: Tristate
E = 158/131 = 1.21
Subclass 326/31: Switching threshold stabilization
E = 104/69=1.51

Interdependence **k** = 4 subclasses /(1.77+2.02+1.21+1.51) = 0.61

Fig. 1. Calculation of interdependence for patent #5,136,185.

a social boundary. Plasmid preparation, for example, a biological technique, involves an intricately intertwined sequence of actions involving various chemicals, reagents and manual operations. As Jordan and Lynch (1992: 84) note, "Although the plasma prep is far from controversial and is commonly referenced as a well established and indispensable technique, how exactly it is done is not effectively communicated, either by print, word of mouth, or demonstration." On the other hand, inventions involving a low degree of interdependence diffuse rapidly. For instance, patent #4,927,016, one of the patents in the bottom quartile of the **k** range, involves the production of monoclonal antibodies. The industry associated with this technology has essentially become a commodity business since one can easily acquire all the necessary knowledge components by reading a textbook and piece them together without concern for sensitive interdependencies. Polymerase chain reaction, a technique for amplifying DNA sequences, has followed a similar route. Or, one might think of Sun's workstation technology. The modular design of its system has allowed rivals to match the performance of its hardware quickly, limiting the company's ability to maintain an advantage in the hardware market.

### 2.4. *Social proximity*

The analyses investigate the effect of knowledge complexity on the diffusion of knowledge to individuals whose close social connections to the source of knowledge give them better access to the template than individuals with distant or no connections have. For each of our 72,801 patent dyads, we develop one direct and two indirect indicators of social proximity between the inventors of the two patents in the dyad.

#### 2.4.1. *Proximity in a collaboration network*

Our most direct indicator measures the distance between inventors in a network of patent collaborators. The idea underlying this indicator is that an inventor gains access to a template via collaborators, collaborators of collaborators, collaborators of collaborators' collaborators, and so forth. Closer connections grant better access. To measure collaborative proximity, we use the methods and data of Singh (2005).[10] Consider the dyad consisting of patent $i$ issued in May or June of 1990 and patent $j$ issued at a later time $t$ (before 1996). To compute the distance between $i$ and $j$, Singh first constructs a network with a node for each discrete inventor who has been listed on any patent from 1975 until time $t$. An edge connects two inventors if they have collaborated on a patent during that period. The collaborative distance of a patent dyad is then the minimum number of intermediaries required to connect a member of the team of inventors listed on patent $i$ to a member of patent $j$'s team. If the two teams share a member, for instance, the distance is zero. If the teams have no common members but an individual listed on neither patent has collaborated with members of both $i$'s and $j$'s teams, the distance is one, and so forth. If no path connects members of the two teams, the distance is $\infty$. See Singh (2005) for a complete description of his approach.

Based on the distance measure, we construct three indicator variables for each dyad[11]:

- Close Collaboration$_{ij} = 1$ if the distance between patents $i$ and $j$ is less than 4; 0 otherwise.
- Far Collaboration$_{ij} = 1$ if the distance between $i$ and $j$ is 4 or greater but less than $\infty$; 0 otherwise.
- Unconnected$_{ij} = 1$ if no path connects $i$ and $j$.

The shorter the path between $i$ and $j$, the better the access to the template enjoyed by the team involved in patent $j$. Our core hypothesis is that this superior access translates into a higher probability of citation especially when the components of patent $i$ display intermediate interdependence. Accordingly, we expect the gap in citation probability between a close and a far inventor – the probability that a close inventor cites a focal patent minus the probability that a far inventor cites the patent – to peak at an intermediate level of **k**.

Although our collaborative distance measure provides direct evidence of access and we believe that it captures many of the important connections between inventors, inventors also have many other types of relations that might also facilitate access. For example, a potential recipient might be a friend of the source even if they have never collaborated. Attempting to identify all of the potential relationships existing in any population of individuals is not feasible, but we can examine two

[10] Breschi and Lissoni (2002) independently developed an equivalent approach.

[11] Though the magnitude of the gap shrinks, our results remain qualitatively robust to shifting the dividing line between close and far from a path length of three to a length of four. We use three categories rather than the distance measure itself for three reasons: (1) calculating the precise distance for the longer paths in these data would increase the time required to compute it by orders of magnitude (i.e. by months); (2) dummy variables for individual path lengths lead to some small cell sizes and concomitantly unstable coefficient estimates; and (3) given our interaction with a quadratic, we find the results of the categorical coding far easier to interpret and understand.

factors – geographic proximity and joint organizational membership – that tend to structure social relationships and therefore may proxy for unobserved social paths between our source–recipient dyads. As McPherson et al. (1992: 154) note: "Homophily structures the flow of information and other social resources through the network so that the dimensions themselves stand as proxies for the number of intervening steps in transmissions through the system."

### 2.4.2. *Geographic proximity*

Space represents one important dimension that structures social interaction. Indeed, some of the earliest literature on social networks emphasized the dramatic decline in the likelihood of a social relation as two parties became increasingly distant (Park, 1926; Bossard, 1932). Accordingly, we develop a measure of geographic proximity for each patent dyad:

- Geographic proximity$_{ij}$ = the natural log of the distance in miles between the first inventors listed on patents *i* and *j* multiplied by negative one (so that larger values indicate greater proximity).[12]

As with our direct measure of social proximity, we expect geographic proximity to have the greatest impact on citation likelihood when the potentially cited patent displays moderate interdependence.

### 2.4.3. *Organizational proximity*

Social networks also concentrate within foci, such as organizations (Feld, 1981). On a daily basis, most fully employed individuals spend more waking hours engaged in work than in any other activity. Employees regularly meet other employees through work to cooperate on projects, to confer on decisions, to transfer information, and to socialize. Hence, we use employment at the same patent assignee as another indicator of social proximity:

- Organizational proximity$_{ij}$ = 1 if the same organization owns both patents in a dyad, 0 otherwise.

We expect common ownership to boost citation likelihood, especially for focal patents of moderate interdependence.

We test our hypothesis by regressing Cite$_{ij}$ on the indicators of social proximity directly, the indicators interacted with $\mathbf{k}$, and the indicators interacted with $\mathbf{k}^2$. We expect social proximity to boost citation probability directly. The core test of our hypothesis resides not in the direct effects but in the interaction terms: the impact of proximity on citation probability should have an inverted-U relationship with respect to interdependence $\mathbf{k}$.[13]

In light of our empirical context, patent citations, it is useful to elaborate our expectations about the direct effect of $\mathbf{k}$ on citation likelihood. Our hypothesis describes the impact of interdependence on the *gap* between near and distant actors' success in receiving and building on knowledge. We examine this gap by examining *interactions* of $\mathbf{k}$ and $\mathbf{k}^2$ with social distance. In developing the hypothesis, however, we also paint a picture of the *direct* impact of $\mathbf{k}$ on knowledge reproduction: we suggest that greater interdependence increases the difficulty for a party of receiving and building upon prior knowledge, regardless of the party's distance from the source. This argument concerns an actor's success in receiving and building on knowledge *conditional on an attempt to do so being undertaken*. Patent citation data nevertheless reflect not only success conditional on an attempt being undertaken, but also the sheer number of attempts being undertaken. We have reason to believe that the number of attempts may rise with interdependence, simply because interdependence increases the fertility that comes from mixing and matching components (Fleming and Sorenson, 2001). Accordingly, we offer no hypothesis about the direct effects of $\mathbf{k}$ on citation rates. Instead, we focus on the *gap* between near and distant actors' citation rates, which should have a robust inverted-U relation to interdependence. (See the Appendix A for a more detailed treatment of this point.)

## 2.5. *Controls*

The non-monotonic interactions between interdependence and proximity that we predict – if found in the data – lend themselves to few alternative interpretations. The models nevertheless include as controls several of the most important variables used in prior patent studies (e.g., Lanjouw and Schankerman, 2004).

[12] All patents list the home address of the inventor on the front page of the patent application. To locate each inventor, we match the inventor's 3-digit zip code to the latitude and longitude of the center of the area in which the inventor resides based on information from the U.S. Postal Service. We then use spherical geometry to calculate the distance between the points. The USPTO includes 5-digit zip information, but we choose to reduce measurement error by using cleaned data. CHI, an information provider, has called every patent holder to verify the inventor's location; however, it records this information only at the 3-digit level.

[13] We mean-deviate the variables before creating the interaction terms to facilitate interpretation of the effects (Friedrich, 1982). For collaborative proximity, we use Unconnected$_{ij}$ as the excluded category.

Table 1
Descriptive statistics and correlations

| | Mean | S.D. | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. **k** | 0.49 | 0.30 | 0.03 | −0.02 | −0.07 | 0.00 | −0.11 | 0.10 | 0.07 | −0.03 | −0.05 | −0.29 | −0.35 |
| 2. Close collaboration | 0.07 | 0.25 | | −0.21 | 0.14 | 0.30 | 0.17 | −0.00 | 0.08 | 0.01 | 0.01 | −0.01 | 0.00 |
| 3. Far collaboration | 0.23 | 0.42 | | | −0.06 | 0.01 | 0.01 | 0.12 | 0.15 | −0.01 | 0.03 | 0.02 | 0.05 |
| 4. Organizational proximity | 0.10 | 0.33 | | | | −0.09 | −0.03 | −0.06 | −0.07 | 0.02 | −0.04 | −0.02 | −0.03 |
| 5. Geographic proximity | −6.50 | 1.96 | | | | | 0.20 | 0.00 | 0.05 | 0.05 | 0.01 | 0.00 | 0.01 |
| 6. Same class | 0.26 | 0.44 | | | | | | 0.12 | 0.08 | 0.04 | 0.02 | −0.11 | −0.01 |
| 7. Activity control | 1.25 | 0.42 | | | | | | | 0.42 | 0.01 | 0.06 | −0.02 | 0.08 |
| 8. Recent technology | 3.97 | 0.62 | | | | | | | | −0.14 | 0.09 | 0.05 | 0.09 |
| 9. Backward patent citations | 9.83 | 8.88 | | | | | | | | | 0.13 | 0.07 | 0.12 |
| 10. Backward non-patent citations | 1.46 | 4.24 | | | | | | | | | | 0.06 | 0.10 |
| 11. Number of classes | 1.85 | 0.97 | | | | | | | | | | | 0.49 |
| 12. Number of subclasses | 4.53 | 3.43 | | | | | | | | | | | |

### 2.5.1. *Activity control*

The activity control accounts for the typical number of citations received by a patent in the same technological areas as the focal patent. In a first stage, we calculated the average number of citations that each patent in a particular USPTO class received from patents granted between January of 1985 and June of 1990 (Eq. (4)).[14] We then weighted these parameters according to the patent's class assignments (Eq. (5)), where $p_{ik}$ indicates the proportion of patent $k$'s sub-class memberships that fall in class $i$:

$$\text{Average citations in patent class } i \equiv \mu_i = \frac{\sum_{j\in i} \text{Citations}_j \text{ (before } 7/90)}{\text{Count of patents } j \text{ in subclass } i} \quad (4)$$

$$\text{Technology mean control patent } k \equiv M_k = p_{ik}\mu_i \quad (5)$$

The models also include controls for several other factors. *Same class* is a dummy variable denoting whether the two patents in each dyad belong to the same primary technological class. *Recent technology* is the mean of the patent numbers of the focal patent's prior art (higher numbers indicating more recent technology).[15] The models include counts of two types of backward patent citations. First, they include a tally of the number of citations to patent *prior art*. Second, the models include a control for the number of *non-patent prior art* citations (e.g., references to published articles). *Number of classes* is a count of the number of major classes and *number of subclasses* is a count of the number of subclasses to which the focal patent is assigned. Descriptive statistics appear in Table 1.[16]

## 3. Results

The results appear in Table 2. Model 1 estimates the effects of the control variables alone, and Model 2 introduces interdependence, **k**.

Model 3 provides the first test of our core hypothesis by interacting interdependence with collaboration-based indicators of social proximity. The results provide three pieces of support for the hypothesis. First, the positive sign on **k** × Close collaboration coupled with the negative sign on $\mathbf{k}^2$ × Close collaboration indicates that the gap in citation probability between close and unconnected inventors rises and then falls, peaking when the source knowledge displays moderate interdependence. (Recall that Unconnected is the excluded category, so the coefficients related to Close collaboration capture differences between close and unconnected inventors.) Second, by subtracting the coefficients for Far collaboration from the coefficients for Close collaboration, we see that the largest gap between close and far inventors also appears for moderate **k**. Third, the coefficient estimates suggest that the greatest gap between far and

[14] We allow all patents issued between January 1985 and June 30, 1990 to enter the estimation of the activity control, meaning that the patents used to calculate it vary in the time during which they can receive citations. Alternatively, we could select a small set of patents from 1985 and base the measures on the subsequent 5 years of citations; however, this approach would ignore the patent activity just prior to our sample.

[15] This variable made use of the fact that the USPTO assigns patent numbers sequentially. This assignment pattern generates a correlation between a patent number and the grant date of the patent of 0.98.

[16] We also considered as a control variable the time between the issuance dates of the focal and potentially citing patents in each dyad. Exploratory analysis revealed small effect sizes (though typically significant), and inclusion of the time control had no meaningful impact on the coefficients of central interest.

Table 2
Rare events logit models of the likelihood of a focal patent receiving a citation from a future patent[a]

| | Model 1 | Model 2 | Model 3 | Model 4 |
|---|---|---|---|---|
| **k** | | 0.863 (0.257)*** | −1.599 (0.644)* | −1.305 (0.378)** |
| $\mathbf{k}^2$ | | −0.203 (0.086)*** | 1.116 (0.209)*** | 1.051 (0.156)*** |
| **k** × Close collaboration | | | 3.242 (1.670)* | 4.327 (1.659)** |
| $\mathbf{k}^2$ × Close collaboration | | | −3.428 (0.708)*** | (0.725) −4.881*** |
| **k** × Far collaboration | | | 1.569 (0.573)** | 1.899 (0.627)** |
| $\mathbf{k}^2$ × Far collaboration | | | −.802 (0.162)*** | −1.056 (0.262)*** |
| **k** × Geographic proximity | | | | 0.325 (0.078)*** |
| $\mathbf{k}^2$ × Geographic proximity | | | | −0.241 (0.031)*** |
| **k** × Organizational proximity | | | | 0.547 (0.679) |
| $\mathbf{k}^2$ × Organizational proximity | | | | −0.508 (0.232)* |
| Close collaboration | 3.952 (0.628)** | 3.979 (0.618)*** | 3.660 (1.148)*** | 2.925 (1.135)** |
| Far collaboration | 0.224 (0.090)* | 0.249 (0.089)** | 0.244 (0.089)** | −0.359 (0.246) |
| Geographical proximity | 0.041 (0.012)*** | 0.041 (0.012)*** | 0.045 (0.011)*** | 0.053 (0.012)*** |
| Organizational proximity | 0.457 (0.118)*** | 0.431 (0.119)*** | 0.423 (0.116)*** | (0.355) 0.292 |
| Same class | 4.800 (0.084)*** | 4.820 (0.085)*** | 4.797 (0.083)*** | 4.784 (0.083)*** |
| Activity control | 0.503 (0.097)*** | 0.515 (0.098)*** | 0.469 (0.095)*** | 0.481 (0.096)*** |
| Recent technology | 0.268 (0.147) | 0.278 (0.144) | 0.226 (0.161) | 0.245 (0.147) |
| Backward patent citations | 0.022 (0.005)*** | 0.021 (0.005)*** | 0.020 (0.005)*** | 0.021 (0.005)*** |
| Backward non-patent citations | 0.011 (0.009) | 0.014 (0.009) | 0.010 (0.008) | 0.010 (0.009) |
| Number of classes | 0.184 (0.047)*** | 0.209 (0.048)*** | 0.204 (0.047)*** | 0.201 (0.046)*** |
| Number of subclasses | 0.184 (0.013)*** | −0.016 (0.014) | −0.017 (0.013) | −0.023 (0.013) |
| Constant | −12.28 (0.586)*** | −12.89 (0.657)*** | −12.21 (0.746)*** | −11.91 (0.697)*** |
| Log-likelihood | −33772.4 | −33751.1 | −33738.2 | −33720.2 |

[a] 72,801 dyads (52% realized ties vs. 0.0004% in population); * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

unconnected inventors arises for moderate **k** (though with much smaller magnitude; see below). In sum, our primary measure for social proximity provides strong support for our core hypothesis.[17]

Model 4 adds interactions of interdependence with geographic and organizational proximity. Both proxies for social proximity display the expected inverted-U relationship, though only the results for geographic proximity show strong statistical significance. Coefficients for the collaboration-based measures retain their signs and significance, as do most of the coefficients for the control variables.

Based on Model 4, Fig. 2 traces out as a function of interdependence, how many times more likely a citation

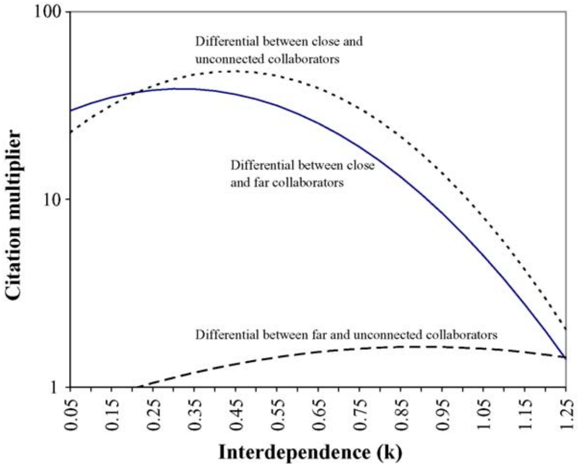


Fig. 2. Citation multiplier for proximate vs. distant actors in the collaboration network as a function of interdependence. *Note*: The line labeled "differential between close and unconnected collaborators" shows, as a function of **k**, how many times more likely a citation is in a dyad of patents whose inventors can reach one another through the collaboration network (path length < 4) relative to a dyad whose inventors are unconnected in the network. When **k** = 0.45, for instance, a citation is 48 times more likely. The other two lines provide the same information for pairs of actors who are close vs. far (path length between 4 and ∞) in the collaboration network and for pairs of actors who are far vs. unconnected. The figure is based on Model 4 of Table 2 for inventors from different organizations, with all variables other than **k** and the collaboration network indicators set to their mean values.

[17] Since the high correlation between a term and its square can force estimates to take opposing signs, we further tested the validity of our non-monotonic effect in two ways: (1) in unreported estimates (available from the first author), we re-estimated the models using a log-quadratic specification and found qualitatively identical results. Since this functional form can capture decreasing returns without a significant coefficient on the quadratic term, it is less sensitive to these problems. (2) We estimated a model with only the linear term and interactions and then entered the quadratic terms. In all cases, the addition of the quadratic terms significantly improved the model. (For example, in Model 4, the addition of the quadratic **k** and its interactions has a $\chi^2 = 70.4$, significant at $p < 0.00001$ with five degrees of freedom.)

is for collaboratively close pairs of inventors than for unconnected pairs, for close pairs than for far pairs, and for far pairs than for unconnected pairs. (We set all other variables to their mean values for the purpose of creating this chart.) The figure shows vividly that the maximal difference in citation probabilities between close pairs and unconnected pairs arises when the focal patent displays moderate interdependence. The same is true of the difference between close and far pairs. Fig. 2 also shows that the citation difference between far and unconnected inventors – while consistent with our hypothesis and statistically significant – is much, much smaller. This suggests that for access to knowledge, the value of a social connection to the source drops off rapidly with the number of intervening intermediaries, echoing the findings of Singh (2005).

Fig. 3 shows, as a function of interdependence, how many times greater the probability of citation is between geographically proximate actors than between geographically distant actors. It does likewise for pairs of inventors in the same organization versus pairs in different organizations. In both cases, the benefits of social proximity rise and then fall with **k**, peaking when the source knowledge displays moderate interdependence. This provides graphical affirmation of our hypothesis. In both Figs. 2 and 3, the peak differences fall within the range of actual **k** in our data—in fact, within one standard deviation above the mean.

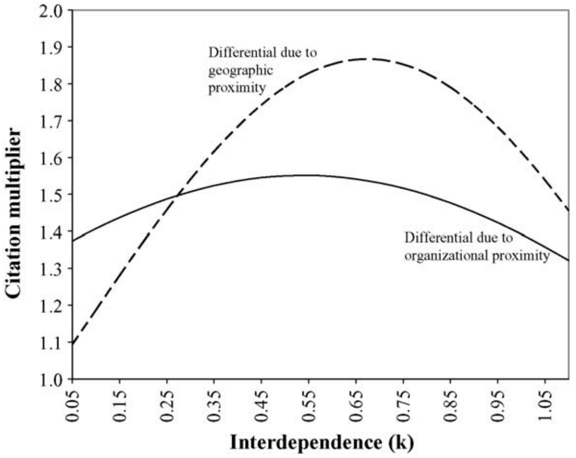


Fig. 3. Citation multiplier for proximate vs. distant actors (in geography and organizational space) as a function of interdependence. *Note*: The line labeled "differential due to geographic proximity" shows, as a function of **k**, how many times more likely a citation is in a dyad of patents when the inventors' addresses on the patents reside 10 miles apart than when they reside 3000 miles apart. When **k** = 0.65, for instance, the multiplier is 1.87 (i.e. 87% more likely). The line labeled "differential due to organizational proximity" shows, as a function of **k**, how many times more likely a citation is in a dyad of patents when the same organization owns both patents relative to when they are owned by different organizations. The figure is based on Model 4 of Table 2, with all variables other than **k** and geographic and organizational proximity set to their mean values.

In addition to being significant, the effects associated with our hypothesis can have substantial economic import. For source knowledge that is simple ($\mathbf{k} \sim 0$), an inventor close in the collaboration network is 30 times more likely than a far inventor to cite a focal patent. For knowledge of moderate interdependence at the gap-maximizing level of **k** shown in Fig. 2, this number rises to 39 times. As knowledge becomes more complex, the number falls, becoming a mere seven times at $\mathbf{k} = 1$. For close and unconnected inventors, the figures are 23 times, 48 times, and 11 times, respectively.[18] Similarly, contrast an inventor 10 miles from the source of knowledge and another 3000 miles away (both collaboratively-unconnected to the source and in different organizations). When $\mathbf{k} \sim 0$, the first inventor is 9% more likely that the second to cite the source. When **k** is at the gap-maximizing level, the probability rises to 87%. It then falls to 61% for $\mathbf{k} = 1$. Such differences in citation likelihood are far from negligible.

Despite the apparent consistency of our results with our expectations, proximity – collaborative, geographic, or organizational – may reflect factors other than the strength of social connections, factors that might also influence the quality of one's access to the template. Actors proximate to a given patent might, for instance, work on similar technical problems and therefore more readily absorb the knowledge embodied in the patent (Cohen and Levinthal, 1990). Any factor that improves access to the template should have the effect that we hypothesize. It is natural to interpret the proximity measures as indicators of social contact, as we do. It is difficult, however, to rule out all other factors that the proximity measures might reflect.

Similarly, our interdependence measure may capture not only the complexity of an item of knowledge but also its breadth of applicability. Our results might then reflect a process in which low-**k** knowledge is broadly applicable and diffuses widely; moderate-**k** knowledge is of particular interest to select groups who tend to be socially proximate to the inventor; and high-**k** knowledge is of such narrow application that it diffuses very narrowly. This would produce a pattern in which actors socially proximate to a source of knowledge most frequently receive and build on it if the knowledge has

[18] These figures assume that the two inventors are 665 miles from one another (the average distance in our sample) and work for different organizations.

moderate **k**. The driving force under this alternative interpretation is not the relative ability of different actors to search in the face of complexity but the relative interest that different actors have in obtaining knowledge. The alternative interpretation raises the question of precisely what makes an item of knowledge broadly or narrowly applicable. Knowledge becomes broadly applicable in part because it is modular and therefore can mix and match with other pieces of knowledge across a wide range of circumstances. Applicability, then, may capture the interdependence of a piece of knowledge (especially if one defines interdependence broadly and not in a narrow technological sense). To the extent that applicability reflects interdependence, we return to our original core hypothesis: individuals proximate to the source of some knowledge have the greatest advantage in receiving and building on knowledge of moderate interdependence/applicability.

## 4. Discussion

The analysis of patent citation patterns supports our basic theoretical perspective on knowledge diffusion: search in the space of possible combinations of ingredients offers a useful lens for understanding the flow of knowledge. Recipients socially proximate to the source of the knowledge have preferential access to the original success, which serves as a template during efforts to receive and build on the knowledge. All recipients, socially near and far, compete on equal footing when receiving and extending simple knowledge; incremental search suffices to reproduce simple knowledge, so guidance from a prior success has little value. Highly complex knowledge, on the other hand, equally resists diffusion to both classes of would-be recipients. Hence, at both extremes of complexity, the close recipient has no lasting advantage over the distant. In contrast, for knowledge whose ingredients display a moderate degree of interdependence, superior but imperfect access to the template translates into greater success in receiving and building on preexisting knowledge. The close recipient can complete its initially imperfect replica via local search, but local search alone cannot guide the distant recipient to an accurate replica. Thus in our patent data, the largest gap between the ability of a close recipient to receive and build on prior knowledge relative to the ability of a distant recipient arises when the cited patent involves moderate interdependence. This result appears when social distance is measured by proximity in a collaboration network as well as when geographic and – to a lesser extent – organizational proximity proxy for social distance.

Our findings have an array of practical and theoretical implications, especially for the issue of knowledge inequality across social borders. Consider the graph of a typical social network. It is quite common in such a graph to observe patches of actors with dense connections amongst themselves and areas of sparse connections between patches (Owen-Smith and Powell, 2004). The dense patches may reflect firms, for instance, or geographic regions. Actors within each patch sit socially proximate to one another but relatively distant from actors in other patches. A question of great practical importance is: When does knowledge diffuse within the patch where it originated but not across the thin areas into other patches? When will knowledge diffuse within a firm but not to competitors, or within a region but not to other locales? When is inequality of knowledge sharpest across social borders? Our results suggest that the nature of the knowledge, specifically its degree of complexity, plays a critical role. One might initially suspect that highly complex knowledge, the most difficult to reproduce, would create the greatest inequality across boundaries. Yet this intuition ignores the fact that inequality in its sharpest form requires *some* diffusion: to create the most inequity across social boundaries, knowledge must creep up to the edge of the thick patch of connections in which it originated but not beyond. This phenomenon, we have argued, most likely occurs for moderately complex knowledge.

Accordingly, the results suggest a resolution to the replication/imitation dilemma that has puzzled evolutionary economists and strategy scholars. To achieve a competitive advantage from knowledge, a firm must typically leverage that knowledge across multiple applications, for example, across all its production facilities (Winter, 1995). Yet any would-be replicator with a valuable piece of knowledge faces a dilemma: the profits produced by its original knowledge attract the envious attention of imitators. Valuable knowledge provides a source of *sustained* advantage only to the extent that it lends itself to replication yet defies imitation. Unfortunately for the innovator, replication and imitation typically go hand-in-hand (Nelson and Winter, 1982). Our results suggest, however, that replication-without-imitation is especially likely when the target knowledge entails moderate complexity. This micro-level phenomenon may manifest itself in outcomes at the industry level. One might expect that, *ceteris paribus*, industries based on moderately complex knowledge will display especially wide intra-industry dispersion in long-run financial returns. We leave this promising hypothesis for future research.

The results also speak interestingly to the literature on the geographic agglomeration of industries. Researchers frequently cite knowledge spillovers as a prominent reason that firms within an industry cluster together (Marshall, 1890; Krugman, 1991) and congregate near universities (Zucker et al., 1997). Our results certainly support this point of view: dense social networks, which tend to localize geographically, give firms and individuals close to the source of knowledge an important advantage in reproducing and building upon the knowledge. This begs the further question, why do some industries cluster while others do not? Though research on economic geography points out that knowledge spillovers can contribute to agglomeration, it does not identify *what type of knowledge* most likely engenders these clusters. Our findings suggest that industries that rely on moderately complex knowledge more commonly form industrial districts (cf. Sorenson, 2004). Simple knowledge can diffuse far and wide because incremental search efforts can substitute for high-fidelity communication. As the complexity of knowledge increases, a gap emerges between local diffusion and distant diffusion; thus, the potential return to locating near to innovators rises.

In addition to influencing geographic agglomeration and industry structure, the nature of the underlying knowledge used by a firm may have implications for organizational design. Firms have both formal and informal structures that influence the degree to which actors within the firm interact with each other. Managers can influence who likely interacts with whom through the assignment of individuals to facilities, the design of laboratories and factories, and the structure of reporting relationships (Allen, 1977). To distribute knowledge effectively, a firm might usefully expend resources to foment close and dense social connections between sources and intended recipients of complex knowledge, while letting networks remain sparse elsewhere. Indeed, leaders might fruitfully construe the task of knowledge management *not* as the construction of central databases of information (as sometimes presented today), but rather as an effort to build social networks that match the nature and intended flow of knowledge. Effective organizational design, however, surely requires a deeper understanding of how social structure affects knowledge diffusion than considered here; networks have subtle features and nuances that doubtlessly influence their ability to convey knowledge, both simple and complex (Hansen, 1999).

To this point, our argument has assumed that the degree of interdependence between combinations of components remains fixed. In the long term, however, the effective interdependence of knowledge may change. Firms and inventors can invest in R&D to specify interfaces and embed knowledge within physical components, thereby reducing the difficulty of combining a particular combination of components with other elements in the future (Baldwin and Clark, 2000). In structuring knowledge, managers must perform a delicate balancing act. Isolating interdependencies within substructures has important attractions, including the ability to perform a greater number of independent experiments (Baldwin and Clark, 2000) and the capacity to adjust more readily to environmental shifts (Levinthal, 1997). Engineering curricula support this preference with a strong emphasis on reliability, black box design techniques, and the re-use of previously combined components (e.g., Mead and Conway, 1980). Such modularization, however, also entails frequently overlooked costs. Designing and implementing an architecture that isolates interdependencies within substructures involves considerable engineering costs (O'Sullivann, 2001). But those direct costs potentially pale in comparison to the indirect costs—the opportunities that the lack of complexity opens for new entrants (Rivkin, 2000), the reduction in variety from which developers can select (Christensen et al., 2002), and the constraints on potential performance (Fleming and Sorenson, 2001). Managers who manipulate interdependencies should recognize that they simultaneously alter the propensity of knowledge to flow to actors near *and* far.

Despite the costs of modularizing, a secular trend towards modularization may influence the evolution of industries, creating a distinctive pattern. Direct costs likely strike firms as more tangible than indirect costs as they decide where to direct R&D effort. Thus, firms may over-invest in less complex technology as they seek to maximize efficiency. As this process reduces the effective interdependence of the knowledge being diffused, knowledge should flow more easily, generating two industry-level patterns. First, an industry that begins its life in a concentrated region should become less concentrated geographically as the advantage of preferential access to the template declines (for related ideas, see Audretsch and Feldman, 1996; Stuart and Sorenson, 2003).[19] Second, the move towards less complex knowl-

[19] This pattern seems consistent with the evolution of the software industry, for instance. Early on, knowledge localized to an extreme: understanding of a new piece of code resided in the head of a single developer or a small group of developers in a university, government, or large corporate computing facility. Inventors developed local languages for specific hardware. Over time, programmers developed techniques for reducing the interdependencies in code. Higher-level languages such as Cobol and C allowed programmers to divorce

edge likely reduces differentiation across firms' products over time, leading to more intense price competition and efforts to control standard interfaces and key modules—a pattern identified in the product lifecycle literature.

To reiterate, our results demonstrate that knowledge complexity importantly influences the dynamics of diffusion. Specifically, a socially proximate actor's advantage over a distant actor in obtaining and building on knowledge peaks when the components underlying the knowledge display intermediate interdependence. Though our empirical results come from patent data alone, the basic logic of our hypotheses applies to knowledge in general, not just the knowledge underlying inventions. Hence, future research might usefully examine these dynamics across a wide range of applications—including organizational learning, the diffusion of management practices, knowledge management, and the sustainability of knowledge-based competitive advantage.

## Acknowledgments

All parties contributed equally to this research. We thank Jasjit Singh for generously sharing his inventor collaboration network data. We also appreciate the helpful comments of George Baker, Matt Bothner, Koen Frenken, Bob Gibbons, Jerry Green, Rebecca Henderson, Bruce Kogut, Dan Levinthal, Woody Powell, Rosemarie Ham Ziedonis, three anonymous reviewers, and seminar participants at Harvard, the University of Michigan, MIT, New York University, Ohio State University, the University of Toronto, Washington University, and Wharton. All errors remain our own. Harvard Business School's Division of Research provided financial support. Earlier versions of this paper have been presented at the annual conferences of the Academy of Management, the European Group on Organizational Studies, and the European Meeting on Applied Evolutionary Economics, as well as at the ESF Exploratory Workshop on Evolutionary Economic Geography.

## Appendix A. Simulation of knowledge flow

A simple simulation of knowledge flow serves two purposes. It clarifies further why the value of social proximity reaches its peak in the transfer of knowledge with intermediate interdependence. It also identifies the range of empirical results consistent with our theoretical model. Specifically, the theoretical model yields a unique prediction about the impact of knowledge interdependence on the *gap* between citation rates of socially close actors and socially distant actors, but can encompass a range of findings about the effect of interdependence on close-actor citation rates alone or on distant-actor rates alone.

### *A.1. Model*

#### *A.1.1. Superstructure*

The model employs Kauffman's (1993) NK approach, which a growing number of researchers have used to simulate technological or organizational search. The simulation unfolds as follows. First, we choose two parameters: $N$, the number of components or ingredients that comprise a piece of knowledge, and $K$, the degree to which those components interact in determining the utility of the knowledge. Using techniques described below, a simulation then generates – in a stochastic manner – a mapping from each possible way of configuring the $N$ components (i.e. each conceivable recipe) to a measure of utility. One can visualize the mapping as a landscape in a high-dimensional space. Each discrete component constitutes a "horizontal" axis, and each possible manner of using the component represents a point along that axis. The vertical axis records the usefulness of the resulting piece of knowledge.

Next, we assume that some firm has happened upon the most useful possible piece of knowledge—the best way to configure the components (i.e. the template described in the main paper).[20] Two new parties then enter the landscape. One party, a close actor, has access to the owner of the template, presumably through a social tie, while the other, a distant actor, cannot access the original template through his social network. Both strive to rediscover the original success—the model's equivalent to the efforts to receive and build on knowledge discussed in the main text. Thanks to its superior access to the template, the close actor enjoys an advantage in this search process. The close actor may begin its search closer to the original success, reflecting the better information it receives or its superior ability to interpret the transmission. Or, it may move toward the success with greater speed and accuracy, reflecting its ability to seek advice from the owner of the template. The simulation mod-

code from specific hardware. Meanwhile, software production has dispersed geographically—beyond Silicon Valley, Route 128, and IBM's Armonk home, to Seattle, Austin, and even Bangalore.

[20] Our focus on the global maximum simplifies the simulation, but the results remain qualitatively robust to a wide range of alternative assumptions.

els these mechanisms and records the relative success of the close actor and the distant actor in rediscovering the original piece of knowledge.

Following this first iteration, the simulation generates a second mapping that, though it differs in its particulars, has the same degree of interdependence as the first. A second pair of close and distant actors tackle the second problem, and the program records their relative success. The simulation iterates through this process hundreds of times. From the repetition emerges a profile of how close and distant actors fare relative to one another for a given degree of interdependence. We then adjust *K*, the parameter that governs interdependence, and repeat the process. By doing so, we build an understanding of how interdependence affects the relative ability of close and distant actors to rediscover the original success.

This description of the model's superstructure leaves two aspects of the simulation unspecified: how we generate landscapes and how actors search to rediscover the original success.

### *A.1.2. Generation of landscapes*

Each piece of knowledge consists of *N* components, and each component $j$, $j \in \{1, 2, \ldots, N\}$, can be configured in two ways. Hence a particular piece of knowledge $\boldsymbol{s}$ is an *N*-vector $\{s_1, s_2, \ldots, s_N\}$ with $s_j \in \{0, 1\}$. In the knowledge germane to a chemical process, for instance, component *j* might indicate the inclusion or exclusion of a particular catalyst. Similarly, a string of four components could represent which of $2^4 = 16$ shades a heated mixture must turn before being removed from a flame. For any set of components, $2^N$ possible pieces of knowledge (recipes) exist. We assign a utility value to each of these as follows. Assume that each component contributes $C_j$ to utility. $C_j$, depending not only on the configuration, 0 or 1, of component *j*, but also on the configuration of *K* other randomly assigned components: $C_j = C_j(s_j, s_{j1}, s_{j2}, \ldots, s_{jK})$. For each possible realization of $(s_j, s_{j1}, s_{j2}, \ldots, s_{jK})$, we draw a contribution $C_j$ at random from a uniform distribution between 0 and 1. The overall utility associated with a piece of knowledge, then, averages across the *N* contributions:

$$U(\boldsymbol{s}) = \frac{[C_j(s_j, s_{j1}, s_{j2}, \ldots, s_{jK})]}{N}.$$

*K*, the parameter that governs interdependence, ranges from 0 to $N-1$.[21] $K=0$ corresponds to a simple situation in which the contribution of each component depends only on the configuration of that component. $K=N-1$ captures a complex setting in which the contribution of each component depends delicately on the configuration of every other component.

Once the modeler sets *N* and *K* and the simulation generates a particular landscape (i.e., a utility $U(\boldsymbol{s})$ for each of the $2^N$ possible pieces of knowledge), the simulation notes the piece of knowledge $\boldsymbol{s}^*$ that produces the greatest utility, which serves as a template in subsequent search efforts.

### *A.1.3. Search*

A modeled close actor and a modeled distant actor enter the landscape, and each struggles to rediscover the original success. Reflecting the reasoning on page 6 of the main text, neither begins precisely atop the peak at $\boldsymbol{s}^*$. Rather, each receives an imperfect transmission of the effective knowledge and begins some distance *d* from $\boldsymbol{s}^*$ (i.e. *d* of its *N* components differ from $\boldsymbol{s}^*$). It must then correct its understanding through search. We consider two types of search. A party involved in incremental search adjusts one component, accepts the adjustment if it produces an improvement in utility, and ceases to search when no improvement opportunities remain. A party engaged in long-jump search changes multiple decisions at once, leaping toward $\boldsymbol{s}^*$. Its leap typically misses the target; it replicates each component of $\boldsymbol{s}^*$ with probability $\theta$. $\theta < 1$ reflects imperfect access to the template. After its leap, the long jumper improves incrementally until it exhausts opportunities. Note that either type of search could terminate on a local peak, instead of at $\boldsymbol{s}^*$.

Though both parties have imperfect access, the close actor has better access due to her social proximity to the original success, which serves as a template. We model the impact of social proximity in three ways. The proximate actor may begin her search closer to $\boldsymbol{s}^*$ ($d_{\text{close}} < d_{\text{distant}}$), leap toward $\boldsymbol{s}^*$ with greater accuracy ($\theta_{\text{close}} > \theta_{\text{distant}}$), or – in leaping toward $\boldsymbol{s}^*$ – may know which components she has gotten "right" and "wrong." These benefits reflect both the more accurate transmission the close actor receives originally and her ability to consult with the owner of the template as she tries to correct the original transmission.

## *A.2. Interdependence and the landscape*

Much of the intuition of the results flows from an understanding of the impact of *K* on the topography of the typical landscape. Four effects strike us as espe-

[21] Note that the empirically derived measure of coupling in the main text, **k**, corresponds to the parameter *K* in the simulation model, but the two differ at least in terms of scaling. For more on this relationship, see footnote 9.

cially germane.[22] First, as $K$ increases, the landscape shifts from being smooth and single-peaked to being rugged and multi-peaked. When $K = 0$, the $N$ components contribute independently to knowledge utility. In that situation, alteration of a single component changes the contribution of that component alone. From any initial location on a landscape, then, a close or distant actor can climb to the global peak via a series of utility-improving, single-component tweaks to its knowledge. In contrast, when $K = N - 1$, every component influences the contribution of every other component. Then a small step on the landscape – a change in a single component – alters the contributions of all $N$ components. Consequently, adjacent pieces of knowledge have altogether uncorrelated utilities, producing a very rugged surface with many local peaks.

Second, as $K$ rises, not only do local peaks proliferate, but also the height of the average peak declines. As the web of connections across components thickens, it becomes possible to exhaust opportunities for incremental improvement even at low levels of performance. Hence, interdependence decreases the fruitfulness of incremental search.

Third, though the height of the average peak falls as $K$ rises, the heights of the highest peaks rise with $K$. When components interact with one another more richly, the amount of variety attainable by mixing and matching components increases, and the quality of the best combination within that variety improves. Rugged landscapes, though challenging to navigate, offer greater fertility than smooth ones—in other words, they more likely produce at least one exceptional peak. More mechanically, recall that we drew a contribution $C_j$ for each possible realization of ($s_j$, $s_{j1}$, $s_{j2}$, ..., $s_{jK}$). The number of possible contributions for each component ($2^{K+1}$) rises sharply with $K$, increasing the available variety.

Finally, as $K$ increases, the high peaks on the typical landscape spread apart from one another, shifting from a situation in which peaks cluster in mountain ranges to one in which peaks spread uniformly across the terrain.[23] With greater interdependence, high peaks carry less and less information about the location of other high peaks. This effect undermines long-jump search, decreasing the likelihood that a jump that aims for but misses the global peak will nonetheless land on high ground.

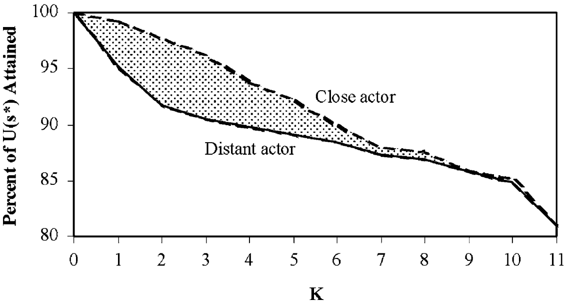


Fig. A1. Incremental search. *Note*: Parameter values for each simulation are given in text. Each data point is an average over 100 landscapes.

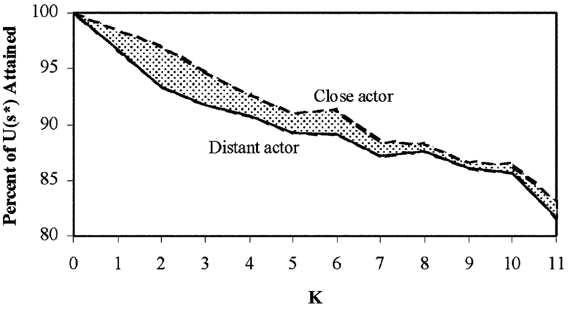


Fig. A2. Long-jump search. *Note*: Parameter values for each simulation are given in text. Each data point is an average over 100 landscapes.

### *A.3. Simulations and results*

#### *A.3.1. Percent of template performance attained*

We explored the model under a wide variety of assumptions regarding $d_{\text{close}}$, $d_{\text{distant}}$, $\theta_{\text{close}}$, and $\theta_{\text{distant}}$. ($N = 12$ throughout. All results average over 100–200 landscapes.) Results remained similar throughout the parameter space so we report only a handful of representative cases here (see Rivkin, 2001, for further robustness checks). Figs. A1–A3 show, as a function of $K$, the utility attained by the close actor and the distant actor as a percentage of the utility of the template. Fig. A1 considers the case of incremental search with $d_{\text{close}} = 4$

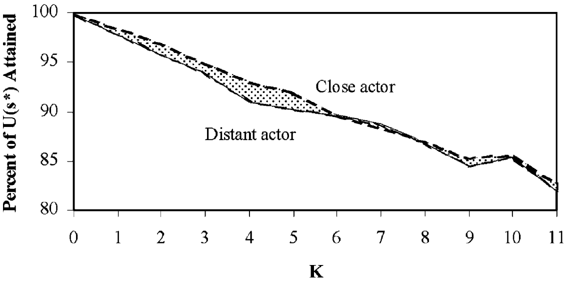


Fig. A3. Long jumps with vs. without knowledge of errors. *Note*: Parameter values for each simulation are given in text. Each data point is an average over 100 landscapes.

[22] Kauffman (1993) explores these effects further.

[23] For the intuition behind this effect, see Rivkin (2001), p. 283.

and $d_{\text{distant}} = 10$. Fig. A2 examines the case of long-jump search with $d_{\text{close}} = d_{\text{distant}} = 12$, $\theta_{\text{close}} = 0.6$, and $\theta_{\text{distant}} = 0.4$. Fig. A3 considers a situation in which both parties start with a poor replica ($d_{\text{close}} = d_{\text{distant}} = 12$), each tweaks uphill to a local peak, each then leaps toward $\boldsymbol{s}^*$ with equal accuracy ($\theta_{\text{close}} = \theta_{\text{distant}} = 0.5$), but in taking the leap, only the close actor knows which of its components matches the components of $\boldsymbol{s}^*$.

In all cases, greater interdependence undermines both close and distant actors, but the greatest gap between the two arises at an intermediate level of $K$. To see why, consider three situations:

- When $K = 0$, the close actor has no advantage at all. The smooth landscape allows both firms to discover the global peak eventually.
- As $K$ rises, a gap emerges between the close actor's performance and that of the distant actor. The landscape is rugged enough that the distant actor becomes stranded far from the global peak, and peaks cluster enough that average peak height declines with distance from the global peak. The landscape is sufficiently smooth and clustered, however, that the close actor – starting near $\boldsymbol{s}^*$ or leaping toward $\boldsymbol{s}^*$ accurately – can scale $\boldsymbol{s}^*$ or a nearby, nearly-as-high peak.
- As $K$ approaches $N$, the gap closes. The landscape becomes so rugged that even the close actor becomes stranded on a peak other than $\boldsymbol{s}^*$. The close actor may finish closer to $\boldsymbol{s}^*$ than the distant actor does, but with high peaks no longer clustered together, this proximity has little benefit. When components depend on each other delicately, superior but slightly imperfect access to the template has little more value than highly imperfect access.

### *A.3.2. Adjusting for frequency of attempts*

The results so far report the knowledge-rediscovery success of the close actor versus the distant actor conditional on both parties attempting to rediscover the knowledge embodied in the original success. In our empirical tests, however, we examine the rates of patent citations by close and distant actors. We interpret these rates as an indication of the number of times the knowledge underlying the focal patent has been received and built upon. Accordingly, the rates reflect not only the degree of success conditional on an attempt at rediscovery being made, but also the frequency with which attempts are made. If, for instance, the frequency of attempts varies systematically with $K$, then the graphs of close- and distant-actor patent counts versus $K$ might reveal shapes that differ in important ways from the pattern shown in Figs. A1–A3. In this light, we consider three scenarios.

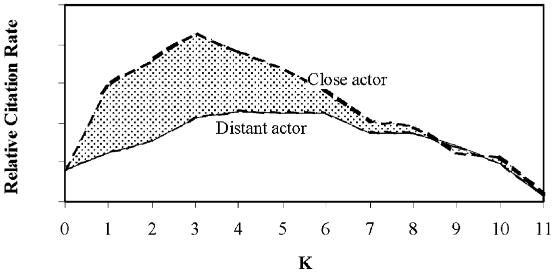


Fig. A4. Incremental search with number of attempts proportional to utility of template. *Note*: Parameter values for each simulation are given in text. Each data point is an average over 100 landscapes.

First and most simply, suppose that the number of attempts made by close and distant actors is independent of $K$. Then we would expect the graphs of citation rates to resemble Figs. A1–A3 without modification. In other words, the frequency of both close- and distant-actor citation would decline with $K$, and the maximal difference would occur at intermediate $K$.

Second, assume that the number of attempts made by socially close and distant actors increases in proportion to the utility associated with the original success (i.e., more useful pieces of knowledge attract more attempts at rediscovery). Recall that the utility of the best piece of knowledge – the height of the global peak on the landscape – rises with $K$, reflecting the greater variety that comes from mixing and matching more interdependent components. When we adjust Figs. A1–A3 to incorporate more frequent rediscovery efforts on high-$K$ landscapes, the citation pattern shifts to that shown in Figs. A4–A6. In contrast to Figs. A1–A3, the frequency of close- and distant-actor citation now rises at first, reflecting the fertility of higher-$K$ landscapes, but then declines. In line with Figs. A1–A3, the largest gap between close- and distant-actor citation arises for knowledge of intermediate interdependence.

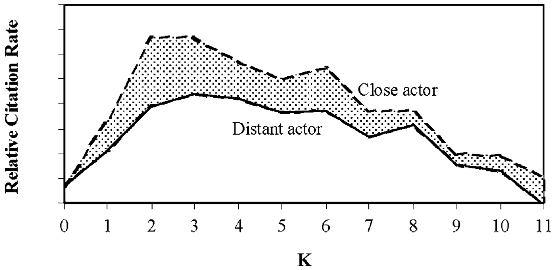


Fig. A5. Long-jump search with number of attempts proportional to utility of template. *Note*: Parameter values for each simulation are given in text. Each data point is an average over 100 landscapes.

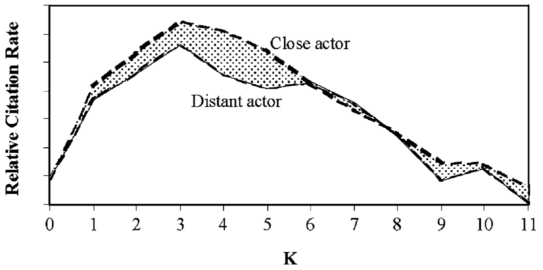


Fig. A6. Long jumps with vs. without knowledge, number of attempts proportional to utility of template. *Note*: Parameter values for each simulation are given in text. Each data point is an average over 100 landscapes.

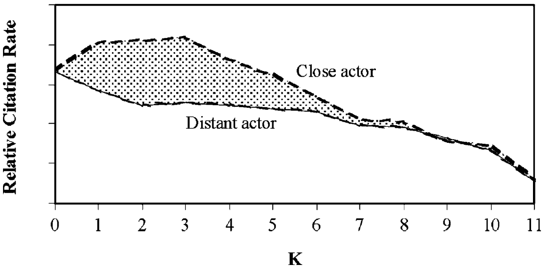


Fig. A7. Incremental search with number of attempts proportional to expected utility of attempt. *Note*: Parameter values for each simulation are given in text. Each data point is an average over 100 landscapes.

Finally, suppose that the number of attempts made by close and distant actors reflects the utility that each expects to attain in a rediscovery attempt. In deciding whether to engage in an attempt, parties not only understand that potential utility increases with *K*, but they also adjust for the odds that they succeed. For instance, distant actors understand they have lower odds of success and therefore make fewer attempts than do close actors. When we adjust Figs. A1–A3 in this manner, we project the citation pattern shown in Figs. A7–A9. Now the distant-actor citation rate declines monotonically with *K* while the close-actor citation rate has an inverted-U shape. Still, the gap between the two reaches its peak at an intermediate value of *K*.

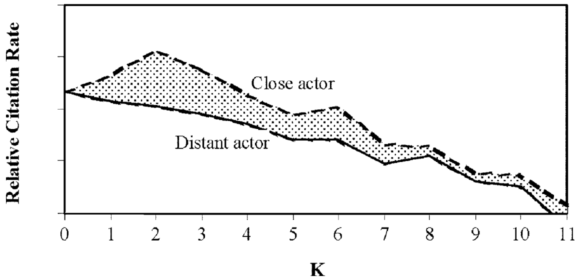


Fig. A8. Long-jump search with number of attempts proportional to expected utility of attempt. *Note*: Parameter values for each simulation are given in text. Each data point is an average over 100 landscapes.

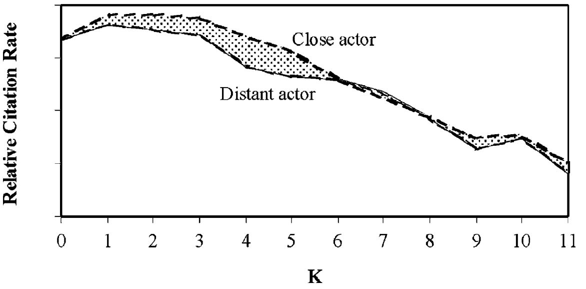


Fig. A9. Long jumps with vs. without knowledge, number of attempts proportional to expected utility. *Note*: Parameter values for each simulation are given in text. Each data point is an average over 100 landscapes.

In sum, the robust prediction of our theory concerns *the gap between citation rates of close and distant actors*, not close-actor citation rates by themselves or distant-actor citation rates alone. The gap between the two citation rates should have an inverted-U relationship with respect to interdependence.

## References

Alcacer, J., Gittelman, M., in press. How do I know what you know? Patent examiners and the generation of patent citations. Review of Economics and Statistics.

Allen, T.J., 1977. Managing the Flow of Technology: Technology Transfer and the Dissemination of Technological Information Within the R&D Organization. MIT Press, Cambridge, MA.

Argote, L., 1999. Organizational Learning: Creating, Retaining and Transferring Knowledge. Kluwer, Boston.

Arrow, K.J., 1962. Economic welfare and the allocation of resources for invention. In: Nelson, R. (Ed.), The Rate and Direction of Inventive Activity. Princeton University, Princeton, NJ, pp. 609–624.

Arrow, K.J., 1974. The Limits of Organization. Norton, New York.

Audretsch, D.B., Feldman, M.P., 1996. Innovative clusters and the industry life-cycle. Review of Industrial Organization 11, 253–273.

Baldwin, C.Y., Clark, K.B., 2000. Design Rules: The Power of Modularity. MIT Press, Cambridge, MA.

Baker, W.E., Faulker, R.R., 2004. Social networks and loss of capital. Social Networks 26, 91–111.

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

Bossard, J.S., 1932. Residential propinquity as a factor in marriage selection. American Journal of Sociology 38, 219–224.

Breschi, S., Lissoni, F., 2002. Mobility and Social Networks: Localized Knowledge Spillovers Revisited. Working paper. Bocconi University.

Burt, R.S., 1987. Social contagion and innovation: cohesion versus structural equivalence. American Journal of Sociology 92, 1287–1335.

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

Chew, W.B., Bresnahan, T., Clark, K., 1990. Measurement, coordination, and learning in a multiplant network. In: Kaplan, R.S. (Ed.), Measures for Manufacturing Excellence. Harvard Business School, Boston, pp. 129–162.

Christensen, C.M., Verlinden, M., Westerman, G., 2002. Disruption, disintegration, and the dissipation of differentiability. Industrial and Corporate Change 11, 955–993.

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

Cohen, W.M., Nelson, R.R., Walsh, J.P., 2000. Protecting their intellectual assets: appropriability conditions and why U.S. manufacturing firms patent (or not). W7552, National Bureau of Economic Research.

Coleman, J.S., Katz, E., Menzel, H., 1957. The diffusion of an innovation among physicians. Sociometry 20, 253–270.

Coleman, J.S., Katz, E., Menzel, H., 1966. Medical Innovation: A Diffusion Study. Bobbs-Merrill, New York.

Davis, G.F., Greve, H.R., 1980. Corporate elite networks and governance changes in the 1980s. American Journal of Sociology 103, 1–37.

Duguet, E., MacGarvie, M., 2005. How well do patent citations measure flows of technology? Evidence from French innovation surveys. Economics of Innovation and New Technology 14, 375–393.

Durkheim, E., 1912. The Elementary Forms of Religious Life.

Feld, S.L., 1981. The focused organization of social ties. American Journal of Sociology 86, 1015–1035.

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

Fleming, L., Sorenson, O., 2004. Science as a map in technological search. Strategic Management Journal 25, 909–928.

Friedrich, R., 1982. In defense of multiplicative terms in multiple regression equations. American Journal of Political Science 26, 797–833.

Gilfillan, S., 1935. Inventing the Ship. Follett, Chicago.

Griliches, Z., 1957. Hybrid corn: an exploration in the economics of technological change. Econometrica 25, 501–522.

Gulati, R., 1995. Social structure and alliance formation patterns: a longitudinal analysis. Administrative Science Quarterly 40, 619–652.

Hägerstrand, T., 1953. Innovation Diffusion as a Spatial Process. University of Chicago, Chicago.

Hall, B.H., Jaffe, A.B., Trajtenberg, M., 2001. The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools. National Bureau of Economic Research Working Paper No. 8498.

Hansen, M.T., 1999. The search-transfer problem: the role of weak ties in sharing knowledge across organization subunits. Administrative Science Quarterly 44, 82–111.

Hargadon, A., 1998. Diffusion of innovations. In: Dorf, R.C. (Ed.), The Technology Management Handbook. CRC/IEEE, Boca Raton, FL.

Hedström, P., 1994. Contagious collectives: on the spatial diffusion of Swedish trade unions. American Journal of Sociology 99, 1157–1179.

Henderson, R., Cockburn, I., 1996. Measuring competence? Exploring firm effects in pharmaceutical research. Strategic Management Journal 15, 63–84.

Homans, G.C., 1950. The Human Group. Harcourt, World and Brace, New York.

Irwin, D.A., Klenow, P.J., 1994. Learning-by-doing spillovers in the semiconductor industry. Journal of Political Economy 102, 1200–1227.

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

Jordan, K., Lynch, M., 1992. The sociology of a genetic engineering technique: ritual and rationality in the performance of the 'plasma prep'. In: Clarke, A., Fujimara, J. (Eds.), The Right Tools for the Job: At Work in the Twentieth Century Life Sciences. Princeton University Press, Princeton, pp. 77–114.

Kauffman, S.A., 1993. The Origins of Order. Oxford University, New York.

King, G., Zeng, L., 2001. Logistic regression in rare events data. Political Analysis 9, 137–163.

Kogut, B., Zander, U., 1992. Knowledge of the firm, combinative capabilities, and the replication of technology. Organization Science 3, 383–397.

Krugman, P.R., 1991. Geography and Trade. MIT, Cambridge, MA.

Lanjouw, J.O., Schankerman, M., 2004. Patent quality and research productivity: measuring innovation with multiple indicators. Economic Journal 114, 441–465.

Lazarsfeld, P.F., Berelson, B., Gaudet, H., 1944. The People's Choice: How the Voter Makes Up His Mind in a Presidential Election. Duell, Sloan, and Pearce, New York.

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

Levinthal, D., 1997. Adaptation on rugged landscapes. Management Science 43, 934–950.

Lim, K., 2001. The Many Faces of Absorptive Capacity: Spillovers of Copper Interconnect Technology for Semiconductor Chips. Working paper. Singapore National University.

Lippman, S., Rumelt, R., 1982. Uncertain imitability: an analysis of interfirm differences in efficiency under competition. Bell Journal of Economics 13, 418–438.

Mahajan, V., Muller, E., Bass, F.M., 1990. New product diffusion models in marketing: a review and directions for research. Journal of Marketing 54, 1–26.

Mansfield, E., 1968. Industrial Research and Technological Innovation. W.W. Norton, New York.

Manski, C.F., Lerman, S.R., 1977. The estimation of choice probabilities from choice based samples. Econometrica 45, 1977–1988.

March, J.G., Simon, H.A., 1958. Organizations. Blackwell, Cambridge, MA.

Marsden, P.V., Friedkin, N.E., 1993. Network studies of social influence. Sociological Methods and Research 22, 127–151.

Marshall, A., 1890. Principles of Economics. MacMillan, London.

McEvily, S.K., Chakravarthy, B., 2002. The persistence of knowledge-based advantage: an empirical test for product performance and technological knowledge. Strategic Management Journal 23, 285–306.

McPherson, J.M., Poplielarz, P.A., Drobnic, S., 1992. Social networks and organizational dynamics. American Sociological Review 57, 153–170.

McPherson, J.M., Smith-Lovin, L., Cook, J.M., 2001. Birds of a feather: homophily in social networks. Annual Review of Sociology 27, 415–444.

Mead, C., Conway, L., 1980. Introduction to VSLI Systems. Addison-Wesley, Reading, MA.

Nelson, R.R., Winter, S.G., 1982. An Evolutionary Theory of Economic Change. Belknap, Cambridge, MA.

O'Sullivan, A., 2001. Achieving Modularity: Generating Design Rules in an Aerospace Design-build Network. Working paper. University of Ottawa.

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

Park, R.E., 1926. The urban community as a spatial pattern and a moral order. In: Burgess, E.W. (Ed.), The Urban Community. University of Chicago, Chicago, pp. 3–18.

Perrow, C., 1984. Normal Accidents: Living With High-risk Technologies. Basic Books, New York.

Prentice, R.L., Pyke, R., 1979. Logistic disease incidence models and case-control studies. Biometrika 66 (3), 403–411.

Podolny, J.M., 1994. Market uncertainty and the social character of economic exchange. Administrative Science Quarterly 39, 458–483.

Polanyi, M., 1966. The Tacit Dimension. Anchor Day, New York.

Porter, M.E., Rivkin, J.W., 1999. Matching Dell. Harvard Business School Case, 158–799.

Reed, R., DeFillippi, R.J., 1990. Causal ambiguity, barriers to imitation, and sustainable competitive advantage. Academy of Management Review 15, 88–102.

Reinganum, J.F., 1981. Market structure and the diffusion of new technology. Bell Journal of Economics 12, 618–624.

Rivkin, J.W., 2000. Imitation of complex strategies. Management Science 46, 824–844.

Rivkin, J.W., 2001. Reproducing knowledge: replication without imitation at moderate complexity. Organization Science 12, 274–293.

Rogers, E.M., 1995. Diffusion of Innovations, 4th ed. Free Press, New York.

Romer, P., 1987. Growth based on increasing returns due to specialization. American Economic Review 77, 56–62.

Ryan, B., Gross, N.C., 1943. The diffusion of hybrid seed corn in two Iowa communities. Rural Sociology 8, 15–24.

Sampat, B.N. 2004. Examining Patent Examination: An Analysis of Examiner and Applicant Generated Prior Art. Working paper. Georgia Institute of Technology.

Scherer, F.M. 1984. Innovation and Growth: Schumpeterian Perspectives. Cambridge, MA.

Schumpeter, J., 1939. Business Cycles. McGraw-Hill, New York.

Scott, A.J., Wild, C.J., 1997. Fitting regression models to case-control data by maximum likelihood. Biometrika 84 (1), 57–71.

Singh, J., 2005. Collaboration networks as determinants of knowledge diffusion processes. Management Science 51, 756–770.

Simon, H.A., 1962. The architecture of complexity. Proceedings of the American Philosophical Association 106, 467–482.

Smith, J.K., Hounshell, D.A., 1985. Walter H. Corrothers and fundamental research at DuPont. Science 229, 436–442.

Sorenson, O., 2004. Social networks, informational complexity and industrial geography. In: Fornahl, D., Zellner, C., Audretsch, D. (Eds.), The Role of Labour Mobility and Informal Networks for Knowledge Transfer. Springer-Verlag, Berlin, pp. 79–96.

Sorenson, O., Fleming, L., 2004. Science and the diffusion of knowledge. Research Policy 33, 1615–1634.

Sorenson, O., Stuart, T.E., 2001. Syndication networks and the spatial diffusion of venture capital investments. American Journal of Sociology 106, 1546–1588.

Stern, S., 2001. Personal communication.

Strang, D., Soule, S.A., 1998. Diffusion in organizations and social movements: from hybrid corn to poison pills. Annual Review of Sociology 24, 265–290.

Stuart, T.E., Sorenson, O., 2003. The geography of opportunity: spatial heterogeneity in founding rates and the performance of biotechnology firms. Research Policy 32, 229–253.

Szulanski, G., 1996. Exploring internal stickiness: impediments to the transfer of best transfer within the firm. Strategic Management Journal 17, 27–43 (winter special issue).

Teece, D.J., 1977. Technology transfer by multinational firms: the resource cost of transferring technological know-how. Economic Journal 87, 242–261.

Tomz, M., 1999. Relogit (Stata ado file). Available at http://gking.harvard.edu/stats.shtml.

Tushman, M., Anderson, P., 1986. Technological discontinuities and organization environments. Administrative Science Quarterly 31, 439–465.

Ulrich, K., 1995. The role of product architecture in the manufacturing firm. Research Policy 24, 419–440.

Usher, A., 1954. A History of Mechanical Invention. Dover, Cambridge, MA.

von Hippel, E., 1988. The Sources of Innovation. Oxford University, New York.

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

Winter, S.G., 1995. Four Rs of profitability: rents, resources, routines, and replication. In: Montgomery, C. (Ed.), Resource-based and Evolutionary Theories of the Firm: Towards a Synthesis. Kluwer, Boston.

Womack, J.P., Jones, D.T., Roos, D., 1990. The Machine that Changed the World. Rawson, New York.

Zander, U., Kogut, B., 1995. Knowledge and the speed of transfer and imitation of organizational capabilities: an empirical test. Organization Science 6, 76–92.

Zimmerman, M.B., 1982. Learning effects and the commercialization of new energy technologies: the case of nuclear power. Bell Journal of Economics 13, 297–310.

Zucker, L.G., Darby, M.R., Brewer, M.B., 1997. Intellectual human capital and the birth of U.S. biotechnology enterprises. The American Economic Review 88, 290–306.