Paper to be presented at the DRUID 2011

on

INNOVATION, STRATEGY, and STRUCTURE -
Organizations, Institutions, Systems and Regions
at
Copenhagen Business School, Denmark, June 15-17, 2011

Furtune favors the well-located firm: Absorptive capacity and the geography if inter-firm alliances

Michelle Gittelman
Rutgers Business School
Management and Global Business
michelle.gittelman@business.rutgers.edu

Abstract

This study proposes that local and distant locations provide distinct opportunity spaces for sourcing external knowledge through alliances: local spaces create opportunities for serendipitous, unplanned encounters, whereas the ?invisible colleges? of researchers working on common problems connect readily across spatial boundaries. I hypothesize that different kinds of prior knowledge will be important in exploiting these distinct opportunities for partnering. The context is the geography of alliances between small biotechnology firms and the world?s largest pharmaceutical firms. The results show that absorptive capacity exerts a strong influence on the geography of alliances, and that the heterogeneity of prior knowledge matters for the choice of partners in space: for large firms, broadscope knowledge predicts alliances with co-located firms whereas domain-specific knowledge predicts non-local alliances. The results support the hypothesis that local spaces represent opportunistic, unplanned search that generates technological variety whereas distant ties reflect problem-oriented search that exploits firms? prior R&D investments in domain-specific knowledge. There is no support for two alternative mechanisms that proximity lowers transactional hazards or facilitates collaboration.

Jelcodes:A10,-

Fortune favors the well-located firm: Absorptive capacity and the geography of inter-firm alliances

Michelle Gittelman
Rutgers Business School
Department of Management and Global Business
Newark-New Brunswick
michelle.gittelman@business.rutgers.edu

Abstract. This study proposes that local and distant locations provide distinct opportunity spaces for sourcing external knowledge through alliances: local spaces create opportunities for serendipitous, unplanned encounters, whereas the "invisible colleges" of researchers working on common problems connect readily across spatial boundaries. I hypothesize that different kinds of prior knowledge will be important in exploiting these distinct opportunities for partnering. The context is the geography of alliances between small biotechnology firms and the world's largest pharmaceutical firms. The results show that absorptive capacity exerts a strong influence on the geography of alliances, and that the heterogeneity of prior knowledge matters for the choice of partners in space: for large firms, broadscope knowledge predicts alliances with co-located firms whereas domain-specific knowledge predicts non-local alliances. The results support the hypothesis that local spaces represent opportunistic, unplanned search that generates technological variety whereas distant ties reflect problem-oriented search that exploits firms' prior R&D investments in domain-specific knowledge. There is no support for two alternative mechanisms that proximity lowers transactional hazards or facilitates collaboration. An unexpected finding is that knowledge variables operate differently for small and large firms in predicting the geography of alliances, suggesting that local spaces represent different opportunities for exploration and exploitation for these sets of firms.

"Pfizer's Research Technology Center (RTC) is located in Cambridge, MA, one of the richest scientific environments in the world, surrounded by more than 300 biotechnology companies and thousands of scientific innovators in world-leading research hospitals and academic institutions. A key element of our success is the growing, dynamic partnerships with these innovators to address the needs of patients worldwide." *Pfizer website*

"We were working on a new drug compound, competing with a team at Genentech. They located a Swiss scientist who was working on a molecule that turned out to be critical to the project, and they made a collaboration with his lab. We were told not to collaborate on the project. Genentech won the race to develop that compound." Research scientist, French biotechnology company (personal communication)

Geographic space is an important dimension patterning firms' search for external knowledge. Despite the increasing ease with which people can communicate almost costlessly across space, evidence indicates that physical space matters for firm-level learning. The concept of embeddedness suggests that it is difficult to separate knowledge from the context in which it is made, such that the locations of firms' R&D labs situate a firms' technological learning in space. The geographic terrain has been likened to the physical dimension of a firm's search in technology space, with distant locations representing novelty and the local space generating opportunities that reinforce a firms' existing knowledge and social connections (Rosenkopf and Almeida, 2003, Phene et al, 2006, Narula and Santangelo, 2008, Sorenson and Stuart, 2008).

The analogy between localized learning in technology and geographic space is problematic, however, for several reasons. Localized learning involves the search for knowledge that builds on a firms' expertise accumulated through its R&D. Indeed, a firm's absorptive capacity is a calculus for linking the rate and direction of its future learning to its past R&D investments. Yet a key reason that firms invest in formalized research is to create a search capability that is free of the cognitive constraints imposed by the embeddedness or situated properties of knowledge. A byproduct of research is a capability to search broadly in geographic space, such that the deeper a firm's engagement in R&D, the greater the relative payoff to localized *technological* search, but the less the relative advantages gained from localized *geographic* search.

This is particularly salient for science-based research. Scientific principles are explicitly geared towards producing and communicating knowledge in a form that is stripped of contextual details, so that

information can be readily accessed and absorbed by others with related expertise. Scientific research can broaden the scope across which firms search in technology space (Fleming and Sorenson, 2003; Fabrizio, 2009) and connect firm scientists to broader "invisible colleges" of researchers working in outside organizations (Cockburn and Henderson, 1998, Gittelman, 2007). The sociologist of science Diana Crane wrote that "Scientific communities have been international since their emergence in the seventeenth century. Basic science today consists of hundreds of research problem areas in which groups of scientists study similar problems and exchange information across national boundaries" (Crane, 1971, p. 585). Engaging in R&D creates a capability to learn that is forged in common interest in specific problems, rather than shared geographic contexts.

Local spaces form a different context for accessing external knowledge. The economist Jane Jacobs, writing about cities as engines of economic growth, pointed to both diversity and density as key ingredients in stimulating new knowledge combinations that spark creativity and innovation in urban spaces (Jacobs, 1969). In this account, productive spillovers – Jacobs externalities – arise from unplanned encounters between co-located firms with diverse but related specializations. Later work by Saxenien and others also stress the importance of informal, unplanned interactions between engineers in a region as stimulating rich information flows and cross-fertilization of ideas (Almeida and Kogut, 1999, Saxenien, 1994). Bringing together the ideas of Crane and Jacobs, economic geographers have distinguished between distant and local spaces as distinct modes of knowledge transmission and retrieval: firms acquire information through local "buzz" – unplanned, informal information flows in a region – as well as through global pipelines: intentional communications that link firms to distant sources of knowledge (Bathelt et al, 2004, Storper and Venables, 2004).

I develop the idea that local and distant locations provide distinct opportunity spaces for sourcing external knowledge through alliances. I propose that local partnerships arise from unplanned, serendipitous encounters, while distant partnerships are generated through the purposeful search for solutions to problems. In this perspective, co-location matters not so much because it affects *what* firms can know, as much as who they can know and how they access those partners. I develop models that test

for the effect of proximity on alliance formation, and also test for alternative mechanisms that generate the localization of inter-firm alliances, distinguishing between learning, transactional hazards, and knowledge effects. The models show that the greatest impact on localization of ties stems from prior knowledge of partners, and the results are consistent with the claim that local spaces are contexts for chance meetings, supported by their broadscope knowledge, whereas firms' R&D in specific domains link them to "invisible colleges" that are geographically dispersed.

The context is alliances in the biopharmaceutical industry. I focus on vertical alliances between two distinct groups of firms: small biotechnology firms and the worlds' largest pharmaceutical firms. Alliances between these groups of firms span important boundaries in both technology and geographic space. The industry is perhaps the best example of recent theorizing of a new "innovative division of labor" in which markets for technology and inter-firm collaborative networks have increasingly replaced internal firm hierarchies in the innovative process of discovery and subsequent commercialization (Arora and Gambardella, 1990, Arora et al 2001). The alliances I focus on join together firms that possess different knowledge sets: entrepreneurial biotechnology firms specialize in genetics-based approaches to drug discovery, while large pharmaceutical firms possess expertise of drug development and the markets associated with specific diseases. Moreover, small firms tend to be clustered in locations that are close to specialized inputs (e.g., scientific labor markets and venture capital) and exhibit strong tendencies to agglomerate in regions that are outside the traditional geographic footprint of large pharmaceutical firms. Alliances between these groups of firms thus provide a fruitful context in which to study the relationship between firm's prior knowledge and the geography of their external alliances.

Geography and inter-firm alliances

Alliances are an important means by which firms invest in uncertain technological and market environments and learn from other firms (Kogut, 1991, Mowery et al, 1996, Gomes Casseres et al, 2006). They are a particularly useful mechanism by which firms break path-dependent learning and acquire knowledge that is different from their own (Rosenkopf and Almeida, 2003). In fields where patents protect a large portion of the technology space, licenses embedded in alliances are frequently necessary to

move forward along a particular technological or market trajectory. Although alliances do not entail full organizational commitment, and may only involve access to partners' knowledge rather than learning from them (Mowery et al, 1996), they nonetheless impose significant opportunity costs in the form of commitment of capital, managerial attention and (frequently) contractual exclusivity, and reflect firms' bets on future technological developments. Alliances thus trace out firms' search for external knowledge, often to complement their own internal R&D projects.

Vertical technology alliances are a specific form of alliance, in which a firm licenses an early-stage technology to another firm that purchases rights to further development and commercialization. Vertical alliances have become increasingly important in industries where small, research-specialized firms have emerged as important sources of new discoveries. Because they lack downstream development and marketing capabilities they frequently rely on license and alliances with large firms (Lerner and Merges, 1998, Arora et al, 2001).

A recent study of patent licensing finds that firms face numerous challenges in sourcing knowledge through vertical alliances. Gambardella and colleagues (2007) show that small firms are much more likely to license out their patented technologies than large firms, but that a significant share of technologies that have the potential to be licensed do not find buyers. Their results highlight the asymmetries of these markets in terms of the size distribution of firms, and the difficulties in contracting for knowledge in such settings. Large firms face significant search costs in identifying projects among hundreds or thousands of small firms that are often privately held and who may have few outputs other than intellectual property. In addition to search costs, technology alliances can suffer from information asymmetries and a "lemons problem", particularly when sellers are small research-intensive firms that pose a risk of over-selling their innovations, and who may not have the managerial or scientific competence to successfully carry out their end of an alliance agreement (Guedj and Scharfstein, 2005, Arora et al, 2007). Partnerships between large, hierachical firms and small entrepreneurial firms frequently join together organizations with different cultures, research paradigms, and geographic locations, posing difficulties for coordination, communication and knowledge transfer.

Researchers have proposed several mechanisms by which geography can reduce these frictions. We can group these mechanisms into two broad categories. The first maintains that co-location reduces the hazards of transacting for intangibles by restraining "bad" behaviors opportunism and other forms of moral hazard - that may arise in such transactions. Proximity facilitates alliances by allowing local firms to gather "soft" information that enables more accurate ex ante evaluation of partners and projects. Peterson and Rajan (2002) find that distance between lenders and small-firm borrowers increases as codified credit information becomes more available; Malloy (2005) shows that equity analysts are able to predict stock prices more accurately as distance to firms declines, and that the effect is strongest for small firms. In a study of technology licensing, Gans et al(2007) find that co-location in Silicon Valley, where reputation effects are strong, substitutes for formal institutions in lowering the risk of opportunism by creating a credible threat of reputational loss in the region.

A second mechanism by which proximity can lower frictions in vertical alliances is by stimulating "good" behaviors associated with cooperation and learning. In this perspective, face to face contact enables for better communication, coordination, and transfer of tacit knowledge that is difficult to accomplish at a distance. Indeed, this mechanism has been identified as a key reason that biotechnology start-up firms cluster near universities and key scientists (Audretsch and Stephan, 1996; Zucker et al, 1998). Within a shared institutional, social, and cultural space, firms can develop similar organizational routines and cognitions that can facilitate inter-firm learning (Phene et al, 2006, Narula and Santangelo, 2008). Entrepreneurs have strong links to business networks, finance providers, and other sources of resources in the home region, and shared location is theorized to foster social and technological homophily, increasing trust and lowering the perceived riskiness of collaboration (Stuart and Sorenson, 2003, Sorenson and Stuart, 2008).

A third mechanism relaxes the normative stance that proximity improves information and knowledge flows, and maintains that localization of ties is driven primarily by the greater likelihood that firms will come in contact when they are co-located. In this case, the clustering of ties may arise simply because of the higher opportunities for random encounters within short geographic distances (Ellison and

Glaeser, 1997). Borgatti and Cross (2003) state that "proximity leads to chance meetings in which people gradually come to learn about each other, become comfortable with each other, and develop bonds that enable future access." That is, co-location creates opportunities for unplanned, serendipitous encounters, some of which will later sprout into formal alliances.

Viewed through this lens, proximity matters to inter-firm ties because it increases firms' exposure to a diverse set of potential partners in a region. This conceptualization squares with empirical work on the localization of knowledge flows through informal communication channels, e.g. the "after hours" socializing of engineers (Saxenien, 1994, Almeida and Kogut, 1999). Economic geographers have described local "buzz" as the unplanned circulation of updated news and technological information through open, informal social and networks in a local space (Storper and Venables, 2004, Owen-Smith and Powell, 2004). Co-located firms may be first to learn of news about promising business leads and new technological developments that bubble up in a region, positioning them to seize promising opportunities before more distant firms even know of them. Economic geographers emphasize diversity in local spaces as stimulating innovation, through the cross-fertilization of ideas across firms in different sectors (Jacobs, 1969; Storper and Venables, 2004). Localized information flows are likely to be particularly important for learning about opportunities to partner with small, research-intensive firms whose reputation may not extend beyond the region.

Two recent studies by Agrawal and colleagues (Agrawal et al, 2006, Agrawal et al, 2008) provide evidence that regions provide spaces for serendipitous encounters, connecting firms that would not otherwise know of eachother. In an analysis of patent citations, they find that spatial proximity is far more important in connecting individuals if they are not already socially connected, and co-location is particularly salient for bringing together firms across fields rather than within them. Distant contacts, in contrast, are more likely to have their origin in membership in common technological communities. However, when the origin of distant contacts was prior co-location, these are more likely to be cross-field connections. These findings suggest that local spaces substitute for, rather than complement, shared

technological knowledge in connecting firms to one another, and that a firms' distant ties are likely to be in related fields.

We can now summarize the mechanisms by which proximity can facilitate alliances between firms:

H1: Co-location reduces transactional hazards of contracting for intangibles

H2: Co-location facilitates inter-firm learning, coordination, and tacit knowledge transfer

H3: Co-locations allows for chance encounters with other firms in a region

In testing for these effects it is important to control for the regions in which alliances occur. Places differ in terms of the number, density and attractiveness of potential partners, as well as the institutional environment supporting inter-firm collaboration. Some regions may have a particularly high bandwidth of localized information flows and dense local networks of information flow; if the economic space is attractive enough, firms seek to enter such regions to gain first-mover advantages in learning about promising deals before others know about them. Thus co-location and the attractiveness of a region for knowledge sourcing and learning are endogenous, such that it is important to empirically disentangle effects that stem from location in specific regions from those that arise from the benefits of proximity across regions.

Absorptive capacity and the geography of alliances

Firms' ability to identify and exploit external knowledge is strongly conditioned by their prior knowledge, such that firms have a tendency to search locally with respect to their past R&D investments (Cohen and Levinthal, 1994). At the same time, firm knowledge is heterogeneous, and different kinds of expertise can be leveraged for different learning tasks (Zahra and George, 2002). Investments in R&D can generate knowledge of specific technologies or markets that enable a firm to advance within a pre-existing technological or market trajectory. It can also create fundamental knowledge of principles that allow a firm to learn across different technologies and markets. I refer to these as domain-specific and broadscope knowledge, respectively.

Domain-specific knowledge allows for increasing returns to scale from search in a given technological trajectory, and can be leveraged to identify external partners working in a specific technology or market niche. Firms accumulate expertise in specific market niches, e.g. cardiovascular disease or diabetes, and this knowledge is specialized in terms of product markets and end users. Specialized expertise lowers search costs and raises the effectiveness with which firms can evaluation and absorb new information that builds on prior knowledge. In addition, working in a particular area or research domain can create a scanning capability – "know-who" – that is valuable in identifying others working on similar problems. Researchers working in product or technology-specific domains become acquainted with, and frequently interact with, the "invisible colleges" of researchers working in a similar area. These larger communities open conduits of information about new development in a field, and allow firms to solve specific problems related to their own R&D efforts (Cockburn and Henderson, 1998, Gittelman, 2007). Product market competition also raises awareness of "who is working on what" in a particular field. Domain specific knowledge should thus be comparatively useful for "problem-oriented" search: locating external technologies that address specific needs of the firm that arise in the course of its R&D.

Broadscope knowledge enables firms to search and absorb external knowledge across multiple markets and technologies, and is thus particularly useful in varied and heterogeneous learning environments where the type of knowledge that the firm receives cannot be planned in advance. Cohen and Levinthal (1990) write: "When information flows are somewhat random and it is not clear where in the firm or subunit a piece of outside knowledge is best applied. . .it is best for the organization to expose a fairly broad range of prospective "receptors" to the environment". Shipilov (2009) finds that firms with broad prior knowledge are better able to absorb heterogeneous information, granting them advantages from brokerage positions in open networks. Broadscope knowledge has been identified as important in firm's ability to adapt to radical technological change (Tripsas, 1997), keep pace with the evolution of technologies as they branch into new directions (Kogut and Kim, 1996); learn from alliance partners

(Lane and Lubatkin, 1998), acquire complementary knowledge through joint ventures (Shenkar and Li, 1999) and search broadly in technology space (Fleming and Sorenson, 2004, Fabrizio, 2009).

If local and distant spaces represent different kinds of environments for information flow and search, the heterogeneity of knowledge should influence the geographic distribution of firms' research alliances. Specifically, Broadscope knowledge should be more useful in partnering in the local environment, since information about possible partnering opportunities is hypothesized to arise in an unplanned, informal manner, and be relatively heterogeneous with respect to the firm's own R&D activities at any point in time. Domain-specific knowledge will be relatively less useful in local spaces, as the set of local partners working on a specific problems at a particular point in time is likely to be small relative to the full set of possible partners. Domain-specific knowledge should be comparatively useful for intentional, problem-oriented search for partners who are more likely to be located outside the region.

The discussion leads to two testable hypotheses that are corollaries of H3:

H3a: Firm's accumulated broadscope knowledge will be associated with local partnerships

H3b: Firms accumulated domain-specific knowledge will be associated with non-local partnerships

Industry Context: Vertical alliances in the pharmaceutical industry

The biopharmaceutical industry is perhaps the best example of recent theorizing of a new "innovative division of labor" in which markets for technology and inter-firm collaborative networks have replaced internal firm hierarchies in the innovative process of discovery and subsequent commercialization (Arora, et. al. 2001). Biotechnologies are process technologies based on molecular biology and genetics; their utility is to create more efficient and precise means of finding drugs, rather than targeting specific disease markets. The new technologies were pioneered by small firms that were frequently spun out of university research, and typically did produce tangible products other than patented technologies. Large pharmaceutical firms, in contrast, had expertise in traditional discovery technologies and specific disease markets; they also possessed downstream development and marketing capabilities. The two sets of firms thus possess distinct but complementary capabilities to innovate; over time,

however, many large pharmaceutical firms have accumulated in-house capabilities in the new biotechnology discovery techniques as well.

Alliances between these sets of firms are an important mechanism by which the discoveries of small firms are developed for a market: the 324 biopharmaceutical drugs in development in 2004, 72% were developed by small, specialized biotechnology firms and 16% by large pharmaceutical firms; however, pharmaceutical firms accounted for 45% of 108 already-approved biotechnology drugs, suggesting that a large proportion of biopharmaceutical drugs are discovered by small biotechnology firms and developed and/or subsequently marketed by large pharmaceutical firms.[1]

Alliances between these sets of firms must bridge not only significant organizational and technological boundaries, but geographic boundaries as well. Small firms tend to be agglomerated nearby universities, to allow for close interactions between firms and academic researchers working on joint projects (Audretsch and Stephan, 1996, Zucker et al, 1998, Gittelman, 2007). Thus, the dis-aggregation of the value chain means that the geographic configuration of innovative activity is shifting away from traditional locations (e.g., New Jersey) that reflected the early roots of the industry in chemistry, and towards new locations such as California and the Boston/Cambridge region that have strong basic research in the life sciences. In response to these shifts, many large firms have opened R&D labs in these new locations. Therefore, the geographic distribution of firms' alliances at any point in time reflect the set of local and non-local partnering opportunities that are generated by their historical R&D locations as well as more recent investments in locations to develop and access new sources of knowledge.

To explore the mechanisms that might influence the geography of inter-firm ties, informal interviews were conducted with executives in charge of alliances and licensing at several large pharmaceutical firms. Managers were asked about the processes by which they search for and identify new partnering and licensing opportunities. With growing pressures to access outside expertise in research and compound discovery, large pharmaceutical firms are increasingly investing in formalized

[1] Figures from authors' analysis of data published by Phrma, Medicines in Development: Biotechnology. 2004 Survey.

alliance processes. The search for an alliance partner is frequently triggered by a specific requirement of researchers working in a specific problem area, for instance the need to round out a portfolio of drug compounds in a disease category or to access a specific technology that the firm has identified as important to a particular research track.

Managers stressed that co-location was not a primary consideration in partner selection, for several reasons. With the increasingly distributed nature of innovation and discovery, large pharmaceutical firms have made significant investment in global scanning and screening capabilities: "We want the best, wherever it is located" is a common theme, and large firms actively "shop the globe" for new opportunities. Scientific connections are an important means by which firms identify the need for a partner, but formal search, due diligence, and negotiations are typically carried out by dedicated licensing and related personnel with legal and financial expertise. Alliances frequently involve access to rights to a specific technology, rather than collaboration or learning, reducing the need for co-location. In cases where teams do collaborate, advanced communication technology, e.g. video-conferencing, is a useful substitute for face-to-face meetings. The institutional environment also reduces the need for face to face interaction and "soft" information gathering: patents and published research codify much valuable information about the feasibility of projects, and significant resources are expended to gather analyze the relevant data in the pre-deal due diligence stage.

Empirical analysis of co-location

I first describe data and measures, followed by the empirical analysis of co-location. The analysis is staged in four parts. The first part compares the univariate distribution of distances between allied and non-allied firms, in the aggregate and disaggregated by major regions. It is instructive to visually inspect these distributions: if localization is important to alliances, allied firms should show greater clustering in space than non-allied firms. The univariate analysis by region graphically underscores the heterogeneity across regions in the localization of ties. I then estimate two sets of regression models. The first are conditional logit models to estimate whether co-location between two firms increases the probability of an alliance between them. I then estimate logit models for the set of all

alliances, in which the dependent variable indicates whether two firms are co-located, conditional on variables that capture the hypothesized effects of prior knowledge, transactional hazards and learning effects. I follow with an analysis of the technological distance firms' between firms' alliances and their drug portfolios to provide additional evidence of the core hypotheses.

Data and measures

Alliances

Data on alliances were collected from Recombinant Capital, a specialized database that has covered the biotechnology industry since the late 1980s (see, e.g., Lerner and Merges, 1998). The first step was to collect data on all alliances that included a biotechnology firm and a large pharmaceutical firm[2]. This resulted in 5445 alliances spanning 1993 to 2008. Several screening criteria were then applied to arrive at the final set of alliances. Biotechnology firms that are based outside the United States are eliminated, as are biotechnology firms whose locations could not be identified from several sources. The largest pharmaceutical firms were selected based on their ranking in terms of drugs in development, and a significant presence of R&D labs in the United States[3]. After applying these screens, the sample consists of 1038 alliances involving 14 of the largest pharmaceutical firms and 380 US-based biotechnology firms. The large firms account for 45% of alliances and 28% of all drugs developed worldwide.

Sixty seven percent of the alliances list a targeted disease in their description, while the remaining one-third either don't specify a disease or list more general objectives. For instance, an alliance described as "Human hyaluronidase (rHuPH20) drug delivery" does not have a specific disease

[2] These categories are defined in the Recombinant Capital database, and screened for accuracy.

[3] The large pharmaceutical firms included are: Abbott, AstraZeneca. Bayer, Bristol-Myers Squibb, Eli Lilly, GlaxoSmithKline, Hoffmann-La Roche, Johnson & Johnson, Merck & Co, Novartis, Pfizer, Sanofi-Aventis, Schering-Plough, Wyeth. Each of these firms account for at least 1 per cent of new drug compounds developed worldwide, with the exception of Schering Plough (0.5 percent) which is included because of its major R&D presence in the US and high number of alliances in the data.
There are 624 alliances that included firms that later merged with or were acquired by one of the fourteen firms in the sample, for instance Pharmacia was acquired by Pfizer in 2002. Alliances of the acquired firms are not included prior to the merger, since the acquired firm did not qualify for the screen as one of the largest pharmaceutical firms as independent entities. Alliances and drugs of the acquired firms are grouped with the parent firm after the date of the merger.

specified, unlike the following description: "FGF-21 Type 2 Diabetes Drug Discovery Collaboration". I identify all diseases targeted in the alliances and group these into sixteen categories that can be matched to data on drugs in development, as described below.[4]

Drugs and drugs in development

Data on drugs in development are used to develop measures of firm-level knowledge accumulated through R&D. The source of the data is Pharmaprojects, a proprietary database that tracks the progress of new drug compounds. The company records, collects and tracks data on all new compounds, starting from an early (pre-clinical research) stage of development. Importantly, compounds whose development is ceased for whatever reason are not removed from the data, so drugs that failed or were withdrawn are included, along with those that are launched or in active development.[5] In total, there are 36,646 compounds in the dataset (included discontinued and ceased compounds), spanning the late 1980s through 2006.

The Pharmaprojects data give information on the pharmacological properties of drugs, as well as the technique used to discover the compound. I utilize these codes to categorize drugs into two main technological categories. A compound is coded as a *Biotechnology drug* if it falls into one of a number of categories that are distinctive to biotechnology methods of drug discovery[6]. Each drug is also coded for the disease market it targets, and diseases are grouped into the same sixteen categories that match the diseases listed in the alliances.

[4] The disease categories and number of alliances in each category are: Cancer (149); Infectious Disease (117) CNS (110); Autoimmune/Inflammatory (81); Cardiovascular (61) Endocrinological/Metabolic (31) Respiratory (30); Hematologic (28); Dermatologic (18); Gastrointestinal (18); Bone disease (16); Urinary/Gynecologic (12); Psychiatric (12); Ophthalmic (8); Allergic (6); Dental (1).

[5] The Pharmaprojects data has an advantage over other data sources, eg, patent applications, in that Pharmaprojects screens out compounds that are not considered to be viable drug candidates and thus present a more realistic picture of actual product pipelines. Sources of the data are company contacts, websites, press releases, presentations, journal references, and published scientific reports.

[6] Biotechnology drugs are identified as belonging to the categories: biological proteins (including sub-categories antibodies and recombinant proteins); Virus particles; Nucleic acids (including sub-categories viral vectors and non-viral vectors); recombinant peptides; cellular therapy; bacterial cells.

Geographic locations and distances between firms

To calculate the geographic distances between alliance partners, the locations of both the biotechnology firms and the R&D labs of the pharmaceutical firms are identified at the time of an alliance. For the large pharmaceutical firms, the headquarters locations are included, as are US-based R&D labs engaged in drug discovery research. Non-R&D facilities are not included, on the assumption that sales branches, manufacturing plants, and other facilities are unlikely to be important sites for identifying and forming alliances with research-intensive biotechnology firms.[7]

Locations are identified using the Directory of Corporate Affiliates and, where not listed in the DCA, other sources: annual reports, websites, news reports and addresses listed on published scientific papers. The year of establishment of each lab is identified from three sources: DCA, scientific papers listed in the ISI, and company 10K reports.[8] In total, the locations of 59 R&D labs are identified for the fourteen firms.

Tables 1A and 1B shows the distribution of locations of the two sets of firms. The highest concentration of R&D labs of the large pharmaceutical firms is in New Jersey, reflecting the early roots of the industry in the chemical sector. However, many have recently opened labs in California and Massachusetts, which likely reflects the motivation to tap into specialized labor markets and clusters of biotechnology firms, hospitals, and universities that have expertise in biotechnology-based methods of drug discovery. The importance of California in the locations of the small firms is much higher than for the large firms.

[7] It is plausible that non-R&D sites initiate alliances. However, our interviews with alliance executives at major pharmaceutical firms indicate that research scientists are the primary instigators of these relationships. Moreover, in all 14 cases the US headquarters of the firms are included in the location analysis, and these employ key decision-making personnel across all functions, not just R&D. Including all facilities may create a spurious co-location effect, since major pharmaceutical firms maintain branches, sales offices and manufacturing facilities in many locations. Some are also diversified into consumer products and other markets that are not the focus of this study. Hence, including only pharamcutical-based R&D labs is a much more conservative test of co-location than if all sites of the pharmaceutical firms were included.

[8] In most cases the establishment date of the R&D labs correspond across all these sources: in case of a discrepancy I use the earliest date listed on any of those sources.

Variables

Co-location – To calculate distance between two firms, I first identify the longitude and latitude coordinates of all biotechnology firms and the R&D labs of the pharmaceutical firms. For each pair of firms, I calculate all possible distances in miles between them at the time of an alliance. That is, for an alliance in 1997 I calculate the distance between the biotechnology firm and each R&D lab of the pharmaceutical partner that existed in 1997. From the set of all possible distances between two firms, I select the minimum distance and use this as the distance measure between firms.[9] This measure only infers that the R&D lab included in the minimum distance calculation initiated the alliance. While it is possible that the alliance was initiated elsewhere in the organization, the selection criteria for inclusion of the R&D labs of the large firms are conservative, and controls are included for the locations of the labs: both of these factors work against a statistically significant finding of co-location. Therefore, if significant co-location effects are nonetheless found, that raises confidence that the distance measure does in fact pick up on the locations of alliances.

Broadscope and domain-specific knowledge - The accumulated knowledge of firms are constructed from their prior experience in drug discovery. Two sets of variables are constructed, to distinguish between broadscope and domain-specific knowledge, respectively:

Biotechnology knowledge is the proportion of drugs developed by a firm that utilize biotechnology techniques. It is calculated as the number of drugs a firm developed using biotechnology techniques, cumulated up to the year of the alliance, divided by the total number of drugs developed by the firm up to that year. Because biotechnologies are process technologies that can be utilized across many disease markets, this variable captures the accumulated broadscope knowledge of a firm up to the time of an alliance.

[9] The set of possible distances vary over time because of shifts in the locational networks of the pharmaceutical firms. If an R&D lab opened subsequent to the year of the alliance, it is not included in the distance calculations. For instance, Bayer opened a NJ location in 1998; that location would not be included in the calculations for Bayer prior to that year.

Disease-specific knowledge For those alliances where a target disease is listed, this variable is the proportion of all drugs a firm developed in that disease category, cumulated up to the year of the alliance. This variable captures the accumulated prior knowledge of a firm of markets and drugs in the disease targeted in an alliance. It is set to 0 for alliances where no disease is specified, and in regressions using this variable I either control for whether disease is specified in an alliance, and also estimate models on the subsample that specify a target disease.

Transactional hazards, learning and coordination variables

Alliance experience is the count of the number of a firm's prior alliances with all other firms, up to the time of an alliance. Past alliance experience should reveal information about a firms' reputation and the quality of firms' management. If proximity reduces information asymmetries and opportunism, there should be a negative relationship between prior alliances and the distance between partners.

Collaboration - This variable is coded 1 if the project lists Codevelopment or Collaboration in the project description, 0 otherwise. If proximity facilitates inter-firm learning and lowers coordination costs, this variable should be positively related to co-location.

Early Stage Project – Drug development is a staged process, with uncertainty about the likelihood of future success declining with each successive stage. To capture differences in technological uncertainty, the variable is coded 1 if an alliance was signed at one of the following stages: Lead Molecule, Development, Preclinical or Phase 1 or clinical trials. The variable is coded 0 for later stages: Phase 2 and 3 of clinical trials, new drug application (IND) Filed, Approved Drug, Formulation. This variable captures both transactional hazards and learning effects. Early-stage technologies developed by small firms may be subject to moral hazard and information asymmetries if entrepreneurial firms "oversell" their projects. The more early-stage and uncertain a project, the more it may benefit from close interaction between collaborating firms. If proximity facilitates monitoring, information flows, and coordination, this variable should be positively related to co-location.

Prior Alliances is a control variable that is coded 1 if the two firms had any prior alliance, 0 otherwise. Everything else equal, it is expected to have a positive impact on the likelihood of follow-on alliances.

Univariate analysis of co-location

For each alliance, I calculate the distance between the biotechnology firm partner and each of the fourteen pharmaceutical firms. Figure 1 displays kernel density estimators (smoothed histograms) of distance distributions between pairs of firms that had an alliance (dashed line) and those that did not (solid line). Both distributions show a bimodal pattern, with a large share of distances under a few hundred miles, and another group of distances greater than 2000 miles, with few distances between these groups. This distribution likely reflects the concentrations of the industry on the two coastal regions of the United States. The figure also shows that the across the full sample, the incidence of proximity between alliance partners is *lower* than the proximity between non-allied firms. This is contrary to the expectation that proximity should be higher for allied partners than the baseline distribution of firms in space and provides some refutation of the idea that on the margin proximity increases the likelihood of a tie forming between two firms.

As discussed above, locations likely matter a great deal to the incidence of co-location, not least because they exhibit different concentrations of firms with which partner. These locational effects are sharply illustrated in Figures 2A-2D, which show the distributions for four important locations of the pharmaceutical partner R&D lab: Massachusetts, Northern and Southern California, and New Jersey. The charts indeed indicate that the localization of ties is strongly conditioned by the actual locations of the firms, with localization very strong in Massachusetts (primarily the Cambridge/Boston region) and weakest in New Jersey, where most large pharmaceutical firms' headquarters are located. Therefore, in all models I control for the locations of the R&D lab of the pharmaceutical firms.

Choice models of alliance formation

I first estimate conditional logit models of the probability of an alliance between two firms, conditional on their being co-located. For computational tractability and to minimize missing data, models are estimated from the perspective of the biotechnology firm choosing to partner with one of the fourteen pharmaceutical firms.

I estimate the following model:

$$P(Allianceijt) = \frac{EXP(\alpha Prioralliance_{ijt} + \beta Colocation_{ijt} + \gamma_{jt} + \mu_t + \varphi_{jt} + \varepsilon}{\sum_{j=1}^{J} EXP(\alpha Prioralliance + \beta Colocation + \gamma + \varphi + \varepsilon)} \quad [1]$$

Where P(*Allianceijt)* is the probability that biotechnology firm *i* enters an alliance with pharmaceutical firm *j* in year *t.* The dependent variable is a coded 1 if a pharmaceutical firm was selected for an alliance and 0 for the 13 other possible partners. Independent variables capture the effect of geographic distance between partners, and variation across each of the fourteen pharmaceutical firms.

Distance between partners is measured in two ways. In Model 1 it is a continuous variable that measures the miles between two firms. In model 2, distance is conceptualized as a discrete variable, to capture the non-linear and discontinuous relationship between distance and face to face interactions: after a certain point, personal contact becomes impractical and its frequency drops sharply. Prior research suggests that localization is specific to the nature of the interaction; I therefore allow the data to indicate the natural cutoff point for this to occur[10]. Figure 1 indicates that the majority of distances in the co-located mode fall below 250 miles (for all distances). I therefore estimate models using 250, 100, and 50 miles as cutoff points for co-location. The results on the co-location variable are nearly identical for the 50 mile and 100 mile cutoffs, but at the 250 mile cutoff no model specifications report significant proximity effects. This is intuitively consistent with the idea that distances greater than 100 miles are impractical for routine interactions and chance encounters. I therefore report regressions with the dependent variable set to 1 if two firms are located within 100 miles of eachother.[11]

The set of variables in γ_{jt} are time-varying measures for each of the fourteen pharmaceutical firms. To capture the impact of a firms' prior knowledge on its ability to identify and exploit external knowledge (Cohen and Levinthal, 1990; Zahra and George, 2002), I include the variables *Biotechnology prior knowledge* and *Disease-specific prior knowledge* for each of the pharmaceutical firms. Arora and

[10] For example, Allen (1977) finds that the probability of communication among engineers working at the same firm drops precipitously after a distance of only 30 meters; Bottazi and Peri (2003) find that research and development spillovers occur within a boundary of 300 miles and fall sharply beyond that point; Gittelman (2007) finds that the cutoff for collaborative research teams is 50 miles.

[11] Results with co-location measured at 50 miles available on request.

Gambardella (1990) find that pharmaceutical firms with higher in-house technological assets in biotechnology are more likely to form alliances with small biotechnology firms. Increasing disease-specific experience can generate targeted search for additional or complementary technologies from external sources. *Alliance experience* may also impact the likelihood of an alliance: to the extent that firms have prior alliance experience, they may be more likely to engage in an alliance. However, it is possible that firms reach a saturation point in the number of alliances they can manage, such that this variable may be negatively related to alliance formation. *Prior alliance* is a control variable coded 1 if the biotechnology and the pharmaceutical firm ever had an alliance up to time t, which is expected to increase the likelihood of follow-on alliances, everything else equal. Additional controls include μt, year fixed effects, and φjt which are fixed effects for the state locations of the pharmaceutical firms' R&D lab used to calculate the co-location variable.

Table 2 shows summary statistics and correlations of the variables. Table 3 reports results of the conditional logit estimations. Both specifications indicate that decreasing distance increases the likelihood of an alliance: in model 1, increasing (decreasing) distance lowers (raises) the probability of alliance ($p<0.01$) and in Model 2, the likelihood of an alliance is greater if two firms are located within 100 miles ($p<0.05$). However, the effects of distance are small compared to the other variables, in particular whether two firms had a prior alliance and the prior knowledge of the pharmaceutical firm. Both knowledge variables are positive and, as expected, the stronger effect is on the prior disease-specific knowledge of the pharmaceutical firm ($p<0.01$)[12]. Alliance experience is negative ($p<0.01$), indicating that there may be decreasing returns to collaboration. The control for prior alliance between two firms is, as expected, positive ($p<0.01$).

Models of co-location: Transactional hazards, collaboration, and knowledge effects

The second set of models explores the mechanisms that influence the probability that, conditional on two firms having an alliance, they will be co-located. Independent variables test for the mechanisms

[12] Because the variable disease_specified is invariate across choices in the models, it is not included as a regressor. Results of both regressions are robust in subsamples that only include alliances for which a target disease is specified.

of lowering transactional hazards, facilitating learning and collaboration, and absorptive capacity effects. In order to capture effects that operate through personal contact and face-to-face interaction, the dependent variable is a discrete measure of co-location that is coded 1 if two alliance partners are within 100 miles of one another.[13] Models are estimated with logit regression, and all specifications include fixed effects for the locations of the pharmaceutical firms' R&D lab used to calculate distance from the biotechnology firm. Since there are multiple alliances for each of the large pharmaceutical firms, the dependent variable is likely to violate the independence assumption, so I estimate random-effects models to account for within-firm correlation for each of the 14 pharmaceutical firms.

Table 4 shows summary statistics and correlations, and table 5 shows the results of the logit models. Model 1 reports effects for firm experience and project-specific variables that test for transactional hazards and collaboration effects. Alliance experience is a proxy for firm-level reputation, so increasing experience should lessen the importance of co-location. However, the results show no significant effect for this variable, so hypothesis 1 is not supported.[14] Increased project uncertainty was predicted to have a positive effect on the likelihood of collocation because of the increased need for coordination, information flows, and monitoring. The coefficient on the variable measuring early-stage technology is negative ($p<001$): the more early stage the project, the *less* likely partners are co-located, which is opposite of the theorized effect. However there is support for Hypothesis 2 that predicted that the need to collaborate increases the importance of co-location: the coefficient for collaboration is positive ($p<0.05$), but the effect weakens in subsequent specifications. Together these results provide no or weak support for the idea that co-location lowers transactional hazards or facilitates coordination and learning.

Models 2 and 3 add the absorptive capacity variables. In Model 2, all dyads are included, but since not all of them list a specific disease I add a variable *disease_specified* which is set to one if an

[13] Results are consistent in regressions estimating co-location as less than 50 miles.
[14] In alternative specifications I estimate the models using the logged number of drugs developed by the firms. Results are similar to those reported using alliances. Since this variable is collinear with number of alliances for the biotechnology firms, I do not include it in the specifications reported here.

alliance includes a target disease in its description. In model 3 I estimate the same model on a subsample of alliances that have disease specified. The results are similar to Model 2, which I discuss. The variables measuring prior knowledge of the pharmaceutical firm have opposite effects on the likelihood of proximity: a firms' prior knowledge of biotechnology has a large positive coefficient ($p<0.01$) but expertise in a specific disease category is negative, indicating a lower likelihood of co-location ($p<0.01$). These findings are consistent with the hypothesized effects of broadscope and specific knowledge on the localization of alliances (H3a and H3b).

However, the variables on firm-level knowledge for the biotechnology firms differ from the predicted effects from those found for large firms. The variable on prior knowledge in biotechnology drugs is not significant, indicating that prior experience in biotechnology technologies is not a differentiating factor for the localization of alliances of small biotechnology firms. This could reflect the fact that small firms as a group are specialized in biotechnologies, so individual variation on this dimension does not provide meaningful additional information about their capabilities. The disease-specific knowledge of the biotechnology firm partner, on the other hand, is positive, increasing the likelihood of co-location ($p<0.10$ in Model 1, $p<0.05$ in Model 2). These results suggest that the impact of domain-specific knowledge on the geography of ties is asymmetric for large and small firms. Whereas disease-specific knowledge in large firms is associated with non-local partnering, disease-specific knowledge of small firms increases localization. The results are consistent with a pattern in which small firms with domain-specific knowledge are preferentially selected by large firms in the local region, but large firms with specialized expertise search broadly to partner in those areas.

If co-location is important in generating serendipitous encounters, the effects should be strong for the initial ties between two firms. Therefore, in Model 4 I restrict the sample to firms that that had no history of a prior alliance. The core results on firm knowledge effects remain.

Additional test: technological distances between firm's drug portfolios and alliances

The results show that the broadscope knowledge of large firms is associated with localized alliances, whereas domain-specific knowledge increases distant ties. The results are consistent with the

hypothesis that their localized ties are unplanned and opportunistic whereas distant ties reflect intentional, problem-oriented search. A corollary of this finding is the expectation that localized ties will be more heterogeneous than distant alliances with respect to a firms' internal R&D.

To explore this further, I measure the technological distance between a pharmaceutical firms' portfolio of drugs and its co-located and distant alliances. I adopt Jaffe's distance measure of the angular separation (uncentered correlation) between firms' portfolios of technologies (Jaffe, 1986). For each of the fourteen large pharmaceutical firms, I construct three vectors. The first is a count of the firms' drugs developed in each of the sixteen disease categories. The second vector counts the local alliances in each of those disease categories, and the third is a count of non-local alliances by disease category. Following the regression models, a local alliance is defined as an alliance with a firm within 100 miles of an R&D lab[15]. Correlation coefficients are calculated between the firm's portfolio of drugs and its local and non-local alliances. The coefficient varies between 0 for no overlap to 1 for perfectly correlated. If localized opportunities are more heterogeneous with respect to firms' internal R&D, then the correlations should be lower for the co-located alliances as compared to the distant alliances.

Figure 3 shows the correlations for each of the 14 firms. The chart indicates that for each firm, the localized alliances have lower correlation coefficients than distant alliances with respect the firms' drug portfolios. This result, while not a formal test of H3, provides additional support for the idea that regions generate chance encounters between firms, whereas distant alliances are motivated by specific problem-oriented search.

Discussion

"Chance favors the prepared mind" is a succinct summation of the process of scientific discovery: discovery builds on the unexpected occurrence, but scientists need prior knowledge to identify and exploit these opportunities. This idea has been extended to the organizational level: firms' prior investments in R&D positions them to not only absorb new information but to place educated bets on a risky future (Cohen and Levinthal, 1994). I extend this idea to geographic space. I propose that local

[15] Similar results are obtained with localization defined at 50 miles.

spaces provide opportunities for serendipitous, unplanned encounters with firms they may otherwise not meet. While this represents a form of spillovers, "being there" is not enough: firms need in-house knowledge to take advantages of new partnering opportunities. Indeed, the strongest effects on both the formation and the geographic location of alliances are variables that measure firm's prior knowledge, with little support for alternative mechanisms that co-location reduces transactional hazards or facilitates learning. For large firms, broadscope knowledge increases the likelihood of a local partner, whereas domain-specific knowledge is negatively related to localization. These differently signed knowledge effects are consistent with two contrasting views of space articulated by Jane Jacobs and Diana Crane - that regions spark innovation by bringing together diverse firms in new combinations, whereas "invisible colleges" of researchers working on common problems can connect readily across spatial boundaries.

Two new and surprising results emerge from this. The first is that for large firms with multiple R&D sites, localized search in technological and geographic space may be inversely related. Whereas local environments are is often conceptualized as offering limited new knowledge, the variety made possible from interactions in a region can result in new combinations that would otherwise not occur. The results can be interpreted as showing that for large firms, local regions are sites of exploration and generate technological variety, while distant ties entail search along specific market trajectories and localized learning.

A second surprising result is that the knowledge variables operate differently for small and large firms in predicting localization of ties. Whereas the domain-specific knowledge of large firms is associated with distant partnerships, for small firms it increases the likelihood of localization. The different effects of knowledge on localization indicate that for small firms, local ties deepen prior specific knowledge. These results square with prior literature on the embeddedness of small firms in local networks that reflect shared technological specializations and access to local pools of human capital and other resources (Almeida and Kogut, 1997, Owen Smith and Powell, 2004, Zucker et al, 1998). The co-occurrence or geographic clustering and technological specialization among small, knowledge-intensive firms can serve to attract R&D investments by large firms precisely to generate variety and gain exposure

to new technological directions in a field. The findings here indicate that prior investment in broadscope knowledge enhances their ability to exploit these opportunities.

The findings have the appeal of being consistent with accounts from practitioners. The univariate analysis shows that a significant share of partnerships is localized, and the models of alliance formation show that co-location does increase the likelihood of a partnership between two firms (though the effect is not very strong). In discussions with licensing executives, however, they relate that co-location does not play a major role in their alliance strategies, and emphasize that they scan globally in their search for partnering opportunities. It is possible to reconcile these accounts with the observed importance of co-location, if we consider the multiple channels by which partnerships are formed, and the specific channels through which licensing managers operate. Alliances can arise through the formal, intentional search and negotiations processes carried out by dedicated personnel, and they can also arise through informal, unplanned encounters at multiple levels and locations of the organization. The data analyzed in this study capture alliances that arise from both of these processes, but cannot distinguish between them; however, the results are suggestive that formalized processes of search lead to alliances that are broadly distributed in space, whereas informal encounters are more likely to occur in localized spaces. This opens up questions for future research into the formal and informal organization routines that generate firms' portfolio of alliances. The results here suggest that these processes will lead to different sets of partnerships that are not only geographically distinctive but represent different kinds of technological opportunities as well.

Finally, the results underscore how heterogeneity in firms' absorptive capacity generates differences in sourcing external knowledge in space. External knowledge is potentially available to all firms in an industry, and in order for firms to gain strategic advantage from such partnerships, they need the capability to locate, evaluate, and contract for relationships in a manner that is distinctive from other firms competing for similar relationships. The results of this study unpack the specific kinds of knowledge that is useful in exploiting external knowledge in geographic space.

Figure 1. Dashed line shows distances for actual alliances between biotechnology firms and the nearest R&D lab of a pharmaceutical firm. Solid line shows distances between biotechnology firms & the nearest R&D lab of 13 other pharmaceutical firms at time of alliance.

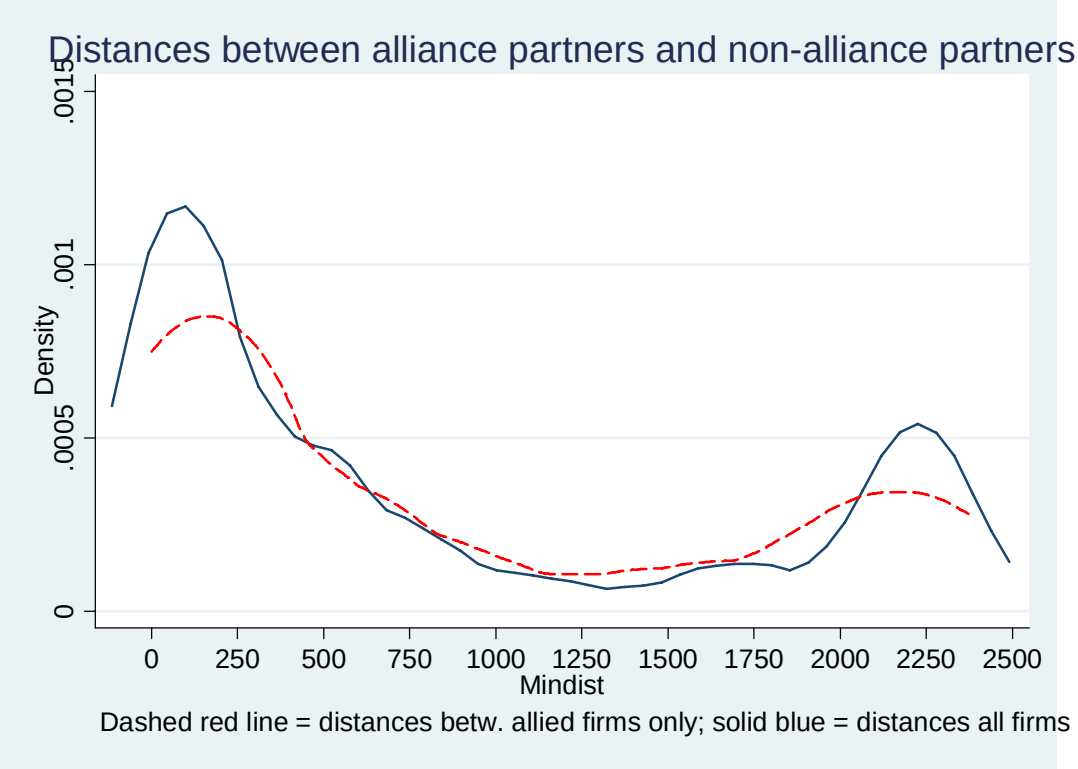
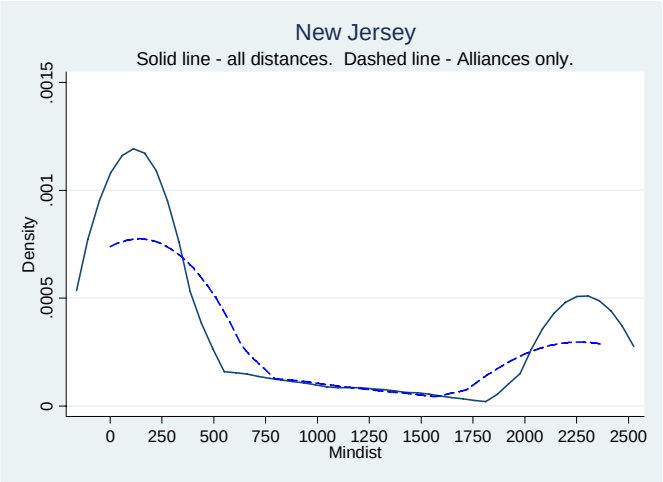
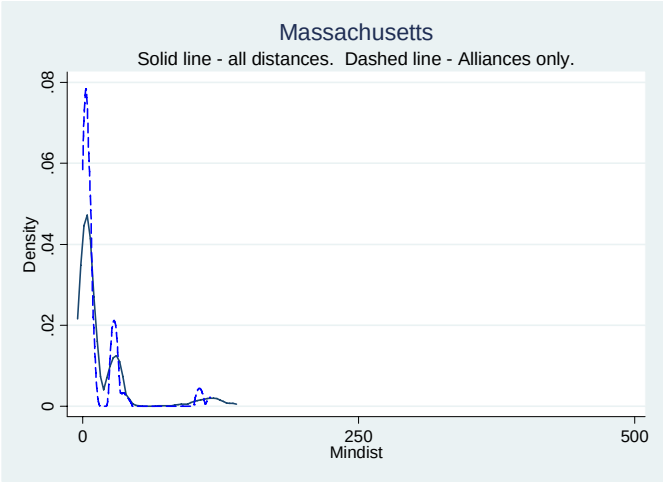
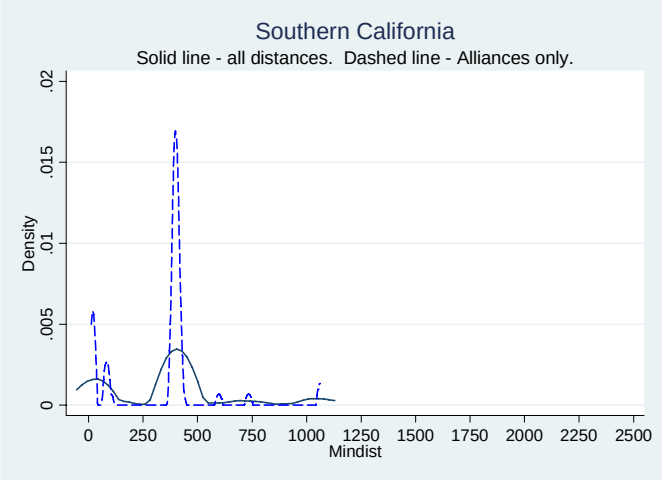


Figure 2. Distribution of distances by major states of R&D labs.
Dashed lines show distances between biotechnology firms and the nearest R&D lab of their alliance partners. Solid lines show distances between biotechnology firms & the nearest R&D lab of 13 other pharmaceutical firms at time of alliance. Only minimum distances that involve an R&D lab from the displayed state are shown.







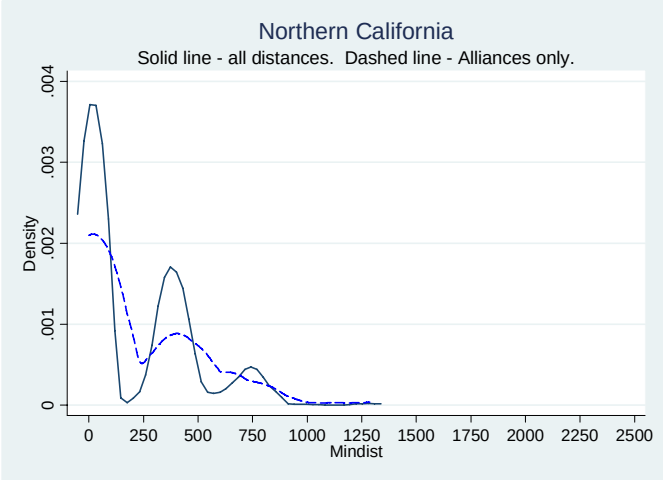


Table 1A. Locations of R&D Labs of 14 large pharmaceutical firms

State	Number of firms with a lab in state	Number of labs in state
New Jersey	10	14
California	8	10
Massachusetts	7	10
Pennsylvania	6	8
New York	4	4
Connecticut	3	3
Illinois	1	3
Other (7 states)	7	7

Table 1B. Locations of biotechnology firms by state

	Number of firms
California	149
Massachusetts	54
New Jersey	35
Pennsylvania	20
Maryland	19
Washington	15
North Carolina	13
New York	11
Texas	9
Colorado	7
Connecticut	6
Utah	5
Georgia	4
Wisconsin	4
Alabama	3
Arizona	3
Florida	3
Illinois	3
Missouri	3
Other	14
Number of biotechnology firms	380

Other states are: Minnesota, Montana, Ohio, Tennessee, Virginia, Indiana, Kansas, New Hampshire, and Rhode Island.

Table 2. Summary statistics and correlations for conditional logit models of alliance formation

Variable	Obs	Mean	Std. Dev	Min	Max	Alliance	Prior Alliance	Distance less than 100 miles	Number alliances pharma firm (log)	Biotech knowledge, pharma firm	Disease knowledge, pharma firm
Alliance	14532	0.07	0.26	0	1	1					
Prior Alliance	14532	0.02	0.13	0	1	0.20	1				
Distance less than 100 miles	14532	0.29	0.45	0	1	0.01	0.01	1			
Number alliances pharma firm (log)	14532	5.39	0.82	2.30	6.68	-0.03	0.04	0.06	1		
Biotech knowledge, pharma firm	14532	0.06	0.04	0.01	0.16	0.02	0.01	0.00	0.26	1	
Disease-specific knowledge, pharma firm	14532	0.07	0.07	0	0.32	0.02	0.00	0.04	0.00	0.01	1

Table 3. Conditional logit models of alliance formation
Include fixed effects for locations of R&D labs of pharmaceutical firm

	Model 1 Dependent variable = Distance in miles		Model 2 Dep. Variable =1 if distance between firms is < 100 miles	
	Coeff	S.E	Coeff	S.E
Distance between alliance partners	-0.00***	0.00	0.23**	0.10
Prior Alliance	2.62***	0.14	2.62***	0.14
Number of alliances, Pharma firm (log)	-0.23***	0.05	-0.24***	0.05
Biotech Knowledge, Pharma firm	2.1**	1.0	2.1**	1.04
Disease-Specific Knowledge, Pharma firm	4.2***	1.1	4.2***	
N	14250		14250	
LR Chi-Square	382.74		378.8	
Log Likelihood	-2527.1		-2529.0	

*p<0.10
** p<0.05
***p<0.01

Note: Locations of the R&D labs are grouped into regions of contiguous states as follows: Region 1 – Illinois, Michigan, Indiana, Ohio. Region 2 – Massachusetts and Connecticut. Region 3 – North Carolina and Virginia. Region 4 – New Jersey, New York Pennsylvania Delaware. Region 5 – Northern California. Region 6 – Southern California. Texas and Washington state are included as separate locations.

Table 4. Summary statistics and correlations of variables in logit models

Variable	Obs	Mean	Std. Dev.	Min	Max	Colocation, less than 100 miles	Year	Alliance experience, Biotech firm	Alliance experience, Pharma firm	Prior Alliances	Early Stage Project	Collaboration	Dis-ease knowledge, pharma	Biotech knowledge, Pharma	Dis-ease knowledge, biotech firm
Colocation, less than 100 miles	1038	0.30	0.46	0	1	1									
Year	1038	2001	3.83	1993	2008	0.12	1								
Alliance experience, Biotech firm[16]	1038	2.32	1.40	0	5.5	0.04	0.13	1							
Alliance experience, Pharma firm[1]	1038	5.31	0.88	2.3	6.7	0.03	0.50	0.04	1						
Prior Alliances	1038	0.11	0.32	0	1	0.01	-0.01	0.21	0.13	1					
Early Stage Project	1038	0.64	0.48	0	1	-0.05	0.07	-0.12	-0.02	-0.04	1				
Collaboration	1038	0.39	0.49	0	1	0.05	0.16	0.04	0.02	0.04	0.25	1			
DiseaseKnowledge, pharma firm	1038	0.09	0.07	0.00	0.32	0.02	0.03	-0.03	0.02	0.00	-0.08	0.02	1		
Biotech Knowledge, pharma firm	1038	0.07	0.04	0.02	0.16	0.05	0.09	0.06	0.29	0.00	-0.08	-0.05	0.05	1	
Disease Knowledge, biotech firm	1038	0.15	0.27	0.00	1.00	0.09	0.06	-0.05	0.03	-0.03	-0.05	0.05	0.35	0.03	1

[16] Calculated as the log of 1 plus the number of all prior alliances of the firm

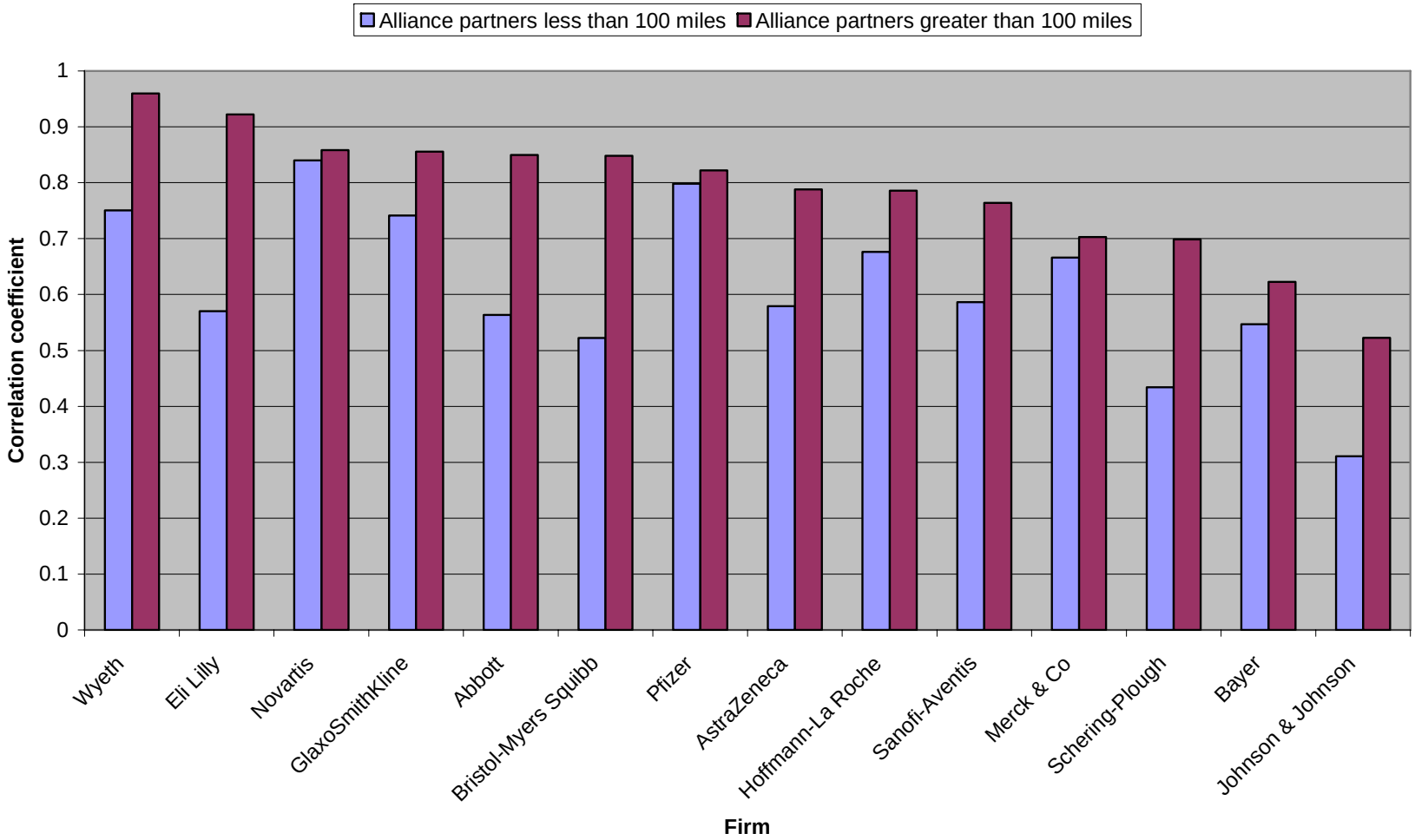
Table 5. Random effects logit Models. Dependent variable: Co-location = partners within 100 miles. All models are estimated with fixed effects for locations of pharmaceutical firms' R&D labs and random effects for each of the 14 pharmaceutical firms

	Model 1. Firm Alliance experience and project characteristics All Alliances		Model 2. Firm Knowledge Variables All alliances		Model 3. Alliances with Disease Specified Only		Model 4. First time alliances only	
	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE
Alliance experience, Biotech firm	-0.07	0.06	-0.08	0.06	-0.06	0.07	-0.10	0.6
Alliance experience, Pharma firm	0.16	0.15	0.02	0.13	-0.09	0.12	0.05	0.14
Prior Alliances	0.22	0.28	0.20	0.28	0.17	0.34		
Early Stage	-0.61***	0.17	-0.35*	0.18	-0.43**	0.20	-0.27	0.19
Collaboration	0.36**	0.17	0.30*	0.17	0.32	0.20	0.26	0.18
Disease knowledge, pharma firm			-3.41***	1.46	-2.98**	1.44	-3.1**	1.5
Biotech knowledge, pharma firm			11.16***	3.7	9.23***	3.25	10.8***	3.7
Disease knowledge, biotech firm			0.58*	0.31	0.60**	0.31	0.56*	0.33
Biotech knowledge, biotech firm			-0.07	0.26	0.04	0.31	0.02	0.28
Disease Specified			1.05***	0.26			1.05***	0.28
N	1038		1038		709		919	
Wald Chi Square	162.5		169.68		103.39		153.4	
Log Likelihood	-496.6		-479.67		-346.70		-423.9	

*p<0.10** p<0.05***p<0.01

Note: Locations of the R&D labs are grouped into regions or contiguous states as follows: Region 1 – Illinois, Michigan, Indiana, Ohio. Region 2 – Massachusetts and Connecticut. Region 3 – North Carolina and Virginia. Region 4 – New Jersey, New York Pennsylvania Delaware. Region 5 – Northern California. Region 6 – Southern California. Texas and Washington state are included as separate locations.

Figure 3. Correlations between alliances and in-house drugs
Alliances < 100 miles and > 100 miles



References

Agrawal, Ajay, Iain Cockburn, and John McHale. 2006. "Gone but not forgotten: knowledge flows, labor mobility, and enduring social relationships." *Journal of Economic Geography* 6, (5) 571-591.

Agrawal, Ajay, Devesh Kapur, and John McHale. 2008. "How do spatial and social proximity influence knowledge flows? Evidence from patent data." Journal of Urban Economics 64 (2) 258-269.

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

Almeida, P., B. Kogut. 1999. Localization of knowledge and the mobility of engineers in regional networks. *Management Science*. 45(7) 905–918.

Arora, A., Gambardella, A., 1990. Complementarity and external linkages: the strategies of the large firms in biotechnology. *The Journal of Industrial Economics* 38 (4)361–379.

Arora, Fosfuri and Gambardella. 2001. Markets for technology and their implications for corporate strategy. *Industrial and Corporate Change* 10(2) 419-451.

Arora, A., Gambardella, A., Magazzini, L., Pammolli, F. 2007. A Breath of Fresh Air? Firm types, scale, scope and selection effects in drug development. Working paper, Carnegie Mellon University.

Audretsch, D.B., Stephan, P.E., 1996. Company–scientist locational links: the case of biotechnology. *American Economic Review* 86 (3) 641–652.

Bathelt, H., Malmberg, A. & Maskell, P. (2004) Clusters and knowledge: Local buzz, global pipelines and the process of knowledge creation, Progress in Human Geography, 28, pp. 31–56.

Borgatti, Stephen and Rob Cross. 2003. "A Relational View of Information Seeking in Social Networks" *Management Science* 49(4).

Bottazi, L., G. Peri. 2003. Innovation and spillovers in regions: Evidence from European patent data. *Eur. Econom. Rev.* 47 687–710.

Cockburn, Iain M., and Rebecca M. Henderson. 1998. "Absorptive Capacity, Coauthoring Behavior, and the Organization of Research in Drug Discovery." *Journal of Industrial Economics* 46, (2) 157-182.

Cohen, Wesley M., and Daniel A. Levinthal. 1990. "Absorptive Capacity: A New Perspective on Learning and Innovation." *Administrative Science Quarterly* 35 (1) 128-152.

Cohen, Wesley M., and Daniel A. Levinthal. 1994. "Fortune Favors the Prepared Firm." *Management Science* 40 (2) 227-251.

Crane, Diana. 1971. "Transnational Networks in Basic Science." *International Organization* 25, (3) 585.

Ellison, Glen and Edward Glaeser. 1997. "Geographic Concentration in US Manufacturing Industries: A Dartboard Approach", *Journal of Political Economy* 1055.

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

Gambardella, Alfonso, Paola Giuri, Alessandra Luzzi. 2007. "The market for patents in Europe", *Research Policy,* 36:1163-1183

Gans, Joshua, David Hsu, Scott Stern. 2008. "The Impact of Uncertain Intellectual Property Rights on the Market for Ideas: Evidence from Patent Grant Delays", *Management Science* 54(5):982-997.

Gittelman, Michelle, and Bruce Kogut. 2003. "Does Good Science Lead to Valuable Knowledge? Biotechnology Firms and the Evolutionary Logic of Citation Patterns." *Management Science* 49, (4) 366-382.

Gomes-Casseres Ben, John Hagedoorn, Adam B. Jaffe. 2006. Do alliances promote knowledge flows? *Journal of Financial Economics* (80)5–33

Guedj, I., D. Scharfstein. 2005. Organizational Scope and Investment: Evidence from the Drug Development Strategies and Performance of Biopharmaceutical Firms. NBER Working Paper 10933. Cambridge MA.

Jacobs, Jane. 1969. *The economy of cities*. New York: Random House.

Jaffe, Adam B. 1986. "Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits, and Market Values." *American Economic Review* 76, no. 5: 984.

Kim, Dong-Jae, and Bruce Kogut. 1996. "Technological Platforms and Diversification", *Organization Science* 7 3:283-301

Kogut, Bruce. 1991 "Joint ventures and the option to expand and acquire." *Management Science* 37, no. 1:19-33.

Lane, Peter J., and Michael Lubatkin. 1998. "Relative absorptive capacity and interorganizational learning." *Strategic Management Journal* 19, no. 5: 461.

Lerner, Josh, and Robert P. Merges. 1998."The control of technology alliances: An empirical analysis of the biotechnology industry." *Journal of Industrial Economics* 46, no. 2: 125-156.

Malloy, Christopher J. 2005. "The Geography of Equity Analysis." Journal of Finance 60, no. 2: 719-755.

Mowery, David C., Oxley, Joanne E., and Brian S. Silverman. "Strategic Alliances and Interfirm Knowledge Transfer" *Strategic Management Journal* 17 (Winter96 Special Issue 1996): 77-91.

Narula, Rajnesh, and Grazia Santangelo. 2008. "Location, collocation and R&D alliances in the European ICT industry", *Research Policy,* in press.

Owen-Smith, J., W. Powell. 2004. Knowledge networks as channels and conduits: The effects of spillovers in the Boston biotechnology community. *Organization. Science* 15(1) 5–21.

Peterson, Mitchell, and Raghuram Rajan. 2002. "Does Distance Still Matter? The Information Revolution in Small Business Lending", *Journal of Finance* 57(6):2533-2570.

Phene, Anupama, Karin Fladmoe-Lindquist, and Laurence Marsh. 2006. "Breakthrough innovations in the U.S. biotechnology industry: the effects of technological space and geographic origin." *Strategic Management Journal* 27, no. 4: 369-388.

Rosenkopf, Lori, and Paul Almeida. 2003. "Overcoming Local Search Through Alliances and Mobility." *Management Science* 49, (6) 751-766.

Saxenien, A. 1994. *Regional Advantage*. Harvard University Press, Cambridge, MA.

Shenkar, Oded, and Jiatao Li. 1999. "Knowledge Search in International Cooperative Ventures." *Organization Science* 10, (2) 134-143.

Shipilov, Andrew V. 2009. "Firm Scope Experience, Historic Multimarket Contact with Partners, Centrality, and the Relationship Between Structural Holes and Performance." *Organization Science* 20, (1) 85-106.

Sorenson, Olav and Toby Stuart. 2008. "Bringing the Context Back In: Settings and the Search for Syndicate Partners in Venture Capital Investment Networks" *Administrative Science Quarterly,* 53:266-294.

Storper, Michael, and Anthony J. Venables. 2004. "Buzz: face-to-face contact and the urban economy." *Journal of Economic Geography* 4, (4) 351-370.

Stuart, T., O. Sorenson. 2003. The geography of opportunity: Spatial heterogeneity in founding rates and the performance of biotechnology firms. *Res. Policy (*32) 229–253.

Tripsas, Mary. 1997. "Surviving Radical Technological Change through Dynamic Capability: Evidence from the Typesetter Industry." *Industrial & Corporate Change* 6, (2) 341-377.

Zahra, Shaker A., and Gerard George. 2002. "Absorptive Capacity: A review, reconceptualization, and Extension." *Academy of Management Review* 27, (2) 185-203.

Zucker, L.G., Darby, M.R., Brewer, M.B., 1998. Intellectual human capital and the birth of U.S. biotechnology enterprises. *American Economic Review* 88 (1), 290–306.