INTERFIRM COLLABORATION NETWORKS:
THE IMPACT OF SMALL WORLD CONNECTIVITY ON FIRM INNOVATION

Melissa A. Schilling
New York University
40 West Fourth Street
New York, NY 10012
212-998-0249
FAX: 212-995-4235
Email: mschilli@stern.nyu.edu

Corey C. Phelps
University of Washington
Box 353200
Seattle WA 98195
206-543-6579
Email: ephelps@u.washington.edu

This research supported by the National Science Foundation under Grant No. SES-0234075.
The authors are grateful for the suggestions of Juan Alcacer, Laszlo Barabasi, Joel Baum, Bill Greene,
Anne Marie Knott, Dan Levinthal, Bill McKelvey, Mark Newman, Joe Porac, Lori Rosenkopf, Rob
Salomon, Kevin Steensma, Kate Stovel, and Duncan Watts.

Version: December 2005

Please do not cite or reference without permission.
INTERFIRM COLLABORATION NETWORKS:
THE IMPACT OF SMALL WORLD CONNECTIVITY ON FIRM INNOVATION

Abstract

The structure of alliance networks strongly influences their potential for knowledge creation. Dense local clustering provides transmission capacity in the network by fostering communication and cooperation while non-redundant connections contract the distance between firms and give the network greater reach by tapping a wider range of knowledge resources. However, since firms are constrained in forming alliances, there appears to be a trade-off between creating transmission capacity versus reach. We argue that small world connectivity (i.e., simultaneity of high clustering and short average path lengths in a sparse, decentralized network) helps resolve this tradeoff by enabling transmission capacity and reach to be achieved simultaneously. We propose that firms embedded in alliance networks that exhibit high clustering and short average path lengths to a wide range of firms will experience greater knowledge creation than firms in networks that do not exhibit these characteristics. We find support for this proposition in a longitudinal study of the patent performance of 1106 firms in 11 industry-level alliance networks.
Interfirm networks are important engines of knowledge creation and innovation (Ahuja 2000; Freeman 1991). A particular type of interfirm relationship that has become increasingly common and received substantial scholarly attention in the last two decades is the strategic alliance (Gulati, 1998; Hagedoorn, 2002). Alliances enable firms to pool, exchange, and jointly create information and other resources (Gulati 1998). By providing member firms access to a wider range of resources than they individually possess, alliances enable firms to achieve much more than they could achieve individually (Eisenhardt & Schoonhoven 1996).

As firms form and maintain alliances with each other, they weave a network of direct and indirect relationships that enable them to access, disseminate and combine information. The specific pattern that such relationships exhibit represents the structure of the alliance network. The structure of an interfirm network influences the rate and extent of information diffusion through the network, including what types of information firms have access to and how readily they may access it (Rogers 1995; Valente 1995, Yamaguchi 1994). By influencing the rate and extent at which firms can access new information or recombine information in new ways, the structure of the interfirm network influences the utilization and creation of knowledge by the firms in the network (Kogut 2000; Powell, Koput & Smith-Doerr 1996).

While research has long recognized the importance of interfirm networks in firm innovation (see Freeman 1991 for a review), nearly all of this work has treated the network concept as a metaphor, rather than a construct with measurable properties. Only recently have researchers begun to assess the formal structural properties of alliance networks and their impact on firm innovation and knowledge acquisition. Most of this research has focused on a firm’s position within a broader network of relationships or the structure of its immediate network neighborhood rather than the structure of the overall network. For example, studies have examined a firm’s centrality (Smith-Doerr et al. 1999), number of alliances (Ahuja 2000; Deeds and Hill 1996; Shan, Walker, & Kogut, 1994) and the structure of its local network (Ahuja 2000; Baum, Calabrese & Silverman 2000). To our knowledge, empirical research has not yet examined
the impact of the structural properties of *industry-level* alliance networks on member firm innovation. However, in a related line of research, Uzzi and Spiro (2005) examined the network structure of creative artists who made Broadway Musicals from 1945 to 1989, and conclude that the structure of the collaboration network of these artists significantly influenced their creativity, and subsequently the financial and artistic performance of the musicals they produced. This raises the following questions: Does the structure of the industry-level interfirm network influence the rate of knowledge creation among firms in the network? If so, what structural properties are more likely to enhance innovation?

To address these questions, we draw upon recent graph theoretic research on “small-world” networks. A fundamental insight from small-world network research is that a high degree of clustering and a short average path length can coexist in a sparse network. That is, even if a network has relatively few links and many of those links create redundant paths in the network (as when a firm’s partners are also partners of each other), the average number of links required to connect all pairs of firms in the network can still be remarkably short. This finding has important implications for information diffusion. The dense connectivity of clusters creates *transmission capacity* in a network (Burt, 2001), enabling large amounts of information to rapidly diffuse, while short path lengths to a wide range of firms provides *reach* in the network, ensuring that diverse information sources can be tapped. In a network with small-world connectivity, there is almost no trade-off between information transmission capacity and reach -- both can coexist, even in a very sparse network. We argue that small-world properties in interfirm networks will significantly enhance the creative output of member firms, irrespective of other idiosyncratic differences. We test this hypothesis using longitudinal data on the innovative performance of a large panel of firms operating in 11 industry-level alliance networks.

This research offers several important contributions for understanding knowledge creation in interfirm networks, as well as knowledge networks and knowledge creation in general. First, we develop a theory relating the structural properties of industry-level interfirm networks to the innovative performance

---

1 An industry-level network is a specific type of whole network. Wellman (1988: 26) defined a whole network as the relationships that exist among members of a population.
of member firms. Other things being equal, structure matters, and small-world structures have stark advantages for knowledge creation. Second, we find empirical support for the theory in a longitudinal study of the patent output of 1106 firms in 11 industry-level alliance networks. To our knowledge, no other study has attempted to assess the effect of industry-level interfirm networks on the innovation performance of member firms. Third, whereas recent studies have demonstrated the existence of small-world network structures and their possible causes (Baum, Shipilov, & Rowley 2003; Davis, Yoo & Baker 2003; Kogut & Walker 2001; Watts 1999a), little research has examined the consequences of small-world structures in an industrial setting (Uzzi & Spiro 2005 is a recent exception).

We begin by describing recent work on small world networks and demonstrate its implications for diffusion and search within an interfirm network. From this we develop a hypothesis about how the structure of interfirm knowledge networks will influence the innovative output of member firms. We test the hypothesis on a large, unbalanced panel of firms embedded in 11 industry-level alliance networks.

SMALL-WORLD NETWORKS

Small world analysis has its roots in work by mathematical graph theorists (e.g., Erdos & Renyi 1959; Solomonoff & Rapoport 1951), but research specifically on the small-world phenomenon did not commence until the 1960s, when de Sola Pool and Kochen estimated both the average number of acquaintances that people possess and the probability of two randomly selected members of a society being linked by a chain of no more than two acquaintances (this work was published in 1978). At around the same time, psychologist Stanley Milgram was conducting an innovative empirical test of the small-world hypothesis (1967).

Milgram addressed a number of letters to a stockbroker friend in Boston. He distributed these letters to a random selection of people in Nebraska. He instructed the individuals to pass the letters to the addressee by sending them to a person they knew on a first-name basis who seemed in some way closer (socially, geographically, etc.) to the stockbroker. This person would then do the same, until the letters reached their final destination. Many of the letters (29%) did reach the stockbroker, and Milgram found that on average the letters had passed through about six individuals en route. Milgram had demonstrated
that the world was indeed small, and this finding was later dubbed “six degrees of separation” (Guare 1990).

If links in social networks were formed randomly, Milgram’s finding that the average path length across randomly chosen pairs of individuals is fairly short would not be surprising. If people chose their friends randomly, then the probability of any two individuals forming a relationship would be independent of any difference between them, (e.g., geographic distance or demographic dissimilarity) and independent of their friends’ choices (Bollobas, 1985). Thus, the likelihood of a farmer in Nebraska being a friend of a stockbroker in Boston would be equivalent to that of the farmer being friends with his next-door neighbor. Consistent with this logic, Bollobas (1985) showed that for a world with an arbitrarily large number of actors, each with an equal and limited number of ties, a random graph is a particularly good approximation of a structure that exhibits minimal average path length among actors

Similarly, if a single (or few) central nodes connected all other nodes in the network, it would again be expected that every pair of nodes would be connected by a relatively short path length through this central vertex.

Finally, if the number of links relative to the number of nodes were large, we would expect very short path lengths. As the number of links per node approaches the number of nodes in the network (i.e., maximum density), it becomes possible for all nodes to be directly connected to each other.

However, social networks are not random. Instead, they are highly clustered, with many local areas exhibiting significant redundancy (i.e., many of an individual’s acquaintances are also acquainted with each other). Furthermore, social networks tend to be decentralized and extremely sparse. No single individual connects all the others, and the maximum number of acquaintances of any individual in the network is a tiny fraction of the entire population (Watts 1999b). Intuitively, such clustered networks

---

2 Assume there is some number $N$ of people in the world, each of which has an average of $z$ acquaintances. This implies that there are $\frac{1}{2}Nz$ connections between people in the world. In a random graph, these $\frac{1}{2}Nz$ linkages are assigned randomly. Because a single person (node) on the graph has $z$ acquaintances, each person has $z^2$ acquaintances reachable in two steps, $z^3$ acquaintances reachable in three steps, and so on. Assuming people have, on average, between 100 and 1000 acquaintances, the number of acquaintances reachable in four steps ($z^4$) is between about $10^8$ and $10^{12}$ (roughly the world population). In general, the number of degrees of separation increases only logarithmically with the size of the network, causing the average path length to be very small even for very large networks (Bollobas, 1985).
should require a long path to connect individual nodes in different clusters with one another due to the sparseness of connections between clusters. Thus intuition might suggest that sparse and clustered networks would tend to be “large worlds” in that the average path length required to connect any two randomly chosen nodes is quite large. What made the findings of small world network research so surprising is that despite such clustering and sparsity, many real networks demonstrate remarkably short path lengths. Watts and Strogatz (1998) showed how this could occur: as a few random or long-spanning connections are added to a highly clustered network, the average path length drops far more rapidly than the degree of local clustering. In the range between highly clustered (locally ordered) networks and random networks there is an interval in which high clustering and short path lengths can coexist. As shown in figure 1, path length begins to drop sharply with only a few random links; the average degree of clustering only begins to decline after significantly more random links are added. In the interval between the drop in path length and the drop in clustering, high clustering and short path lengths coexist.

To better understand this, consider two stylized and extreme cases. The first is a network consisting of numerous highly clustered cliques that are connected to each other with only one link. Such a network is both highly clustered and extremely sparse. Watts (1999b) referred to such a network as a “connected caveman graph” and argued that it is an appropriate benchmark for a large, clustered graph (see Figure 2, panel a). The contrasting case is a random graph, which exhibits minimal clustering and represents a good approximation for a network with minimal average path length (Figure 2, panel c). Consistent with the intuition above, the connected caveman network has a very large average path length when compared with the random graph. However, highly clustered and globally sparse networks need not be large worlds. Watts and Strogatz (1998) demonstrated that by randomly “rewiring” a very small percentage of links in a highly clustered graph, the network exhibits the small-world properties of high clustering and short average path length. Because nodes that are initially widely separated in the network are as likely to become connected as those that are neighbors, the network’s average path length contracts as ties within clusters are replaced with ties that span them (Kogut & Walker 2001; Watts, 1999). In Figure 2, replacing
three of the links in panel a with randomly-generated links decreases the path length 34%, from five to 3.28, while its clustering coefficient decreases by only 12%, from .75 to .66 (Figure 1, panel b).³

The structure of networks greatly influences their dynamics. Watts (1999b) demonstrated how the topology of a small world network affects the degree to which a contagion (e.g., information, fashion, disease) diffuses throughout the network and the rate at which this diffusion occurs. Watts’ simulation results demonstrate that a contagion can spread completely and far more rapidly in a small world network than in a large world and nearly as fast as in a random network. Yamaguchi (1994) obtained similar results in his examination of the rate of information diffusion in a variety of network structures. Wilhite (2001) extended this work to an industrial setting by using a simulation to explore the impact of small-world properties on bilateral trade networks. Wilhite showed that small-world properties in a bilateral trade network enable agents to quickly find goods at the best price, resulting in an economy that reaches a Pareto optimal equilibrium more rapidly, and with lower search and negotiation costs than those incurred in alternative network structures. In sum, small-world connectivity increases the rate and extent of diffusion and the scope and efficiency of search.

**Structural Properties of Interfirm Networks**

Alliance networks also demonstrate sparsity, decentralization and clustering. First, interfirm networks tend to be extremely sparse because forming and maintaining alliances has a cost in terms of time and effort, and connections that are not reinforced over time diminish (Cummings 1991). When firms forge relationships with other organizations to share and exchange information and knowledge, they face a variety of search, monitoring, and enforcement costs (Williamson 1985). Firms face search costs to find alliance partners that are a good fit with the firm's objectives. Monitoring and managing alliances is also complex and costly, causing the firm’s effectiveness at managing its alliances to decline with the

³ The connected caveman graph is useful as a starting point for illustrating how randomly rewiring (or adding) a few links can greatly alter the average path length of a graph, yet largely preserves its degree of clustering. Watts (2004), however, critiques the empirical validity of this example and provides a thorough discussion of the network substrate and recent advances in the theoretical modeling of large-scale networks. We thank an anonymous reviewer for providing this insight.
number of alliances maintained (Deeds & Hill 1996). Thus, due to the cost constraints in forging and maintaining links, interfirm networks will tend to have far fewer links than if all pairs of firms were directly connected. Second, alliance networks tend to be decentralized. While interfirm networks often have “hub” firms that have very large numbers of connections (Barabasi 2002), most interfirm networks have several “hubs” rather than a single dominant firm that connects all other firms in the network (Baum et al. 2003; Gulati & Gargiulo 1999).

Finally, alliance networks tend to be highly clustered: some groups of firms will have more links connecting them to each other than to the other firms in the network. There are several mechanisms leading to clustering in interfirm knowledge networks, but two of the most common are linking based on similarity or complementarity. Firms tend to interact more intensely or frequently with other firms with which they share some type of proximity or similarity, such as geography or technology (Baum et al. 2003; Rosenkopf & Almeida, 2003; Saxenian 1994). This tends to result in a high degree of clustering.

Networks dominated by a high degree of local clustering and global sparsity often exhibit long path lengths, greatly reducing the overall efficiency of search and diffusion across the whole network (Watts; 1999a). Since clustering is achieved by forming redundant links and short path lengths are achieved by non-redundant links, we would expect alliance networks to exhibit either high clustering or short average path lengths. However, as described above, locally dense and globally sparse networks can also manifest short average path lengths. In the next section, we argue that alliance networks that exhibit such small-world properties will have a positive influence on member firm innovation.

Small-World Connectivity and Knowledge Creation

We adopt a recombinatory search perspective in explaining the process of innovation (Fleming, 2001; Katila & Ahuja 2002; Nelson & Winter, 1982). In the context of innovation, search refers to the attempts on the part of an actor to find or discover a solution to a problem. In this way, innovation is characterized as a problem-solving process in which solutions to problems are discovered via search (Dosi 1988; Vincenti 1991). A long line of research suggests that the search process that leads to the creation of new knowledge, embodied in artifacts such as patents and new products, most often involves
novel recombination of known elements of knowledge, problems, or solutions (Fleming 2001; Gilfillan 1935; Nelson and Winter 1982; Schumpeter 1934; Usher 1954) or the reconfiguration of the ways in which knowledge elements are linked (Henderson and Clark 1990). Critical inputs into this recombinatory process include access to and familiarity with: a variety of knowledge elements (e.g. different technological components and the scientific and engineering know-how embedded in them), novel problems and insights into their resolution, failed recombination efforts, and successful solutions (Hargadon, 2002). Firms that have greater access to and understanding of these recombinatory resources should produce more novel knowledge than other firms.

While firms have tended to pursue the creation of commercially-valuable knowledge through internal research and development activities, organizations have increasingly relied upon extramural sources of knowledge for innovation in the form of strategic alliances (Hagedoorn 2002). As firms form and maintain alliances with each other, they weave a network of direct and indirect relationships. As a result, firms embedded in these networks gain access to information and know-how of direct partners and that of others in the network to which they are indirectly connected (Ahuja, 2000; Gulati & Gargiulo 1999). The network of alliance relationships constitutes a conduit that channels the flow of information and know-how among firms in the network (Ahuja, 2000; Owen-Smith & Powell, 2004), with each member firm acting as both a recipient and transmitter of information (Ahuja, 2000; Rogers & Kincaid, 1981). What types of information, knowledge and other resources flow through these networks and how does this matter for the recombinatory efforts and innovativeness of member firms?

Alliance relationships provide for the social interaction of personnel from two or more firms for a particular purpose (e.g., joint research and development of a new product). As such, they can benefit the recombinatory search efforts of partnered firms in multiple ways. Alliances typically involve some degree of knowledge sharing between the partners, yielding a greater pool of knowledge each firm has to draw on for its recombinatory efforts relative to going it alone. Collaboration increases the depth and diversity of complementary knowledge available to partners’ innovation efforts (Mowery, Oxley & Silverman, 1996; Richardson, 1972; Teece, 1992). Searching diverse domains of knowledge increases the number of
knowledge elements available for recombination, increasing their combinatorial possibilities (Fleming, 2001). Due to the increased interpersonal interaction, enhanced incentives alignment, and monitoring features they provide, alliances are institutions better suited than market transactions for the repeated exchange of tacit, embedded knowledge on a reciprocal basis (Kogut, 1988; Richardson, 1972; Teece, 1992). Access to partners with different knowledge and experience can provide individuals involved in the alliance with multiple interpretations of technical problems and solutions, resulting in increased cognitive variety and quicker identification of potential recombinatorial solutions (March, 1991; Noteboom, 1999). Alliances also enable partners to learn about each other’s failed innovation attempts and dead-ends and identify new projects to undertake (i.e., problems to solve) (Ahuja 2000; Powell et al. 1996). Research by Schrader (1992) and Rogers and Larsen (1984) provides rich qualitative evidence to support these arguments. Large sample studies in different industrial settings have found that alliances facilitate knowledge flows between partners (Gomes-Casseres, Hagedoorn & Jaffe, forthcoming; Mowery et al. 1996) and enhance the innovative performance of firms (e.g., Deeds & Hill, 1996; Sampson, 2005; Stuart, 2000).

Indirect ties in an alliance network can also be beneficial for a firm’s recombination efforts. Ahuja (2000: 430) identifies two primary benefits of indirect alliances relationships for a firm’s innovation efforts. First, indirect ties can provide firms with timely information about the success and failure of numerous innovation efforts of other firms, suggesting that networks act as information gathering mechanism. Second, networks can act as information-processing devices because each additional firm to which a focal firm is indirectly connected can provide alternative interpretations of information about new solutions, opportunities, or failed innovation efforts and imbue this information with new meaning. Firms are thus better able to pursue promising opportunities, avoid mistakes made by others and learn about novel recombination approaches (Hargadon, 2002). Thus, the network can provide a focal firm insight into the efficacy of its own recombinatory efforts. This information processing capability is likely to exceed that of individual firms (Ahuja, 2000). In addition to these benefits, firms can actively search the network to identify source firms that possess the needed information or know-how
for their particular recombinatory efforts.

Though firms may go to great lengths to protect their proprietary information from being transmitted beyond a particular collaboration agreement, much of the information exchanged between firms is considered nonproprietary and thus is not deliberately protected from diffusion. For example, firms engaged in technological collaboration might freely exchange information about their suppliers, potential directions for future innovation, scientific advances in other fields that are likely to impact the industry, etc. Other information exchanged between firms is considered proprietary but is imperfectly protected from diffusion. Even when collaboration agreements have extensive contractual clauses designed to protect the proprietary knowledge possessed by each partner or developed through the collaboration, it is still very difficult to prevent that knowledge from ultimately benefiting other organizations. Secrecy clauses are very difficult to enforce when knowledge is dispersed over a large number of employees or embedded in visible artifacts. The alliance network thus enables a wide range of information – even some that would be considered proprietary technological information – to diffuse to (or be sought out by) firms connected to the network. These firms then, in turn, seek to integrate or recombine the information in ways that create new knowledge, and embody that knowledge in novel and useful innovations.⁴ Consistent with these arguments, prior research shows that the extent to which a firm is indirectly connected to other firms in an alliance network enhances its production of patents and new products (Ahuja, 2000; Owen-Smith & Powell, 2004; Smith-Doerr et al., 1999; Soh, 2003).

Given the role of direct and indirect ties as channels for the flow of information and other resources, we argue that the structure of the interfirm network will significantly influence the recombination process. Integrating the small-world network ideas outlined earlier with existing sociological research on network structure indicates that interfirm networks with small-world network properties should have significant advantages relative to other global network structures in enabling knowledge creation by networked firms.

⁴ Firms may, of course, also use new knowledge in a variety of ways other than creating new innovations; we focus on innovation here because of its significant role in firm performance and economic growth.
Clustering (i.e., local density) increases the information transmission capacity of a network. First, the dense connectivity of individual clusters ensures that information introduced into a cluster will quickly reach other firms in the cluster. The multiple pathways between firms in the cluster enhance not only the speed of information transmission, but also the fidelity of the information received. Firms can compare the information received from multiple partners, helping them to identify ways in which it has been distorted or is incomplete. Second, and related to the previous point, clusters within networks are important structures for making information exchange meaningful and useful. The internal density of a cluster can increase the dissemination of alternative interpretations of problems and their potential solutions, deepening the collective’s understanding and stimulating collective problem-solving (Powell & Smith-Doerr 1994). The development of a shared understanding of problems and solutions greatly facilitates communication and further learning (Brown & Duguid 1991; Powell et al. 1996). Third, dense clustering can make firms more willing and able to exchange information (Uzzi & Spiro, 2005). Sociologists (e.g., Coleman 1988; Granovetter 1992) have suggested that densely clustered networks give rise to trust, reciprocity norms, and a shared identity, all of which lead to a high level of cooperation and can facilitate collaboration by providing self-enforcing informal governance mechanisms (Dyer & Singh 1998). In addition to stimulating greater “transparency” (Hamel 1991), trust and reciprocity exchanges facilitate intense interaction among personnel from partnered firms (Uzzi 1997), improving the transfer of tacit, embedded knowledge (Hansen 1999; Zander & Kogut 1995). Thus, clusters enable richer and greater amounts of information and knowledge to be exchanged and integrated more readily.

Fourth, when dense clusters are only sparsely connected to each other, they become important structures for creating and preserving the requisite variety of knowledge in the global network that enables long-run knowledge creation. As noted by several authors, the internal cohesion of a cluster can lead much of the information and knowledge shared within a cluster to become homogeneous and redundant (Burt 1992; Granovetter 1973; Rosenkopf & Almeida, 2003; Uzzi & Spiro, 2005). The dense links provide many redundant paths to the same actors and thus the same sources of information and knowledge. Worse still, norms of adhering to established standards and conventions can potentially stifle
experimentation and creativity (Uzzi & Spiro, 2005). This limits innovation. On the other hand, clusters of firms will tend to be highly heterogeneous across a network with respect to the knowledge they possess and produce due to the different initial conditions and causes for each cluster to form. The diversity of knowledge distributed in different clusters across the network provides the requisite variety for recombination.

Clustering thus offers both local and global advantages. Firms benefit from having redundant connectivity among their immediate neighbors because it enhances the speed and likelihood of information access, and the depth of information interpretation. Firms also benefit from being embedded within a larger network that is clustered because the information a firm receives from partners that are embedded in other clusters is likely to be more complete and richly understood than information received from partners not embedded in clusters, and because information received from different clusters is likely to be diverse, enabling a wider range of recombinatorial possibilities.

The importance of combining the diverse information distributed across clusters points to the importance of shortcuts between clusters. As Uzzi and Spiro (2005) note in their study of artistic collaboration in Broadway plays, bridges between clusters increase the likelihood that different ideas and routines will come into contact, enabling recombinations that incorporate both previous conventions and novel approaches. Similarly, interfirm networks that contain bridges between clusters of firms provide member firms access to diverse information that exists beyond their local cluster, enabling new combinations with their existing knowledge sets. The number and distribution of these shortcuts strongly influences the average path length of the overall network. As discussed previously, the diffusion of information and knowledge occurs more rapidly and with more integrity in networks with a short average path length than in networks with longer paths (Watts 1999b). A firm that is connected to a large number of firms by a short average path can reach more information, and can do so quickly and with less risk of information distortion than a firm that is connected to fewer firms or by longer paths.5

5 Note that network size is a crucial factor here – short average path lengths in a small network do not afford a firm the same information reach as short path lengths in a large network.
Since forming alliances is costly and constrained, there appears to be a trade-off between forming dense clusters to facilitate rapid exchange and integration of knowledge, versus forging links to create short paths to a wider range of firms. Small world network properties help to resolve this tradeoff by enabling both dense clustering and wide reach to coexist, even in a sparse and decentralized network. By forming a relatively small number of random or atypical inks that provide bridges between clusters, interfirm networks can achieve a short path length to diverse knowledge sources (i.e., reach) while retaining a high degree of clustering (Hansen 2002; Hargadon 1998). The combination of clustering and reach enables a wide range of information to be exchanged and integrated rapidly, leading to greater knowledge creation. In sum, we predict a multiplicative interaction between clustering and reach in their effect on firm knowledge creation. Consistent with the symmetrical nature of such interactions (Jaccard & Turrisi, 2003), we have argued and expect that the effect of clustering on firm knowledge creation will be increasingly positive as reach increases, while the effect of (increases in) reach on knowledge creation will be increasingly positive as clustering increases.

**Hypothesis:** Firms participating in alliance networks that combine a high degree of clustering and short average path lengths to a wide range of firms will exhibit significantly more knowledge creation than firms in networks that do not exhibit these characteristics.

**METHODS**

To test our hypothesis, we constructed a large, unbalanced panel of U.S. firms for the period 1990-2000. The panel includes all U.S. firms that were part of the alliance networks of 11 high technology manufacturing industries: Aerospace equipment (SICs: 3721, 3724, 3728, 3761, 3764, 3769); Automotive Bodies and Parts (3711, 3713, 3714); Chemicals (281-, 282-, 285-, 286-, 287-, 288-, 289-); Computer and Office Equipment (3571, 3572, 3575, 3577); Household Audiovisual Equipment (3651); Medical Equipment (3841, 3842, 3843, 3844, 3845); Petroleum Refining and Products (2911, 2951, 2952, 2992, 2999); Pharmaceuticals (2833, 2834, 2835, 2836); Semiconductors (3674); Telecommunications
Equipment (366-), and Measuring and Controlling Devices (382-). The database also contains the panel firms’ patenting activity.

The choice of industries was of particular importance to this study. The eleven industries selected have been designated as high tech in numerous Bureau of Labor Statistics studies (e.g., Hecker 1999; Luker & Lyons 1997). To be considered high tech, the industry’s employment in both research and development and other technology-oriented occupations must be at least twice the average for all industries in the Occupational Employment Statistics Survey. This set of industries provides an excellent context for the current study for three reasons. First, the creation of knowledge is fundamental to the pursuit of competitive advantage in high technology industries (Teece, Pisano & Shuen 1997). Second, firms in each of these industries make active use of alliances in pursuit of their innovation activities (Hagedoorn 1993; Vonortas 1997). Third, because we use patent data for our dependent variable, it is important to select industries that use patents. There is evidence that firms in these industries actively patent their intellectual property (Levin et al. 1987).

Alliance Networks

We chose to measure the network structure created by publicly-reported strategic alliances for two reasons. First, there is a rich history of research on the importance of strategic alliances as a mechanism for knowledge sharing among firms (Freeman 1991; Gulati 1998; Hamel 1991; Powell et al. 1996). Second, alliances are used by a wide range of firms (including both public and private firms) in a wide range of industries, and are often used explicitly for the exchange and joint creation of knowledge.

Determining the boundaries of interfirm networks is a nontrivial issue (Marsden 1990). Prior research in social networks has identified three procedural tactics for establishing network boundaries for empirical research: attributes of actors that rely on membership criteria, such as membership in an industry; types of relations between actors, such as participation in strategic alliances; and participation in a set of common events (Laumann, Marsden, Prensky 1983). Following these prescriptions, we employed

---

6 We removed high tech manufacturing industries that make very little use of alliances: special-industry machinery (355), electrical industrial apparatus (362), search & navigation equipment (381), and photographic equipment & supplies (386).
two rules to guide our construction of the 11 industry networks analyzed in this study. First, each alliance included at least one participant that was a member of the target industry (indicated by its primary four-digit SIC). Second, to be included in the target industry network each alliance had to operate in that industry, as indicated by its primary four-digit SIC of activity. These rules help to ensure that the industry networks consist of alliance activity focused on the designated industry. Because an industry member’s partners can come from both within or beyond its industry, there is some overlap between the alliance networks. Notably, some well-known firms such as IBM, Hewlett Packard, AT&T, and General Motors appear as alliance partners in multiple networks. We included alliance partners from beyond the target industry because excluding them would eliminate our ability to observe many of the indirect relationships between industry members, thus biasing our measures of network connectivity. Recent alliance research has employed similar network construction criteria (Rowley, Behrens, & Krackhardt 2000).

Alliance data were gathered using Thomson Corp.’s SDC Platinum database. The SDC data have been used in a number of empirical studies on strategic alliances (e.g., Anand & Khanna 2000; Sampson 2004). For each industry, alliances were collected that were announced between 1990 and 1997. We chose 1990 as the initial year for our sample because information on alliances formed prior to 1990 is very sparse in the SDC database. Separate alliance networks were created for each industry according to the primary SIC code of the alliance. Both public and private firms were included. We chose to use data on only U.S. firms because the SDC data on alliances is much more complete for U.S. firms than for non-U.S. firms (Phelps 2003). Furthermore, to avoid overlooking alliances formed by subsidiaries, all alliances were aggregated to the parent corporation.

The resulting data set includes 1106 firms involved in 3,517 alliances. Many of the alliance announcements included more than two participating firms, so the number of dyadic alliance pairs is much higher, totaling 5,306. Since any type of alliance may provide a path for knowledge diffusion, and because prior studies indicate that the breadth of an alliance’s true activity is often much greater than what is formally reported (Powell et al. 1996), we include all alliance types in our analysis. However, it is

---

7 SDC did not undertake systematic collection of alliance data until around 1989 (Anand & Khanna, 2000: 300).
also reasonable to assume that an alliance formed specifically for the purpose of joint research and
development or technology exchange might have more impact on innovation than, for example, a supply
agreement or marketing alliance. We explore this possibility by including a measure of the proportion of
alliances that are coded as R&D, cross-technology transfer or technology licensing agreements.

Alliance relationships typically last for more than one year, but alliance termination dates are rarely
reported. This required us to make an assumption about alliance duration. We took a conservative
approach and assumed that alliance relationships last for three years, consistent with recent empirical
work on the average duration of alliances (Phelps 2003). Other research has taken a similar approach,
using windows ranging from one to five years (e.g., Gulati & Gargiulo 1999; Stuart 2000). We created
six snapshots of the network structure for each industry, for a total of 66 alliance network snapshots. Each
network snapshot was constructed as a binary adjacency matrix. Since we are concerned with whether a
path exists from one firm to another and not with the effect of multiplex relationships, multiple alliance
announcements between the same pair of firms in any time window are treated as one link. Alliance
relationships are considered to be bidirectional, resulting in an undirected unipartite graph (Wasserman &
Faust, 1994). UCINET 6.23, a leading social network analysis software package, was used to obtain
measures on each of these networks, as described below (Borgatti, Everett, & Freeman 2002).

As we focus on publicly-reported contractual alliance agreements, we do not observe the numerous
informal collaborative arrangements that exist between firms in our sample. Such informal arrangements
often lead to the types of formal agreements that we observe (Powell, Koput & Smith-Doerr 1996;
Rosenkopf, Metiu & George 2001). Consequently, our analysis represents a conservative test of our small
world diffusion argument because our data do not include widely used informal relationships that promote
knowledge transfer.

**Dependent Variable: Patents**

---

8 A binary adjacency matrix is a square matrix with nodes (e.g., firms) as rows and columns. The entries in the
adjacency matrix, \( x_{ij} \), indicate which pairs of nodes are adjacent (i.e., have a relationship). In a binary matrix, a value
of 1 indicates the presence of a relationship between nodes \( i \) and \( j \), while a 0 indicates no relationship.
One way that knowledge creation is instantiated is in the form of inventions (Schmookler 1966). Knowledge embedded in artifacts such as inventions can be seen as the “empirical knowledge” of organizations (Hargadon & Fanelli 2002). As such, inventions provide a trace of an organization’s knowledge creation activities. Patents provide a measure of novel invention that is externally validated through the patent examination process (Griliches 1990). Patent counts have also been shown to correlate well with new product introductions and invention counts (Basberg 1987). Indeed, Trajtenberg (1987) concluded that patents are perhaps the most valid and robust indicators of knowledge creation. One of the challenges with using patents to measure innovation is that the propensity to patent may vary with industry sector, resulting in a potential source of bias (Levin et al. 1987). We have addressed this potential bias in three ways. First, we have chosen only high tech manufacturing industries, which helps to ensure a degree of commonality in the industries’ emphasis on innovation. To further capture differences in emphasis on innovation and/or complexity of innovation, we control for industry-level R&D intensity as described in the controls section. Third, to control for other unobserved factors that influence the propensity to patent (e.g., appropriability regimes, etc.) that are likely to be stable within industries (Griliches 1990), we control for industry fixed effects. The propensity to patent may also differ due to firm characteristics (Griliches 1990). We attempt to control for such sources of heterogeneity using a covariate, Presample Patents (described below), and the inclusion of firm fixed and random effects in our estimations.

We measure the dependent variable, Patents$_{it}$, as the number of successful patent applications for firm $i$ in year $t$. For each year, patent data was collected for every firm in the network whose primary SIC code matched the industry, consistent with the way we formed the alliance networks. We used the Delphion database to collect yearly patent counts for each of the firms, aggregating subsidiary patents up to the ultimate parent level. While only patents that were ultimately granted were counted, patents were counted in the year of application. We do so because the time between application and grant varies across patents and using the date of application more precisely captures the time of knowledge creation (Griliches 1990). Yearly patent counts were created for each firm for the time range of 1993 to 2000,
enabling us to assess different lag specifications between alliance network structure and patent output.

**Independent Variables**

**Clustering Coefficient.** To measure the clustering in each network for each time period we used the weighted overall clustering coefficient measure (Borgatti et al. 2002; Newman, Strogatz & Watts, 2002). This measure indicates the transitive closure of a graph and is defined as:

\[
\text{Clustering}_w = \frac{3 \times (\text{number of triangles in the graph})}{(\text{number of connected triples})}
\]

Where a *triangle* is a set of three nodes (e.g., \(i, j, k\)), each of which is connected to both of the others, and a *connected triple* is a set of three nodes in which at least one is connected to both the others (e.g., \(i\) is connected to \(j\) and \(k\), but \(j\) and \(k\) need not be connected). This measure indicates the proportion of triples for which transitivity holds (i.e., if \(i\) is connected to \(j\) and \(k\), then by transitivity, \(j\) and \(k\) are connected). The factor of 3 in the numerator ensures that the measure lies strictly in the range of 0 and 1 because each triangle implies 3 connected triples. This is a “weighted” measure in that each node’s contribution to the overall clustering coefficient is weighted by its number of links (i.e., degree).

In the present context, the weighted overall clustering coefficient is measured as the percentage of a firm’s alliance partners that are also partnered with each other, weighted by the number of each firm’s partners, averaged across all firms in the network. This variable can range from 0 to 1, with larger values indicating increasing clustering. While network density captures *global* density (or sparsity) of the entire network, the clustering coefficient captures the degree to which the overall network is characterized by localized pockets of dense connectivity. A network can be quite sparse globally, and still have a high clustering coefficient.

**Reach.** To capture the reach of each network for each time period, we use a measure of average distance-weighted reach (Borgatti et al. 2002; Borgatti, forthcoming). This is a compound measure that takes into account both the number of firms that can be reached by any path from a given firm, and the path length it takes to reach them. This measure is calculated as:
Average distance weighted reach = \[ \frac{\sum_{n} \sum_{j} 1/d_{ij}}{n}, \]

where \( n \) is the number of nodes in the network, and \( d_{ij} \) is defined as the minimum distance (geodesic), \( d \), from a focal node \( i \) to partner \( j \), where \( i \neq j \). Average distance-weighted reach can range from 0 – \( n \), with larger values indicating higher reach (and smaller average path lengths). Thus, for each node, the measure counts how many other nodes that can be reached by any path, and then divides that number by the average length of those paths. This number is then averaged across all of the nodes in the network. A simple example illustrating the measurement of reach is provided in Appendix A.

A significant advantage of using the reach measure is that it provides a meaningful measure of the overall connectivity of a network, even when that network has multiple components and/or component structure is changing over time. It avoids the infinite path length problem associated with disconnected networks by measuring only the path length between connected pairs of nodes and it provides a more meaningful measure than the simple average path length between connected pairs by factoring in the size of connected components.\(^9\) Since our networks are characterized by multiple components that merge and split apart over time (which we discuss in the results section), this is an important advantage of the reach measure. Furthermore, the reach measure better captures our conceptual argument that short path lengths to a wide range of firms enable firms to have greater reach to a diverse information sources than would a measure of path length that was not scaled by component size.

**Clustering X Reach.** Small-world connectivity enables high levels of reach and clustering to coexist. To capture this, we include the interaction term, *Clustering X Reach*. We expect a positive sign on the estimated coefficient for this variable.

**Firm-Level Control Variables**

**Pre-sample Patents.** To control for unobserved heterogeneity in firm patenting (due, for example, to differences in R&D expenditures, propensity and/or ability to patent, etc.), we follow the pre-sample

\(^9\) We are grateful to Steve Borgatti for pointing this out. We are also grateful to Mark Newman for numerous discussions about how to handle the infinite path length consideration in our networks.
information approach of Blundell, Griffith and Van Reenen (1995) and calculate the variable \textit{Pre-sample Patents} as the sum of patents obtained by a firm in the five years prior to its entry into the sample.

\textbf{Betweenness Centrality.} Research has found that firms that occupy more central positions in alliance networks tend to generate more innovations than more peripheral firms (e.g., Owen-Smith & Powell, 2004; Smith-Doerr et al., 1999; Soh, 2003). The theoretical explanation provided for these findings is that centrally located organizations benefit from substantial and diverse knowledge flows as a function of their connectedness, both directly and indirectly, to a larger number of companies than more peripheral firms have access to. We use the variable \textit{Centrality} to control for the time-varying influence of a firm’s network centrality on its subsequent patenting. Following Owen-Smith & Powell (2004), we operationalize \textit{Centrality} using Freeman’s (1977; 1979) measure of “betweenness centrality,” which captures the extent to which a firm is located on the shortest path (i.e., geodesic) between any two actors in its alliance network. Betweenness centrality indicates an actor’s ability to access diverse information flows and serve as a gatekeeper or broker of such information (Freeman, 1979). Formally, betweenness centrality for firm \(i\) in year \(t\) is calculated as:

\[
\text{Betweenness Centrality}_{it} = \sum_{j<k} g_{jk}(n_i) / g_{jk},
\]

where \(g_{jk}(n_i)\) refers to the number of geodesics (i.e., shortest paths) linking firms \(j\) and \(k\) that contain focal firm \(i\). The term \(g_{jk}(n_i) / g_{jk}\) captures the probability that firm \(i\) is involved in the shortest path between \(j\) and \(k\). Betweenness centrality is the sum of these estimated probabilities over all pairs of firms (not including the \(i\)th firm) in the network.

We use normalized betweenness centrality (i.e., betweenness divided by maximum possible betweenness, expressed as a percentage) to make the measure comparable across time and industry networks.\(^{10}\) Normalized betweenness centrality for firm \(i\) in year \(t\) is calculated as:

\[^{10}\text{We use betweenness centrality rather than a simple count of alliances (i.e., degree centrality) to assess the influence that indirect paths of information flows may have on firm innovation (Ahuja, 2000). In unreported analyses, we found that betweenness centrality and degree centrality were highly correlated (}r=0.70\text{) and that their effects in each of our estimated models were qualitatively similar. We did not employ closeness centrality because this index is only meaningful for a completely connected graph (which our networks are not) (Wasserman & Faust, 1994: 185).\]
Normalized Betweenness Centrality,\( \text{NBC}_{i} = 100 \times \frac{[(\text{Betweenness Centrality})/((g-1)(g-2)/2)]}{}, \)

where \( (g-1)(g-2)/2 \) is the number of pairs of firms, not including \( i \).

**Local Efficiency.** The extent to which a firm’s alliance partners are non-redundant (i.e., not partnered with each other) has also been shown to influence firm innovation (e.g., Ahuja, 2000; Baum et al., 2000). In such ego (or “local”) networks, non-redundant partners are indicative of structural holes (Burt, 1992). While studies have found that the extent to which a firm’s partners are nonredundant enhances its knowledge creation (Baum et al., 2000; McEvily & Zaheer, 1999), other research shows that redundant links improve knowledge transfer and innovation (Ahuja, 2000; Dyer & Nobeoka, 2000). Although the empirical evidence is mixed, controlling for the affect of local structural holes is important if we wish to demonstrate that the global structure (i.e., small worldliness) of the alliance network in which a firm is embedded has an independent and significant influence on its subsequent patenting. We control for the influence of a firm’s local network structure using Burt’s (1992) measure of efficiency. Efficiency captures the extent to which a firm’s partners are nonredundant, indicating the presence of structural holes in a firm’s (ego’s) network. Formally, local efficiency for firm \( i \) in year \( t \) is computed as follows:

\[
\text{Local Efficiency}_{it} = \sum_{j} \left[ 1 - \sum_{q} p_{iq} m_{jq} \right] / N_i, \quad j \neq q,
\]

where \( p_{iq} \) is the proportion of \( i \)’s relations invested in the relationship with \( q \), \( m_{jq} \) represents the marginal strength of the relationship between alter \( j \) and alter \( q \) (as we use binary data, all values of \( m_{jq} \) are set to 1 if the relationship is present and 0 otherwise), and \( N_i \) represents the number of unique alliance partners to which the focal firm \( i \) is connected. This measure varies across firms and time and can vary from 0 to 1, with higher values indicating greater efficiency.

**Industry (Network) Control Variables**

**Network Density.** We control for the overall density of the network with the variable *Network Density*, calculated for each industry network and time period. We do so because the rate and extent to which information diffuses increases with density (Yamaguchi 1994). This variable measures the ratio of
existing links in the network to the number of possible links (i.e., all possible pairwise combinations of firms), and may range from 0 to 1, with larger values indicating increasing density and lower values indicating sparsity.

**Centralization.** The extent to which a network is highly centralized can also influence its diffusion properties. A highly centralized network is one in which all ties run through one or a few nodes, thus decreasing the distance between any pair of nodes (Wasserman & Faust 1994). To control for network centralization, we employ Freeman’s (1979) index of group betweenness centralization, calculated for each industry network and time period. Group betweenness centralization for industry network \( j \) in year \( t \) is calculated as follows:

\[
\text{Betweenness Centralization}_{jt} = 100 \times \left\{ \sum_{i=1}^{g} [C_B'(n^*) - C_B'(n_i)]/(g-1) \right\},
\]

where \( C_B'(n^*) \) is the largest realized normalized betweenness centrality for the set of firms in network \( j \) in year \( t \), \( C_B'(n_i) \) is the normalized betweenness centrality for firm \( i \) (in industry network \( j \) for year \( t \)), and \( g \) is the number of firms (in industry network \( j \) for year \( t \)). This variable is expressed as a percentage and can range from 0, where all firms have the same individual betweenness centrality, to 100, where one firm connects all other firms (i.e., a star graph).

**Industry R&D Intensity.** To control for differences in the emphasis and costliness of innovation across industries, we employ a time-varying measure of industry-level R&D intensity (R&D expenditures/Sales), updated annually. To construct this variable, we collected the annual R&D expenditures and sales of firms in each industry from Compustat. Since our alliance networks include both public and private firms it would have been preferable to use R&D intensity data on both public and private firms. However, R&D expenditures for privately held firms are rarely available. We assume that aggregate R&D intensity for public firms is a good proxy for industry-level R&D intensity.

**Proportion of Alliances for R&D, Cross-Technology Transfer, or Licensing.** While all types of alliances are potential conduits for information about technologies, market opportunities, manufacturing processes, etc., alliances that are established for the purpose of conducting joint R&D activities, cross-
technology transfer, or licensing agreements might be more directly related to rates of patented innovation. To examine this possibility, we include a time-varying measure of the percentage of alliance agreements in each network that were established explicitly for the purpose of joint research and development, cross-technology transfer or technology licensing.

**Model Specification**

The dependent variable in this study, *Patents*, is a count variable and takes on only non-negative integer values. The linear regression model is inadequate for modeling such variables since the distribution of residuals will be heteroscedastic non-normal. A Poisson regression approach is appropriate to model count data (Hausman, Hall & Griliches 1984). However, the Poisson distribution contains the strong assumption that the mean and variance are equal, implying the absence of unobserved cross-sectional heterogeneity. Patent data often exhibit overdispersion, where the variance exceeds the mean (e.g., Ahuja 2000; Hausman et al. 1984; Henderson & Cockburn 1996). In the presence of overdispersion, coefficients will be estimated consistently but their standard errors will generally be underestimated, leading to spuriously high levels of significance (Cameron & Trivedi 1986). Each model that we report, when estimated using the Poisson specification, exhibited significant overdispersion.\(^{11}\)

A commonly used alternative to the Poisson regression model is the negative binomial model. The negative binomial model is a generalization of the Poisson model and allows for overdispersion by incorporating an individual, unobserved effect into the conditional mean (Hausman et al. 1984). The panel data implementation of the negative binomial model accommodates explicit control of persistent individual unobserved effects through both fixed and random effects. In the present study, unobserved heterogeneity refers to the possibility that unmeasured (or unmeasurable) differences among observationally equivalent firms affects their patenting. Unobserved heterogeneity may also stem from unmeasured, systematic time period and industry effects. Failing to control for such unobserved heterogeneity, if present, can result in specification error (Heckman 1979).

We employ a number of strategies to control for these sources of unobserved heterogeneity.\(^{11}\)

\(^{11}\) We used Cameron and Trivedi’s (1990) $T_{opt}$ diagnostic as implemented in Limdep 8.0 to test for overdispersion.
First, we include year fixed effects to control for systematic period effects such as differences in macroeconomic conditions or technological opportunity that may affect all sampled firms’ patent rates. Second, we employ individual firm effects to control for firm-specific unobserved heterogeneity. Firm effects serve as a control for temporally stable, unobserved firm-level differences in patenting performance. We use both firm fixed effects and firm random effects in alternative estimations of our model. The use of firm fixed and random effects in the negative binomial model allows for a firm-specific variance to mean ratio. We use Hausman et al.’s (1984) implementation of fixed effects in the context of a negative binomial model, which employs a conditional maximum likelihood estimation procedure\textsuperscript{12}. We also use Hausman et al.’s random effects specification, which assumes that overdispersion due to unobserved heterogeneity is randomly distributed across firms\textsuperscript{13}. Because the random effects specification assumes that the unobserved firm specific effect is uncorrelated with the regressors, we report the results from both fixed and random effects as a robustness check.

\textsuperscript{12} Allison and Waterman (2002) recently criticized Hausman et al.’s (1984) conditional negative binomial fixed effects model as not being a “true” fixed effects method in that it does not control for all time invariant covariates. Allison and Waterman (2002) developed an unconditional negative binomial model that uses dummy variables to represent fixed effects, which effectively controls for all stable individual effects. This procedure has been implemented in Limdep 8.0. However, estimates of $\beta$ are inconsistent in negative binomial models when using such a dummy variable approach in short panels due to the incidental parameters problem (Cameron & Trivedi, 1998: 282). The number of unit-specific (e.g., firm) parameters ($\alpha_i$) increases with the sample size, while the number of periods ($T$) stays fixed, resulting in a limited number of observations to estimate a large number of parameters. In our data, we would need to estimate 1105 firm-specific parameters using 6 periods of observations per firm. Contrary to linear regression models, the maximum likelihood estimates for $\alpha_i$ and $\beta$ are not independent for negative binomial models since the inconsistency of the estimates of $\alpha_i$ are transmitted into the MLE of $\beta$. Thus, we chose not to employ Allison and Waterman’s (2002) unconditional estimator. Furthermore, given that this method is a true fixed effects specification it does not allow for time-invariant covariates. Considering the importance of controlling for unobserved, time-invariant industry effects in our models as well as our use of a time-invariant covariate (Pre-sample Patents), we were unable to implement the unconditional negative binomial fixed effects specification and report the results using Hausman et al.’s (1984) conditional fixed effects approach. We point out that the results we obtained from both fixed and random effects specifications are highly consistent (see the Results section). Studies that have employed both Hausman et al.’s (1984) negative binomial fixed effects approach and that of Allison and Waterman (2002) have found very similar results (Dee, Grabowski, and Morrisey, 2005; Furman & Stern, 2004; Gordon et al., 2004). Finally, although we had evidence of significant overdispersion, we analyzed the data using a Poisson fixed effects estimation procedure (Hausman et al., 1984). This approach controls for all unobserved time-invariant sources of heterogeneity. In this analysis we excluded all time-invariant variables and obtained qualitatively similar results to those presented in Table 3. We are thankful to William Greene for his insight and advice on this matter.

\textsuperscript{13} In the Hausman et al. (1984) random effects negative binomial model, the firm specific effect is assumed to follow a gamma distribution and is described by two parameters from a beta distribution ($a$ and $b$). These parameters are estimated from the observed data. Limdep 8.0 provides estimates of these parameters and their significance levels, which we report in our results.
As an additional control for firm-level unobserved heterogeneity, we adopt the pre-sample information approach of Blundell et al. (1995). In this approach, unobserved heterogeneity is directly measured and entered into the model as a covariate. Blundell et al. (1995) argued that because the main source of unobserved heterogeneity in models of innovation lies in the different knowledge stocks with which firms enter a sample, a variable that approximates the build-up of firm knowledge at the time of entering the sample is a particularly good control for unobserved heterogeneity. Blundell et al. (1995) suggested that the pre-sample history of the dependent innovation variable is an appropriate proxy variable for a firm’s knowledge stock upon entry into the sample. The *Pre-sample Patents* variable described above serves as a ‘fixed effect’ control for unobserved differences in knowledge stocks between sample firms. Ahuja and Katila (2001) used a similar approach to control for unobserved heterogeneity in firm patenting. Finally, we include industry dummies in our models to control for unobserved industry effects that are not captured by the firm effects.

A final estimation issue concerns the appropriate lag structure of the independent variables. Based on prior research that investigates the relationship between interfirm alliances and innovation (e.g., Ahuja 2000; Sampson 2004; Stuart 2000), we employ alternative lags of our independent variables relative to our dependent variable. Specifically, we estimate three models: the first using a one-year lag, the second using a two-year lag and the third using a three-year lag. We do so to explore the robustness of our findings across alternative specifications. All models were estimated with Limdep 8.0. The model we estimate takes the general form provided below (Aerospace is the omitted industry and 1992 is the omitted year). Variables are indexed across firms (i), industry (j), and time (t):

\[
\text{Patents}_{it+1(2,3)} = f(\text{Clustering}_{jt}, \text{Reach}_{jt}, \text{Clustering} \ast \text{Reach}_{jt}, \text{R} \& \text{DAlliance}\%_{jt}, \text{R} \& \text{DIntensity}_{jt}, \text{Centrality}_{it}, \text{Local Efficiency}_{it}, \text{Centralization}_{jt}, \text{Density}_{jt}, \text{Pre-sample Patents}_{it}, \text{Automotive}, \text{Chemicals}, \text{Computers}, \text{Audiovisual}, \text{Medical}, \text{Petroleum}, \text{Pharmaceuticals}, \text{Semiconductors}, \text{Telecommunications}, \text{Measuring}, 1993, 1994, 1995, 1996, 1997).
\]
RESULTS

A summary of the network statistics and patent counts for each industry is provided in Table 1. As shown, there is substantial variation across industries in the number of firms that participate in alliances. This is largely due to differences in industry size. The average number of alliances per firm within each industry exhibits much less variation. The next column provides the average number of firms in each network. This number includes firms from the industry and their partners, some of which are not in the target industry. The next column indicates what percentage of the nodes in the network that are connected to the single largest (“main”) component. This number varies significantly both across industry and over time (not shown). While researchers often study only the single largest (“main”) component in many network studies, in our study this would have yielded misleading results. Whereas in some industries there is a large main component that is relatively stable over time (e.g., pharmaceuticals), in other industries there are multiple large components, and those components merge and split apart over time. For example, between 1996 and 1997 in the computer industry, a large component broke away from the main component (see Figure 3). If we had focused only on the single largest component, we would have both understated the amount of alliance activity in the industries, and overstated the amount of change in alliance activity over time.

The next set of columns refers to the clustering coefficients of the alliance networks. First, the actual clustering coefficient of the networks (averaged across time) is provided, followed by the clustering coefficient that would be expected of a random graph of similar size and degree (calculated as a ratio of degree over number of firms, k/n), and the ratio of these two coefficients. Notably, each of the industry networks demonstrates significantly more clustering than would be expected in a random graph of the same size and degree. The chemicals, computers and office equipment, and pharmaceutical industries demonstrate particularly high degrees of clustering. In the next set of columns, the actual path length of each network (averaged across time) is provided, with the upper limit of the expected average
path length, the *diameter*, of a random graph of the same size and degree (calculated as log n/log k).\textsuperscript{14} The following column provides the ratio of the actual path length to the random graph path length.

Comparing the clustering ratio to the path length ratio reveals that for some of the industries, the clustering coefficient is much greater than that of a random graph, but the path length is remarkably close to that of a random graph. The industries where the clustering coefficients are very high and the path length is close to that of a random graph (e.g., pharmaceuticals, measuring equipment) appear to be small worlds. Not all of the industries in our sample, however, are small worlds. Without such variation we would likely not be able to statistically detect an influence of *Clustering X Reach* on firm patenting. It is important to note, however, that the statistics provided in Table 1 are averaged across all of the time periods and are only used to provide some illustrative data about the industries. The averages mask considerable variance over time for each of the industries. We take advantage of both the cross-sectional and longitudinal variance in these network measures to test our hypothesis about the impact of small-world properties using a panel model (presented in Table 3).

-------------------------------------------------Insert Table 1 About Here---------------------------------------------------

In the remainder of the analyses we emphasize average distance-weighted reach (or simply *reach*) rather than average path length because it captures both path length and the number of firms that can be reached by any path (which accounts for differences in network size). This is a more meaningful measure in our context than simple path length because a firm that can reach a large number of other firms via a short average path has greater access to a wider range of information than a firm that can reach fewer firms via a similarly short average path. Given the variation in the number of nodes in the networks and in the average path lengths, it is not surprising that there is substantial variation in the average reach of the networks. The column indicates that firms in some industries can reach many others via a short path

\textsuperscript{14} Because the networks are disconnected, we used the harmonic path mean technique to calculate average path lengths for Table 1. This method resolves the infinite path length problem by exploiting the fact that the inverse of infinity is zero. The distance between every pair of nodes is inverted, and then averaged across every pair, and then this average is inverted. The resulting number gives you a meaningful measure of the overall connectivity of the network.
length, while others can reach relatively few. The table also indicates that there is substantial variation in average firm patenting across industries.

Table 2 provides the descriptive statistics and correlations for the variables. Table 3 reports the negative binomial panel regression results for the three dependent variables (\(\text{Patents}_{it+1}; \text{Patents}_{it+2}; \text{Patents}_{it+3}\)). Because the random effects specification assumes that regressors and firm-specific effects are uncorrelated, we also provide results using firm fixed effects as a robustness check. Separate results are provided for three dependent variables. Models 1, 2 and 3 report the results using a one-year lag between the independent variables and firm patenting (\(\text{Patents}_{it+1}\)). Models 4, 5 and 6 report the results using a two-year lag (\(\text{Patents}_{it+2}\)) and models 7, 8 and 9 report the results using a three-year lag (\(\text{Patents}_{it+3}\)). For each dependent variable, the first models (1, 4 & 7) include the constant and control variables only, the second models add the direct effects of Clustering and Reach (models 2, 5 & 8), and the third model adds the interaction term, Reach X Clustering (models 3, 6 & 9). Firm, industry and time period effects, while estimated, are not reported to conserve space.

Our sole hypothesis predicted a positive effect of the interaction of Clustering and Reach on subsequent firm patenting. The interaction term, Clustering X Reach, does not obtain statistical significance at conventional levels in the model specified with a one-year lag, using either fixed or random firm effects (Model 3). However, the coefficient for Clustering X Reach is positive and statistically significant in models using both two- and three-year lags (Models 6 & 9). Furthermore, this result holds for models using both fixed and random firm effects. Thus, our hypothesis received strong support in models using two- and three-year lags.

In order to better understand the meaning of the interaction effect, the nature of the coefficients for Clustering and Reach in models 6 and 9 in Table 3 must be understood. The estimated coefficients for Clustering and Reach in these models are simple effects rather than true main effects due to the significance of the interaction term (Jaccard & Turrisi, 2003). Consequently, the effect of each on Patents...
is conditioned on the other variable taking on the value of 0. For example, the coefficient estimate of -0.022 for Reach in model 6 (Random Effects) assumes that the value of Clustering is equal to 0 (thus removing the impact of the interaction with Reach). Thus, the negative sign on the coefficient for Reach cannot be interpreted as a negative (main) effect of Reach on Patents. While the effect of Reach is indeed negative when Clustering is 0, the effect becomes positive when values of Clustering exceed 0.267 (note that the range of Clustering in the data is 0.0 – 0.8). Similarly, the effect of Clustering is negative (although not statistically significant) when Reach is equal to 0, but becomes positive for values of Reach greater than 1.224 (note that the range of Reach is 1.88 – 61.18). The fact that the effects of both Clustering and Reach become positive when the other obtains a relatively small value and increase in their positive effects with increases in the other provides further support for our hypothesis. These mutually reinforcing effects are consistent with the symmetrical nature of multiplicative interaction effects (Jaccard & Turrisi, 2003).

Plots of the effect of the interaction on predicted values of \( \text{Patents}_{t+2} \) and \( \text{Patents}_{t+3} \) reinforce this interpretation. For ease of presentation and interpretation, we used the log-linear form of the negative binomial models in Table 3 (i.e., where the log of the conditional mean function is linear in the estimated parameters) to calculate these effects. Figure 2 presents the interaction plot of Clustering and Reach to illustrate the magnitude of the interaction effect. The “Low Clustering” line shows the slope of the effect of Reach on Patents when the value of Clustering is set to one standard deviation below its mean (.01). The end points of the line are calculated at one standard deviation below and above the mean of Reach. The “High Clustering” line represents the effect of Reach on Patents when the value of Clustering is set to one standard deviation above its mean (.34). The end points of the line are calculated at ±1 standard deviation from the mean of Reach. Of particular interest is the difference in height between the two lines.

\[ b_1 + b_3Z \]

In order to calculate these effects, we used the log-linear form of the negative binomial models in Table 3 (i.e., where the log of the conditional mean function is linear in the estimated parameters). We then took the first derivative of the linear equation with respect to Clustering and Reach, and algebraically solved for the effect of each variable (Clustering and Reach). As Jaccard and Turrisi (2003: 23) show, the equation for calculating the slope of the predicted effects of X (e.g., Clustering) on Y at any particular value of Z (e.g., Reach) in a linear model is: \( b_1 + b_3Z \) (where \( b_1 \) represents the estimated coefficient for X and \( b_3 \) represents the estimated coefficient for the interaction effect).
If the two lines were equidistant throughout, no interaction effect would exist. However, as the plot shows the “High Clustering” line has a different slope than that of the “Low Clustering” line. Thus, consistent with the results in models 6 and 9 of Table 3, increases in Reach increase the positive effect of Clustering on Patents. The symmetrical case of plotting low and high Reach lines for low and high values of Clustering (not shown) provides similar results.

In order to assess the magnitude of the small world effect we employed the estimated marginal effects ($e^\hat{\beta}$). We calculated the difference between the value of Patents holding Clustering at its mean and Reach at its mean and the value of Patents with Clustering at its mean and Reach at one standard deviation above its mean. We then calculated the difference in Patents when Clustering was high (mean + 1 s.d.) and Reach was at its mean, and when Clustering was high and Reach was high (mean + 1 s.d.). We then calculated the absolute value in the difference in these two numbers to obtain the magnitude of the interaction effect on Patents for a single standard deviation increase in both component variables. For the model specified with a two-year lag and employing firm fixed effects, this yielded an increase of 0.98 patents (for the average firm), or 2.3%. The magnitude of the interaction effect when both component variables increase one standard deviation above their means for the model employing a two-year lag and random effects is 1.00 patents (2.3%). The magnitude of the small world effect is much smaller in the models using a three-year lag: an increase of 0.01 patents for the average firm when the model employs firm fixed effects and 0.07 when the model is specified with random effects. While the effect of small world connectivity on firm patenting was positive and statistically significant, the size of this effect in absolute terms is fairly small in our data and appears to realize its peak within two years. Based on these results, we speculate that the effect of network structure as a medium of knowledge diffusion decays over time. While a particular structure may persist over time, the knowledge that diffuses through it has limited benefit as actors absorb and apply these knowledge flows to productive ends.

The results related to the control variables also merit discussion. The effect of betweenness centrality on subsequent firm patenting failed to achieve statistical significance in any of the estimated models. In contrast, efficiency had a significant negative effect on firm patenting in all models. This result...
suggests that the presence of structural holes in a firm’s ego network of alliance relationships has a deleterious effect on its inventive output. This finding is consistent with results obtained by Ahuja (2000) and Soh (2003). To our knowledge, our study represents the largest panel data investigation of this relationship. More importantly, we find a significant small world effect on firm patenting even after controlling for a firm’s network position and the structure of its ego network of alliances.

Among the other variables in the models, most were not consistent in terms of sign and significance. This might be due, in part, to the moderate-to-large correlations among the network measures (i.e., Centralization, Density, Reach, Clustering, and Clustering X Reach). This multicollinearity might influence the robustness of our main finding because parameter estimates are unstable to very small changes in the data when substantial collinearity is present, sometimes resulting in the signs on estimated coefficients to flip (known as the “wrong sign” problem) (Gujarati, 1995). In order to examine the influence of multicollinearity on our main result, we reran each of the models in Table 3 with Centralization removed and, alternatively, with Density removed (not reported here). The results for Reach, Clustering and Clustering X Reach remained substantively unchanged across all models. Thus, our primary result does not appear to be influenced by multicollinearity.

Finally, the Pre-sample Patents variable was positive and significant in all models, indicating its importance as a control for firm-level unobserved heterogeneity. Furthermore, several time period and industry dummies (not reported) were consistently significant in all models, which indicates it was important to control for unobserved time period and industry effects.

Robustness of Results

One concern regarding our results is that we were not able to control for differences in firm-level R&D since nearly 42% of our sample firms were privately-owned during some portion of the sample. While stocks and flows of R&D have been used to proxy for a firm’s research capability and absorptive capacity (Cohen & Levinthal, 1990; Helfat, 1997), patent stocks have also been used to capture a firm’s ability to learn from external sources of knowledge (Henderson & Cockburn, 1994; Penner-Hahn &

16 We thank one of the anonymous reviewers for recommending this approach to us.
Shaver, 2005). This is consistent with characterizing a firm’s stock of patents as representing its technical competencies (Patel & Pavitt, 1997). Prior research has found that patent stock measures and annual R&D expenditures are highly correlated (Penner-Hahn & Shaver, 2005; Phelps, 2003; Trajtenberg, 1990). To measure a firm’s patent stock \((\text{patstock})\), we used the total number of patents obtained by firm \(i\) in the 4 years prior to and including year \(t\). Due to the extremely high correlation between this variable and \textit{pre-sample patents} \((r = 0.937)\), we re-estimated all of our models using the time-varying patent stock variable in place of \textit{pre-sample patents}. As might be expected (due to the extremely high correlation between the two variables) our results (not reported) did not substantively change from those reported in Table 3.

For our second robustness check we analyzed the data using a Poisson fixed effects estimation procedure (although we had evidence of significant overdispersion). We did so to address the concern identified in footnote 12 above. This approach controls for all unobserved time-invariant sources of heterogeneity. In this analysis we excluded all time-invariant variables and obtained qualitatively similar results (not reported) to those presented in Table 3.

A third concern regarding our results is that they may be influenced by the presence of persistent serial correlation in the residuals. This could result from temporally stable unobserved firm effects (Greene, 1997). Serial correlation may also result from reverse causality running from firm invention to global industry-level network structure (e.g., clustering or reach), manifesting in the lagged network variables. We explicitly address the first potential source of serial correlation by including firm fixed effects, thus controlling for unobserved, temporally persistent firm effects. Prior research on firm patenting has found that the use of a firm fixed effect virtually eliminates persistent serial correlation (Blundell, Griffith, & Van Reenen, 1995). The results from our negative binomial regressions and Poisson regressions with fixed effects are consistent with those obtained using (negative binomial) random effects.

Unreliable estimates may also result from unobserved variables that vary systematically over time. In this case, serial correlation in the errors would persist even after controlling for stable firm effects. This might occur as the result of a dynamic process whereby past network structure depends upon prior firm
patenting, implying the relationship between prior network structure and current firm patenting is spurious and more accurately the result of a dynamic relationship between prior and current firm patenting. A direct relationship between firm patenting and subsequent industry-level network structure seems unlikely because global network structure tends to be generated by complex and emergent interactions among network nodes (see e.g., Robins, Pattison & Woolcock, 2005). Nevertheless, we examined this possibility in two ways. First, we regressed our measures of clustering, reach and their interaction on annual firm patent counts using a linear panel data model. We did so using contemporaneously-measured firm patents, and one-, two- and three-year lags of firms patents. We found no significant relationship between firm patents and clustering, reach or their interaction in any of these models. Next, we aggregated firm patents to the industry-level using the average annual patent count across firms in the industry. The idea here is that as industry inventiveness increases, so does the likelihood that firms in such industries will form small world alliance networks. We ran the same model specifications as those using firm patents and also found no significant relationships.

While reverse causality running from firm patenting to industry-level network structure does not seem to be a concern, a dynamic effect of prior inventiveness on subsequent inventiveness may exist. Although time invariant, we believe our entry stock variable (Pre-sample Patents) adequately controls for such a possibility because firm patenting is highly stable over time in our sample. The correlation between Pre-sample Patents and firm patents measured one year prior to the dependent variable range from 0.95 (for 1992) to 0.78 (in 1997). Furthermore, our results do not change when we use a multiyear lagged dependent variable (Patstock) and this variable is highly correlated with Pre-sample Patents. These correlations suggest that the use of a single-year lagged dependent variable (instead of the entry stock or patent stock variables) would add little additional explanatory variance and not substantially reduce the influence of serial correlation. Blundell, Griffith, and Van Reenen (1995) showed that the use of a pre-sample patent entry stock measure virtually eliminated persistent serial correlation in their panel data models.
DISCUSSION

We began by describing small world networks and discussing their implications for diffusion and search. We argued that two structural properties in particular, clustering and reach, play crucial roles in network diffusion and search. Clustering enables a globally sparse network to achieve high information transmission capacity through locally dense pockets of closely connected firms. Short path lengths to a wide range of firms increase the reach of the network by bringing the diverse information resources of more firms within relatively close range. It is typically assumed that there is a tradeoff between transmission capacity and reach: alliances that create redundant paths within a clique of partners yield transmission capacity but forfeit reach, while alliances that create nonredundant paths to new firms create reach but forfeit bandwidth. However, research in small-world networks reveals this need not be the case. Small-world networks have both high clustering and short path lengths, enabling great reach while forfeiting little information transmission capacity. We thus argued that small-world network properties would enhance the knowledge creation within an interfirm network. We tested this argument using longitudinal data on the innovative performance of a large panel of firms operating in 11 industry-level alliance networks. The results indicated strong support for our argument: the combination of clustering and reach was associated with significantly higher firm patenting. The results were stronger for models employing a two- and three-year lag versus a one-year lag, suggesting firms do not quickly realize the innovation benefits of collaboration (Stuart, 2000).

Our results are robust to a number of controls and model specifications. The use of alternative lag specifications reduces concerns of reverse causality. The inclusion of year fixed effects controls for the influence of unobserved factors on firm patenting that vary over time, but are invariant across firms, such as changes in technological opportunity or macroeconomic conditions. Time varying industry-network controls help alleviate concerns that alternative network structures are confounding our observed results. The industry R&D intensity variable allows us to control for the influence on firm patenting of inter-industry differences in the intensity of investments for innovation. The use of industry fixed effects helps rule out the influence of unobserved time invariant industry effects. We also control for the influence of
two well-known, time varying and firm-specific network predictors of firm innovation (i.e., centrality and efficiency). Using the pre-sample (and time-varying) stock of firm patents and firm fixed and random effects allow us to control for firm-level unobserved heterogeneity. Finally, in unreported analyses, we obtain similar results using a Poisson fixed effects specification.

Our results support much of the line of argument developed in Uzzi and Spiro (2005), though our findings are not entirely consistent with their results. Our theory and evidence accord with Uzzi and Spiro’s argument that the cohesion and connectivity of a small-world network enable the circulation of creative material that can be recombined into new creative products. Furthermore, our argument that the heterogeneity of knowledge distributed across clusters enhances innovation is similar to Uzzi and Spiro’s argument that the different conventions and styles used in different clusters is a valuable source of diversity in the network.

Their data and analysis, however, are different from ours in some important ways. First, as they point out (Uzzi & Spiro, 2005: 470, footnote 8), in a mature small-world network such as theirs, the path length changes little over time, behaving like a fixed effect with a constant value near one. This means that much of the variation in the structure of their network comes from clustering and that their principle finding is driven by temporal variation in clustering. Our networks, by contrast, exhibit significant cross-sectional and temporal variation in path length and network size, leading to great variation in our measure of reach. We thus chose not to use the small-world quotient approach emphasized in Uzzi and Spiro’s study as this compound measure would be more difficult to interpret in our context than the individual measures of reach and clustering and their interaction. Second, our networks are, on average, far less dense than their network. Their network becomes sufficiently dense and clustered that it leads to excessive cohesion and homogenization of material, and a decline in creative performance. In essence, such a globally dense network has the advantages and disadvantages we argued would exist within each cluster. To investigate this effect in our data we re-estimated each of our models, replacing our interaction term with the quadratic version of clustering (i.e., clustering$^2$). This variable was not statistically significant in any model, thus we have no evidence of a parabolic effect of clustering in our data. We
speculate that our networks never reach a sufficiently high level of density, and thus are at less risk of excessive cohesion. Finally, and perhaps most importantly, Uzzi and Spiro’s network is composed of individuals whereas our networks are composed of firms. Some of the dynamics that lead to deleterious effects of cohesion (for example, strong feelings of obligation and camaraderie between friends leading to an “assistance club” for ineffectual members of the network) are far more likely in the relationships between individuals than the alliances among firms.

The results of this study speak to the economics literature on knowledge spillovers. Knowledge spillovers represent an externality in which the knowledge produced by one firm can be appropriated, at little cost, by other firms (Jaffe 1986). Spillovers are made possible by the public good nature of knowledge, which prevents it from being completely appropriated by the inventing firm (Arrow 1962). This reasoning suggests that the R&D efforts of a collection of firms serve as a pool of external knowledge for a focal firm, which may allow it to innovate at much less cost than otherwise possible (Griliches 1979). Thus, knowledge spillovers may serve as a positive externality that enhances aggregate firm innovation (Jaffe 1989). Empirical evidence indicates that spillovers are important in explaining innovation and productivity growth (Griliches 1992). However, spillovers are not equally accessible to or appropriable by all firms. Prior research has shown that spillovers tend to be spatially bounded: their effect is more pronounced for firms conducting research in similar technological domains (Jaffe 1986; 1989) and geographic locations (Feldman 1999). Our results add to this literature as they suggest that interfirm networks are an important mechanism of knowledge spillovers. Our findings indicate that collaborative interfirm relationships can act as channels for knowledge spillovers and the specific pattern that these relationships exhibit can have important consequences for the innovativeness of networked firms. Specifically, interfirm network structures that combine both a high degree of clustering and short average path lengths to a wide range of firms seem to enhance the spillover process and yield higher firm innovation rates.

This research has a number of additional contributions. First, whereas previous alliance network research has examined the impact of a firm’s network position or the structure of its immediate network
neighborhood on firm innovation, our study is the first that we know of to examine the influence of the structure of industry-level alliance networks on firm innovation. Next, while a few recent studies have shown that some individual industries exhibit small-world connectivity properties (Baum, et al. 2003; Kogut & Walker 2001), to our knowledge this is the first study that has measured the structural properties of multiple industry networks over time. More importantly, we link these network properties with knowledge diffusion and search, and demonstrate that small-world connectivity properties in interfirm networks can enhance firm knowledge creation. This result is consistent with the simulation results of Cowan and Jonard (2003), but departs somewhat from Uzzi and Spiro’s (2004) finding of an inverted U-shaped effect of the small world network structure of Broadway musical artists on the financial and artistic success of Broadway musicals. Finally, this research informs the debate over whether innovation is enhanced by network density or efficiency (see Ahuja 2000): both local density and global efficiency can exist simultaneously, and it is this combination that enhances innovation.

These findings have important implications for how interfirm networks might be deliberately structured to enhance knowledge creation. While no single actor may be able to control an entire industry network, there are some that can influence the structure of interfirm networks to take advantage of small world properties. Government agencies involved in industrial policy can exert influence over the collaborative activities of industry members to foster greater innovation and competitiveness. For example, the European Union's EUREKA R&D program plays a large role in organizing the collaborative R&D activities among European companies. MITI performs a similar function in Japan. Research consortia also play a powerful role in structuring relationships among consortia members. All of these organizations are in a position to actively influence the structure of their respective collaborative networks. This suggests that these organizations could benefit from analyzing the structure of interfirm networks from a global network perspective in order to enhance the innovation of participating firms.

A limitation of our theoretical focus is that we ignore the influence of network characteristics other than structure. We do not address the properties of the alliances themselves (e.g., strength, type, governance structure, scope). Different types of relationships may be better or worse for searching for vs.
transferring knowledge (Hansen 1999). In addition, different types of relationships will be more or less costly to maintain and thus affect the efficiency of network structure for knowledge creation. We do not examine how the attributes of the firms shape the flow of knowledge (see Owen-Smith & Powell 2004). We have also not addressed the potential impact of the nature of knowledge that is being accessed, transferred and recombined in the network. Different characteristics of knowledge (e.g., tacit versus explicit, complex versus simple, etc.) can influence the knowledge creation and innovation process (Zander & Kogut 1995). Network structure may also differentially interact with different dimensions of knowledge. For example, the high density of clusters may facilitate the search and transfer of tacit, complex knowledge, but the relatively few connections to other clusters may make such search and transfer problematic. These aspects of relationships and knowledge will likely be important in fully understanding the relationship between interfirm knowledge networks and knowledge creation, but are beyond the scope of our paper. Future theoretical and empirical research should incorporate these lines of inquiry with our emphasis on network structure.

Another limitation of our work is that the generalizability of our theory and main result is likely limited to industries that make frequent use of alliances. While small world structures can exist in sparse networks, networks characterized by extreme sparsity in which only a handful of all possible links exist may not have a sufficient degree of connectedness to observe clustering or calculate path lengths. However, the insights of our theory are not necessarily limited to alliance relationships. Because firms are connected via other relationships, the global structure of such relationships may influence firm innovativeness. For example, firms are often connected by interpersonal collaborative relationships among individual inventors. The extent to which the global structure of these relationships is characterized by small world properties may have implications for the inventiveness of individual inventors and their firms (see Fleming, King & Juda, 2004). Furthermore, because knowledge can flow between firms through other mechanisms such as individual mobility, geographic clustering, participation in technical committees, or learning from information made public through patenting, it is possible that some of the knowledge creation advantages of a particular alliance network structure might spillover to
other industry (or non-industry) participants. This is an exciting area for future research.

An additional opportunity for future research pertains to the possibility of a reciprocal relationship between network structure and firm-level innovation. That is, does firm-level patenting influence industry-level structure? A few studies have investigated the influence of firm inventiveness on subsequent alliance formation (e.g., Ahuja, 2000; Shan, Walker & Kogut, 1994). While it would be theoretically and statistically challenging to assess when and how firm-level innovation influences global network structure, research in sociology has begun to explore such micro-to-macro processes using simulations (see, for example, Robins, Pattison & Woolcock, 2005).
References


Borgatti, S.P. Identifying sets of key players in a network. *Computational, Mathematical and Organizational Theory*. Forthcoming


Figure 1: Interval within which High Clustering and Short Path Lengths Coexist

Figure 2: Connective Properties of "Connected Caveman" and Random Networks

<table>
<thead>
<tr>
<th>Network Type</th>
<th>Path Length</th>
<th>Clustering Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) Connected Caveman</td>
<td>5</td>
<td>.75</td>
</tr>
<tr>
<td>b) Connected caveman with three randomly rewired links</td>
<td>3.28</td>
<td>.66</td>
</tr>
<tr>
<td>c) Random growth network</td>
<td>2.51</td>
<td>.21</td>
</tr>
</tbody>
</table>
Figure 3: Network Component Structure

Computers, 1996

Computers, 1997
Figure 2: Graph of Interactions for Random Effects Models, Patents\textsubscript{t+2} and Patents\textsubscript{t+3}

Log of Conditional Mean of Patents\textsubscript{t+2}

Low Reach | High Reach
---|---
0 | 1.4
0.2 | 1.2
0.4 | 1
0.6 | 0.8
0.8 | 0.6
1 | 0.4
1.2 | 0.2
1.4 | 0

Log of Conditional Mean of Patents\textsubscript{t+3}

Low Reach | High Reach
---|---
-0.3 | 0.3
-0.2 | 0.2
-0.1 | 0.1
0 | 0
0.1 | -0.1
0.2 | -0.2
0.3 | -0.3

High Clustering
Low Clustering
<table>
<thead>
<tr>
<th>Industry</th>
<th>Average Number of Firms from Industry in Alliances*</th>
<th>Average Number of Alliances per Firm</th>
<th>Average Network Size (nodes)</th>
<th>Clustering Coefficient</th>
<th>Average Path Length</th>
<th>Average Distance-Weighted Reach</th>
<th>Average Number of Patents per Firm per Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aerospace</td>
<td>9</td>
<td>3.05</td>
<td>28</td>
<td>0.4</td>
<td>3.67</td>
<td>4.83</td>
<td>134.22</td>
</tr>
<tr>
<td>Automotive</td>
<td>15.67</td>
<td>3.43</td>
<td>53.2</td>
<td>0.47</td>
<td>7.29</td>
<td>5.51</td>
<td>47.04</td>
</tr>
<tr>
<td>Chemicals</td>
<td>45.17</td>
<td>2.97</td>
<td>199.8</td>
<td>0.36</td>
<td>24.22</td>
<td>3.9</td>
<td>27.08</td>
</tr>
<tr>
<td>Computers and Office Equipment</td>
<td>79.67</td>
<td>4.48</td>
<td>347</td>
<td>0.24</td>
<td>18.59</td>
<td>20.64</td>
<td>49.88</td>
</tr>
<tr>
<td>Household Audio - Visual Equipment</td>
<td>9</td>
<td>1.5</td>
<td>28.3</td>
<td>0.13</td>
<td>2.45</td>
<td>2.04</td>
<td>2.74</td>
</tr>
<tr>
<td>Measuring and Controlling</td>
<td>22.67</td>
<td>1.96</td>
<td>48.33</td>
<td>.65</td>
<td>16.14</td>
<td>22.01</td>
<td>40.52</td>
</tr>
<tr>
<td>Medical Equipment</td>
<td>66.17</td>
<td>1.66</td>
<td>172.33</td>
<td>0.06</td>
<td>6.23</td>
<td>2.90</td>
<td>5.81</td>
</tr>
<tr>
<td>Petroleum Refining and Products</td>
<td>5.3</td>
<td>2.65</td>
<td>24.83</td>
<td>0.22</td>
<td>2.06</td>
<td>2.33</td>
<td>19.58</td>
</tr>
<tr>
<td>Pharmaceuticals</td>
<td>218.33</td>
<td>2.54</td>
<td>510</td>
<td>0.09</td>
<td>18.07</td>
<td>46.32</td>
<td>7.46</td>
</tr>
<tr>
<td>Semiconductors</td>
<td>58.67</td>
<td>3.51</td>
<td>204</td>
<td>0.13</td>
<td>7.56</td>
<td>19.65</td>
<td>39.31</td>
</tr>
<tr>
<td>Telecommunication Equipment</td>
<td>44.83</td>
<td>6.53</td>
<td>266.33</td>
<td>0.21</td>
<td>8.56</td>
<td>23.78</td>
<td>28.08</td>
</tr>
</tbody>
</table>

a. This number includes only those firms with the designated primary SICs; it does not include partners in the network that are not in those SICs.

b. Includes all U.S. firms in network, including both those with the designated primary SICs and their alters, regardless of SIC.

c. Since the networks are not fully connected, the average path length is calculated using a harmonic mean technique (see Newman, 2000).

d. This is the expected diameter of a random graph, or the length of the largest geodesic (i.e., shortest path between two nodes), which is the upper limit of the average path length of a random graph (Newman, 2000).
<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Patents_{it+1}</td>
<td>40.372</td>
<td>173.187</td>
<td>0</td>
<td>3638</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Patents_{it+2}</td>
<td>42.726</td>
<td>185.856</td>
<td>0</td>
<td>3638</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Patents_{it+3}</td>
<td>42.715</td>
<td>186.145</td>
<td>0</td>
<td>3638</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Pre-sample Patents</td>
<td>104.443</td>
<td>403.998</td>
<td>0</td>
<td>4191</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Density</td>
<td>0.010</td>
<td>0.014</td>
<td>0.004</td>
<td>0.127</td>
<td>0.068**</td>
<td>0.083**</td>
<td>0.085**</td>
<td>0.193**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Centralization</td>
<td>9.152</td>
<td>5.544</td>
<td>0.000</td>
<td>17.820</td>
<td>-0.016</td>
<td>-0.005</td>
<td>0.010</td>
<td>-0.065**</td>
<td>-0.185**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>R&amp;D Intensity</td>
<td>0.096</td>
<td>0.035</td>
<td>0.010</td>
<td>0.152</td>
<td>-0.087**</td>
<td>-0.083**</td>
<td>-0.085**</td>
<td>-0.159**</td>
<td>-0.353**</td>
<td>-0.681**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>R&amp;D Alliance %</td>
<td>0.783</td>
<td>0.129</td>
<td>0.125</td>
<td>0.937</td>
<td>-0.089**</td>
<td>-0.086**</td>
<td>-0.092**</td>
<td>-0.181**</td>
<td>-0.478**</td>
<td>-0.364**</td>
<td>-0.763**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Efficiency</td>
<td>0.949</td>
<td>0.171</td>
<td>0.000</td>
<td>1.000</td>
<td>-0.035*</td>
<td>-0.037*</td>
<td>-0.036*</td>
<td>-0.046**</td>
<td>-0.185**</td>
<td>-0.034*</td>
<td>0.060**</td>
<td>0.107**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Betweenness</td>
<td>0.514</td>
<td>1.673</td>
<td>0.000</td>
<td>18.970</td>
<td>0.420**</td>
<td>0.429**</td>
<td>0.447**</td>
<td>0.429**</td>
<td>0.127**</td>
<td>0.167*</td>
<td>0.077**</td>
<td>-0.005</td>
<td>-0.021</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Clustering</td>
<td>0.175</td>
<td>0.163</td>
<td>0.000</td>
<td>0.800</td>
<td>0.104**</td>
<td>0.093**</td>
<td>0.083**</td>
<td>0.180**</td>
<td>0.576**</td>
<td>-0.302**</td>
<td>-0.512**</td>
<td>-0.540**</td>
<td>-0.103**</td>
<td>0.006</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Reach</td>
<td>25.948</td>
<td>19.490</td>
<td>1.880</td>
<td>61.180</td>
<td>-0.086**</td>
<td>-0.081**</td>
<td>-0.077**</td>
<td>-0.132**</td>
<td>-0.413**</td>
<td>-0.747**</td>
<td>-0.794**</td>
<td>-0.593**</td>
<td>-0.051**</td>
<td>-0.076**</td>
<td>-0.447**</td>
</tr>
<tr>
<td>13</td>
<td>Clustering X Reach</td>
<td>3.129</td>
<td>1.893</td>
<td>0.000</td>
<td>7.288</td>
<td>0.020</td>
<td>0.020</td>
<td>0.016</td>
<td>-0.027</td>
<td>-0.174</td>
<td>0.675**</td>
<td>0.391**</td>
<td>0.266**</td>
<td>-0.034*</td>
<td>0.079**</td>
<td>0.098**</td>
</tr>
</tbody>
</table>

n = 3444, * p<.05, ** p<.01
Table 3: Panel Negative Binomial Regression Models with Fixed and Random Effects (N=1106; Obs=3444)

<table>
<thead>
<tr>
<th>Fixed Effects</th>
<th>Patents$_{it+1}$</th>
<th>Patents$_{it+2}$</th>
<th>Patents$_{it+3}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>1.136** (.354)</td>
<td>.582 (.359)</td>
<td>.604 (.360)</td>
</tr>
<tr>
<td>Pre-sample Patents</td>
<td>.001** (.000)</td>
<td>.001** (.000)</td>
<td>.001** (.000)</td>
</tr>
<tr>
<td>Density</td>
<td>-2.48 (1.154)</td>
<td>-6.24 (1.358)</td>
<td>-5.27 (1.468)</td>
</tr>
<tr>
<td>Centralization</td>
<td>-0.14 (0.008)</td>
<td>-0.14 (0.008)</td>
<td>-0.12 (0.008)</td>
</tr>
<tr>
<td>R&amp;D Intensity</td>
<td>2.73 (2.668)</td>
<td>2.86 (2.522)</td>
<td>2.87 (2.581)</td>
</tr>
<tr>
<td>R&amp;D Alliance %</td>
<td>-0.112 (2.75)</td>
<td>-0.23 (2.75)</td>
<td>-0.22 (2.89)</td>
</tr>
<tr>
<td>Efficiency</td>
<td>-1.99** (.068)</td>
<td>-1.89** (.072)</td>
<td>-1.90** (.073)</td>
</tr>
<tr>
<td>Betweenness</td>
<td>0.003 (.006)</td>
<td>0.003 (.005)</td>
<td>0.003 (.005)</td>
</tr>
<tr>
<td>Clustering</td>
<td>.420** (.136)</td>
<td>.507* (.235)</td>
<td>.346** (.127)</td>
</tr>
<tr>
<td>Reach</td>
<td>.01** (.003)</td>
<td>.01** (.003)</td>
<td>.01** (.003)</td>
</tr>
<tr>
<td>Clustering X Reach</td>
<td>-0.15 (.030)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Random Effects</th>
<th>Patents$_{it+2}$</th>
<th>Patents$_{it+3}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>1.984** (.307)</td>
<td>1.342** (.303)</td>
</tr>
<tr>
<td>Pre-sample Patents</td>
<td>.001** (.000)</td>
<td>.001** (.000)</td>
</tr>
<tr>
<td>Density</td>
<td>1.444 (0.900)</td>
<td>0.25 (1.092)</td>
</tr>
<tr>
<td>Centralization</td>
<td>-0.21** (.006)</td>
<td>-0.20** (.007)</td>
</tr>
<tr>
<td>R&amp;D Intensity</td>
<td>.887 (2.429)</td>
<td>1.030 (2.408)</td>
</tr>
<tr>
<td>R&amp;D Alliance %</td>
<td>.014 (.230)</td>
<td>.383 (.214)</td>
</tr>
<tr>
<td>Efficiency</td>
<td>-3.34** (.062)</td>
<td>-3.36** (.069)</td>
</tr>
<tr>
<td>Betweenness</td>
<td>0.008 (.005)</td>
<td>0.007 (.004)</td>
</tr>
<tr>
<td>Clustering</td>
<td>.554** (.106)</td>
<td>.548** (.212)</td>
</tr>
<tr>
<td>Reach</td>
<td>.007** (.003)</td>
<td>.008** (.003)</td>
</tr>
<tr>
<td>Clustering X Reach</td>
<td>-0.13** (.003)</td>
<td>-0.22** (.003)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Log Likelihood</th>
<th>Patents$_{it+2}$</th>
<th>Patents$_{it+3}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-4646.65</td>
<td>-4637.32</td>
<td>-4637.12</td>
</tr>
<tr>
<td>-4597.46</td>
<td>-4586.78</td>
<td>-4577.98</td>
</tr>
<tr>
<td>-4468.75</td>
<td>-4464.64</td>
<td>4464.46</td>
</tr>
</tbody>
</table>

Standard errors in parentheses.
*p < .05, **p < .01 (two-tailed tests for all variables)
All models include firm, time period and industry effects.
Appendix A: Measuring Reach

To better understand the reach measure, consider the figure below. There are three components, A-B-C-D-E, F-G-H-I, and J-K.

One can calculate the collective reach of this network as following:

- A can reach three nodes (including self) at path length of one = 3, two more nodes at path length of two - 2/2 = 1; thus A’s *distance weighted reach* = 4
- B can reach four nodes (including self) at path length of one = 4; one more node at path length of two - ½ = .5; thus B’s *distance weighted reach* = 4.5
- C’s *distance weighted reach* = 4
- D’s *distance weighted reach* =4.5
- E’s *distance weighted reach* = 4
- F’s *distance weighted reach* = 3.5
- G’s *distance weighted reach* = 4
- H’s *distance weighted reach* = 3.5
- I’s *distance weighted reach* = 3
- J’s *distance weighted reach* = 2
- K’s *distance weighted reach* = 2
- *Average distance weighted reach of network: 3.55*