DRUID society

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

A Spatial Ecology of Structural Holes

Christopher C. Liu
University of Toronto/Rotman School of Management
Strategic Management
chris.liu@rotman.utoronto.ca

Abstract
A Spatial Ecology of Structural Holes: Scientists and Communication at a Biotechnology Firm

Jelcodes:O32,O31

A Spatial Ecology of Structural Holes: Scientists and Communication at a Biotechnology Firm

Abstract
This paper focuses on the relationships between individuals, their local spatial environment, and their ability to form networks that span the internal structure of an organization. I contend that the prevailing view about spatial proximity – that high spatial density leads to greater opportunities to expand one's network – is at best incomplete and potentially misleading. Instead, I argue that an individual's spatial proximity must be considered in an ecological context – i.e., in relation to the location and expertise of proximate colleagues. Using a unique dataset of email communication patterns, fine-grained blueprints of office-locations, scientific publication records and associated keywords, as well as the formal organization chart from the research division of a biotechnology firm, I find that co-location alone is not associated with an individual's ability to assume network positions that bridge structural holes. In fact, spatial proximity to workgroup (i.e., same laboratory) members inhibits an individual's ability to form networks that reach across the organization. By contrast, proximity conditioned on the expertise overlap of nearby, non-workgroup (i.e., different laboratory) colleagues is associated with having a less constrained network (i.e., one rich in structural holes). Lastly, spatial ecological effects are as important as human capital in providing an individual with opportunities to occupy brokering positions in the network. This paper contributes to the literature on networks and the organization of innovative work by developing a theory of spatial ecology and demonstrating its connection to network structural features that have been repeatedly linked to superior performance in knowledge production contexts.

I. Introduction

A large literature suggests that there are benefits to serving as an intermediary between unconnected parties. These intermediaries, network brokers that span structural holes, have been shown to have better ideas (Burt 2004; Fleming, Mingo and Chen 2007), performance outcomes (Burt 1997), as well as greater success in both job searches (Granovetter 1974) and internal promotions rates (Podolny and Baron 1997). Despite rapid progress on the consequences of brokered positions, we still know relatively little about how these positions arise. This dearth stems not from a lack of interest, but from the empirical difficulties in collecting multi-dimensional network data over time. In particular, what factors enable or constrain an individual's ability to occupy brokerage positions within a network? Furthermore, for settings where expansive networks are widely viewed as beneficial, why aren't all networks rich in structural holes?

Distilled, the burgeoning literature on brokerage antecedents falls into two streams that both emphasize an individual's ability (or inability) to foster relationship partners. Underlying both strands of the literature is a broker's ability to attract a diversity of relationship partners. The very fact that an individual has a wide number of potential relationships allows that actor to be selective in his or her choice of relational partners, to prune their relationships, and to have an unconstrained network (Davis 2009; Ibarra 1999). Thus, an individual's ability to form a diverse array of relationships underlies his or her occupation of brokerage positions within a network.

First, a number of studies emphasize the attributes of the focal actor themselves. Some individuals are more desirable relationship partners than others. For example, a person's association with high status individuals may result in both direct and indirect benefits (Podolny 2001; Stuart, Hoang and Hybels 1999). As a result, high status individuals are often sought out for relationships. Consistent with this notion is the observation that brokers often have high human capital, and partners seek out relationships with brokers to access their superior knowledge stocks (Allen 1977; Burt 1992).

A second body of research stresses environmental factors that shape an individual's ability to form relationships. Most prominently, spatial proximity decreases the costs of initiating and maintaining relationships, suggesting that high spatial density clusters facilitate broader networks (Festinger, Schachter and Back 1950; Sorenson and Stuart 2008). By contrast, an ecological perspective suggests that spatial proximity may also increase competition for relationship partners, contingent on the social overlap between co-located individuals (Hannan and Freeman 1989; Hawley 1986). For homogeneous clusters, the relationships that do form are also likely to be redundant, further inhibiting an individual's ability to form relationships between unconnected parties. Taking into consideration these multi-faceted aspects of co-location, the effects of high spatial density on fostering a network rich in structural holes remain unclear.

With these caveats in mind, I argue that a consideration of either spatial or social proximity alone is insufficient to explain an individuals' ability to form a diverse array of relationships partners. This paper contends that the effects of spatial density on network formation must be considered in concert with the social attributes of spatially agglomerated actors. In contrast to the literature's focus on factors that *enable* relationships, this paper's central concern is the competitive forces that serve to *constrain* an individual's network. In particular, I explore the role environmental factors play in shaping an individual's ability (or inability) to form relationships and navigate through an organization. In deference to Festinger, Schachter, Back's (1950) seminal work on the multi-faceted antecedents to networks, I borrow their use of the term: spatial ecology.

My central concern is the relationship between an individual, his or her localized environment, and the conditions that shape his/her ability to form relationships across the internal structure of an organization. Specifically, to what extent do environmental factors, embodied in the social attributes of an individual's co-located peers, influence an individual's ability to reach across an organization's communication network? How does social and spatial proximity act in concert to either enable or constrain an individual's ability to structure their network? And in the consideration of an organization's communication network, what broader lessons might we derive about the role that internal ecologies play in structuring opportunities for interaction? To address these questions, I exploit a rich and unique dataset that maps expertise and spatial positions onto communication network data for all researchers at a prominent biotechnology firm.

I provide evidence that environmental factors are as important as individual-level attributes in shaping communication flows. Using fine-grained office-blueprints to measure spatial distance and either formal organizational distance or keywords associated with scientific publications to measure expertise overlap, I provide evidence that both spatial and expertise dimensions act in concert to shape an individual's ability to reach through the organization. Moreover, a consideration of spatial density alone has no effect on an individual's brokerage positions. A higher spatial density of scientifically proximate alters *increases* an individual's brokerage opportunities. However, greater co-location with workgroup members serves to *constrain* an actor's network. Thus, geographic effects are contingent: examining spatial density without overlaying the social attributes of co-located partners leads to misleading results. A duality of agglomerative effects, either enabling or constraining an individual's actions, lies at the heart of a spatial ecological framework.

II. Theory and Hypotheses.

Individuals with broad, unconstrained networks have received wide attention as brokers that bridge structural holes (Burt 1992). Individuals with not just a higher number of partners, but a wider array of unconnected partners have greater access to a diversity of information and hence better ideas (Burt 2004; Fleming, Mingo and Chen 2007; Hargadon and Sutton 1997). Although expansive networks clearly yield benefits in a variety of settings, the origins of these brokered positions remain under-explored. Of note, a variety of scholars have focused on the evolution of networks, either through accumulative advantage (Fleming and Waguespack 2007; Merton 1968), structural inertia (Powell et al. 2005; Zaheer and Soda 2009), or entrepreneurial action (Burt 1992). In contrast, this paper focuses on non-structural correlates of the origins of brokerage, particularly environmental factors.

IIa. Spatial Ecology. A prominent environmental factor is geographic proximity (Gieryn 2000). Scholars from a diversity of disciplines have long recognized that geographic distance plays a central role in facilitating the likelihood of relationships. Beginning with Bossard's (1932) study of homogamy and propinquity, many studies have explored the role that geographic distance plays in structuring opportunities for interaction. Physical proximity increases the overall likelihood of interactions, shaping outcomes as diverse as friendship and marriage (Bossard 1932; Festinger, Schachter and Back 1950; Marmaros and Sacerdote 2006), capital flows (Sorenson and Stuart 2001), and competitive outcomes (Baum and Mezias 1992; McPherson 1983).

As geographic distance rises, individual's have less knowledge about the attributes of those distant relationship opportunities, fewer conduits to access that information, as well as greater uncertainty in the veracity of the little knowledge that they do receive (Sorenson and Audia 2000). Moreover, individuals are often unwilling or unmotivated to expend the effort to overcome the difficulties presented by spatial distance (Zipf 1949). For these reasons, spatially distant relationships require greater effort to initiate and maintain. As a result, an individual's activities and relationships are often circumscribed in geographic space. In recognition of the key role that proximity plays in the formation of relationships, the prevailing viewpoint is that a higher geographic density (i.e., number of individuals within a geographic area) is associated with a greater number of relationships and hence, increased information flows (Argote, McEvily and Reagans 2003; Uzzi 1997). Consistent with this perspective, many scholars emphasize the benefits of co-location and high density clusters (Allen 1977; Ancona and Caldwell 1992a; Cummings 2004; Powell, Koput and Smith-Doerr 1996; Zucker and Darby 1996).

By contrast, ecologists emphasize that greater similarity along a vector of attributes increases the overlap in resource dependencies, leading to heightened competition (Hannan and Freeman 1989; Hannan and Carroll 1992). Spatial location is one (of many) prominent characteristics through which competitive pressures may increase (Baum and Mezias 1992; Carroll and Wade 1991; Sorenson and Stuart 2001). Consistent with both the idea of structural equivalence (Burt 1992) and theories of resource-partitioning (Carroll and Swaminathan 2000), individuals in high-density environments may compete with one another for external attention. When faced with a large number of equivalent opportunities, individuals have a tendency to choose the nearest one (Stouffer 1940). As a result, external relationships may accrue to individuals at the periphery of high-density clusters.

Furthermore, a high spatial density of individuals with similar attributes (i.e., homogeneous cluster) is likely to lead to overlapping relationships (Coleman 1988). Although cohesive networks may facilitate trust and the flow of information, extensive similarity between co-located individuals leads to redundant information flows (Hansen 1999; Reagans and Zuckerman 2001). As a result, homogeneity across both spatial and social dimensions may inhibit the ability of co-located individuals to introduce one another to new, non-redundant relationships. Members of homogeneous clusters talk amongst themselves.

The spatial co-location of workgroup members remains a contentious issue. Recently, a number of scholars have suggested that as communication costs decrease, geography plays a less prominent role in the organization of work. This emphasis is widely mirrored in arenas as diverse as open innovation (Chesbrough 2005), distributed teams (Cummings 2004), as well the dissolution of permanent offices

(Elsbach 2003). There is little doubt that modern technology has accelerated the flow of information across geographic and organization space. However, spatial proximity remains an important factor in supporting collaborative work (Kraut, Fussell, Brennan, and Siegel 2002). As members of the same workgroup, there are significant benefits to co-location, including a shared identity, values, norms, as well as ready oversight and support on the part of the group leader (Hackman 2002).

The co-location debate: whether to place laboratory members adjacent to one another or to intercalate them with non-workgroup members is one of the most divisive issues in the organization of scientific work. In university settings, the space allocated to a professor's laboratory is not only indicative of their prestige, but also dictates the capacity of their laboratory, and hence their power. For example, MIT Biology's policy is to allocate two rooms (capacity for a laboratory of ~ 8) to untenured faculty and an additional room as they are promoted. These laboratory rooms are under the sole authority of an individual professor, populated by members of that professor's laboratory. Although this policy of co-location is mirrored in the vast majority of scientific production settings, this dominant policy has recently been called into question. Notably, at UCSF's new Mission Bay campus, the policy is to intercalate individual workspaces on a building-floor between multiple laboratories. Adjacent scientists are no longer members of the same laboratory, ostensibly to promote greater interaction between laboratories.

Within for-profit knowledge-production organizations, the "MIT" or "UCSF" model is contentious. A policy of workgroup co-location is the norm, consistent with an emphasis on the need for oversight and constant feedback. At the same time, the normative structure of scientific workgroups, with frequent meetings and collective work suggest rapid knowledge flows between workgroup members (Owen-Smith 2001). Although there are clear benefits to the co-location of workgroup members, I presume that the benefits of promoting inter-group communication outweigh the costs of decreased intra-group communication, particularly in settings where a diversity of knowledge inputs is valuable. I therefore hypothesize:

H1: Individuals with a lower spatial density of workgroup members have a network rich in structural holes.

On the other hand, co-located individuals that diverge along social dimensions may yield benefits. For example, socially diverse, but spatially proximate clusters may promote a variety of relationships, yielding cross-cutting relationship opportunities (Blau 1977, McPherson 2004). Furthermore, a dearth of similar, co-located peers may increase the geographic range of an individual's relationship partners as he/she search across greater geographic space for homophilous relationships (Stouffer 1940). Moreover, this expanded search may increase the overall set of relationships available to co-located individuals through referral systems. For these reasons, a high spatial density of individuals who are diverse along social dimensions may result in broad networks rich in structural holes.

For scientists, the most prominent social dimension is expertise. A recent stream of research emphasizes the benefits of non-overlapping expertise, or recombining existing knowledge pools (Simonton 1994). Combining separate bodies of knowledge can provide novel perspectives (Janis 1971), better feedback, as well as enhanced creativity (Fleming, Mingo and Chen 2007). In an innovation context, the benefits to relationships with very high expertise overlap are likely to be outweighed by relationships with more distant partners. This argument closely parallels the rationale underlying brokerage positions, which emphasize the benefits of forming relationships between unconnected parties. However, significant social distance between communication partners often leads to miscommunication, or worse, accidental misdirection. Thus, relationships with individuals whose underlying expertise is far afield from an individual's focal interest are unlikely to be beneficial. Consistent with the theory of recombinant invention, most combinations fare poorly (Fleming 2001). Only a minority of combinations, which are both novel and relevant, are ultimately useful. In recognition of these concerns, the policy of many knowledge-production organizations is to co-localize domains of expertise to the greatest extent possible. Thus, I hypothesize:

H2: Individuals with a higher spatial density of non-workgroup members have a network rich in structural holes.

Hypotheses 1 and 2 form the crux of a spatial ecological perspective (summarized in Figure 1). Co-located individuals with high social overlap (i.e., homogeneous clusters) yield an array of strong bonds with one another through homophilous and propinquitous relationships (lower-right quadrant). These individuals comprise the archetype of a cohesive network. They turn inward, forming strong, redundant relationships with one another (McPherson Smith-Lovin, and Cook 2001). Moreover, cohesive individuals compete with one another for external attention, further inhibiting external relationships. As social heterogeneity increases, with no decrease in spatial density, individuals form spatially proximate relationships with a greater diversity of individuals (lower-left quadrant). These relationships, which are propinquitous but not homophilous, facilitate the formation of broad networks rich in structural holes. Physical proximity mitigates both the uncertainty and the costs of forming and maintaining socially diverse relationships. As spatial density decreases between socially similar individuals, each individual is more distinctive, increasing each person's share of external attention (upper-right quadrant). As propinquitous ties weaken, individuals must roam farther afield for form relationships with similar partners, further expanding an individual's network. Lastly, populations that are diffuse along both social and spatial dimensions have low likelihoods of forming relationships (upper-left quadrant). Consistent with an ecological framework, overlap along a multitude of dimensions results in high competitive pressures. Diversity along either social or spatial dimensions can mitigate this competition, enabling an individual to form networks rich in structural holes.

Of central importance to a spatial ecological framework is the ambiguity of co-location effects. As individuals move towards higher spatial density positions (top row to bottom row), they may increase both their likelihood of diverse (left column) and redundant (right column) relationships. Thus, a consideration of spatial density without examining the social attributes of co-located individuals is potentially misleading. Only by separating spatial density along social dimensions can we disentangle the ecological determinants of networks rich in structural holes.

IV. Setting and Data

IVa. Context. To test these hypotheses, I have collected data at a large, first-generation biotechnology firm (hereafter, "BTCO"). As an exemplar of science-based knowledge production, the biopharmaceutical industry has been a fertile setting for a diverse array of studies, including research on firm capabilities and internal organization (Henderson and Cockburn 1996), inter-organizational networks (Powell, Koput and Smith-Doerr 1996), as well as the social construction of knowledge (Knorr-Cetina 1999; Latour and Woolgar 1986). Biotechnology is an industry characterized by multi-disciplinary research, radical technological change, and rapid growth (Henderson, Orsenigo, Pisano 1999).

The composition of human capital and expertise within a biotechnology firm provides a number of attractive attributes. With a historical dedication to in-house research, BTCO's research division is currently composed of a large number of scientists. The mandate of these scientists, distinct from development, is to conduct basic and applied research to feed molecules into the drug development pipeline. The goal of the employees in my dataset is to produce novel science in a for-profit setting.

Although limited to one setting, exploring ecological effects within one firm has a number of attractions. First, I can identify the complete population of firm employees, as well as measure the spatial and social distances between these individuals in detail. Close cooperation with one firm allows the collection of an extraordinarily rich dataset, including spatial distances from office blueprints, social distances from both formal organizational charts and scientific publications, as well as complete network data using electronic mail logs. The ability to both collect fine-grained data along geographic and social (i.e., scientific) dimensions and to map these dimensions onto network data suggest that a biotechnology firm may be a rich setting to explore spatial ecological effects. Moreover, complete human resources data allow the inclusion of a detailed set of controls as well as performance outcomes (proxied by compensation). Lastly, I am able to construct not only a multi-dimensional map of the ecological landscape for a discrete population, but equally important, a complete one.

IVb. Expertise. A spatial ecological framework suggests that social proximity must be overlaid on spatial proximity to yield meaningful results (Figure 1). To examine the expertise overlap between

BTCO employees, I used two methodologies: formal organization and the content of scientific publications.

In concordance with BTCO's (and the biotechnology industry's) historical origins in university laboratories, BTCO has modeled their culture and organizational structures and policies after an academic biology department. The formal chart of BTCO research scientists closely mirrors the flat organizational structure of a university biology department. Research scientists are organized broadly into Divisions, each comprised of oncologists, immunologists, neurobiologists, molecular biologists, or small-molecule chemists (Figure 2). These Divisions constitute the major disciplines where BTCO has chosen to focus its research efforts. Importantly, employees in each Division have distinguishing techniques and logics, their own language and norms, and publish in different venues. As a result, there are significant differences in expertise across Divisional boundaries. Each division is further divided across discrete domains of expertise. The fundamental organizational unit is a laboratory, led (and named after) individual scientists. Individual laboratory heads are hired on a case-wise basis for their unique skills and expertise. As a result, BTCO research is an amalgamation of distinct expertise domains, with expertise overlap closely mirroring distance across the formal organizational chart (see Figure 2).

BTCO laboratories are composed of a single laboratory head and a limited number of research associates, or laboratory members. Just as affiliation with a university professor's workgoup is the single most salient characteristic of a university scientist, BTCO research associates identify themselves primarily through their laboratory affiliations. Each associate works on one or more distinct projects, but all laboratory members draw from core interests and techniques that are specific to each laboratory. Although each associate has the flexibility to communicate and seek direction from whomever they choose, they do not have authority over the strategic direction of their respective projects: associates work on projects designated and overseen by the laboratory head. At the same time, as front-line workers associates retain significant autonomy with regards to their daily workflows, the synthesis and interpretation of results, and their own communication patterns. Embedded in fundamental units, co-laboratory members have significantly overlapping skills. At a minimum, each laboratory member should be cognizant of each co-members skills and overall research agenda. For the purposes of this paper, and with evidenced provided in latter sections, I consider co-laboratory members to be “equivalents”.

I have suggested that BTCO’s formal organizational chart closely mirrors the expertise overlap between each laboratory. To objectively measure the expertise overlap between laboratories, I draw upon data external to BTCO: the scientific publication record of each laboratory head. Towards these ends, I have hand-collected the complete publication history of each laboratory head throughout their career (Azoulay, Stellman and Zivin 2006). As each laboratory head has a record of publishing that spans their graduate education, post-doctoral appointment, and tenure at BTCO, I am able to identify a deep historical record of both the quality and content of scientific output, even for recent hires. Associated with these publications are Medical Subject Heading (MeSH) keywords, expert-curated keywords comprising the National Library of Medicine’s “controlled vocabulary thesaurus.” Importantly, these keywords are assigned to each publication not by the authors themselves, but by employees of the National Library of Medicine. In 2009, there were 25,186 keywords to index journal articles in Medline (Azoulay, Liu and Stuart 2009). Through a comparison of publication-associated keywords, I can objectively construct a continuous measure of each laboratory’s proximity to one another across scientific space.

Lastly, BTCO has provided me with extensive human resource data on all BTCO research employees, including educational background, socio-demographic variables, as well as compensation data. This data also encompasses BTCO’s reporting structure over time.

IVc. Spatial Architecture. BTCO scientists are allocated workspace across eight co-located buildings and 15 building-floors. To explore the (relative) spatial locations of each BTCO employee, I was provided with blueprints of each BTCO building-floor, as well as office assignments over time. As individual office-locations are relatively static, I was provided office assignments in approximately six-month spells. Office location coordinates were derived from blueprints to allow interpersonal distance measurements (in meters). I estimate the error in spatial location to be less than a meter. These data fulfill a number of requirements necessary to explore an individual's position within the internal spatial ecology of an organization. First, there must be fine-grained and complete spatial information on all individuals in the dataset. Second, these path-dependent spatial positions do not vary over time. Lastly, an individual's spatial positions should be randomly assigned within the organization. Although my setting meets the first two criteria, the third (i.e., random assignment) is not met. To the contrary, an individual's workspace location is explicitly designed to mirror the expertise (i.e., formal) organization of BTCO Research. With this caveat in mind, spatial positions within BTCO are the result of "bureaucratic" assignment coupled with a significant number or organizational constraints. I delve into these restrictions in detail to give the reader a greater understanding of the policies and constraints that underlie BTCO's spatial topography.

To the greatest extent possible, BTCO's spatial layout mirrors the organizational chart, following on the belief that co-locating scientists with similar research interests results in greater sharing of knowledge and tools. Consistent with this policy, laboratories members are assigned workspace close to one another, and divisions are assigned to the same, or nearby, building-floors to the greatest extent possible. Although a policy of assigning spatial positions to mirror the expertise layout of the organization appears straightforward, a significant number of current, historical, and infrastructure constraints impede strict parallels between spatial and expertise layouts.

BTCO's policy is to co-locate laboratory (i.e., workgroup) members to the greatest extent possible. As a result, 93% of laboratories are co-located on the same building-floor. BTCO has two heuristics with regards to the assignment of individual office space. First, new employees are given the closest available office to their supervisor, consistent with their policy of minimizing the distance between laboratory members. Although a simple heuristic, gaps in employee departures and arrivals, coupled with minimal spatial slack introduce variation in an individual's laboratory proximity (for a distribution, see Figure 4A). Second, employees are not allowed to re-locate offices, creating idiosyncrasies in each individual's spatial position over time. In the words of the central facilities planner: "We don't move them around as you might think… … Just because a seat opens up closer to the lab, we don't move them." Taken together, these two policies inhibit lobbying for specific office locations in the workplace.

Scientists need a place to work and the spatial distribution of these work locations are organizational choices. Unlike other knowledge production settings such as design firms or management consulting firms, the internal topography intrinsic to a biotechnology firm places severe restrictions on the number of potential spatial configurations. Although BTCO's internal environment, like almost all other organizations, is planned to the greatest extent possible, idiosyncratic elements remain.

The most prominent restriction is the proportion of square footage that must be set aside for laboratory space. This "wet" space is used for such basic, noxious tasks as the construction of genetically-modified organisms, radioactive labeling of proteins, and the use of carcinogens to track DNA. As well as general laboratory space, all biology laboratories are supplemented by warm-rooms (to speed bacteria growth), cold-rooms (to slow down chemical reactions), as well as dark-rooms. Furthermore, space is dedicated to housing large core equipment, such as centrifuges and microscopes, as well as storage space for glassware and chemicals (see Figure 3 for the layout of a building-floor). A major implication of these extensive, specialized work-areas is to limit the number of employees on each building-floor. Thus, it is plausible that BTCO scientists are familiar with all other individuals co-located onto the same building-floor.

Second, offices are clustered within floors, resulting in strong discontinuities in inter-office distances. Figure 4B presents the bi-modal distribution of interpersonal distances within all building-floors. Within a cluster, the median distance is ~8 meters. Between clusters, this spatial distance rises substantially (see Figure 4B). Given the environmental constraints on laboratory space, one might expect that office configurations are more malleable than laboratory configurations. However, the need to co-locate an individual's "wet" and "dry" spatial allocations binds these two configurations tightly together.

A second constraint on BTCO's spatial configurations is the requirement of specialized facilities by specific sub-disciplines of biology. For example, individuals who use model organisms (mouse, etc.) must have laboratory space in close proximity to the vivarium (Rader 2004). The distance that animals travel from the core vivarium to the laboratory workspace must be minimized. At the other end of the experimental spectrum, medicinal chemists are anchored to chemical fume hoods, specialized exhaust systems that draw hazardous chemicals away from common areas. Not only must chemists be co-located with this hardware, fume hoods require significant ductwork that exhaust to chemical "scrubbers" on the rooftop. As a result, chemists are almost always located on the top floors of a building. As both vivariums and fume hoods are fixed in the spatial architecture (it is estimated that the de novo installation of 24 chemical fume hoods into an existing building cost $10-11 million), scientists who are dependent on these resources are anchored in the spatial topography of the firm.

IVd. Network Data. To explore communication patterns, I obtained complete daily electronic mail log files for all BTCO research employees. I have taken two steps to insure the privacy of company employees. First, before transferring the email logs to us, BTCO's IT staff stripped the subject headings and email content from all files. Second, in constructing the dataset I analyze, the company assisted in replacing all names with hashed identification numbers.

I was assured by multiple parties that BTCO is an "email place" and that a great deal of the research division's business is conducted over email exchanges on the company's servers. Given the highly visual (i.e. microscope slide pictures) composition of knowledge flows in the life sciences, it is likely that communication partners exchange one or more emails at some point in time. Although the nature of information being transferred through electronic communications (i.e. jokes, data, golf-dates) has been removed, electronic mail presents a lower bound on communication flows. Not only are the costs to communication low relative to face-to-face communication, BTCO has a policy of standardizing email addresses (lastname.firstname@BTCO.com), further easing communication patterns.

Lastly, electronic communications allow the ready identification of the full set of both potential and realized communication patterns within the dataset, a task that is difficult with other data collection methods. All emails originated from a unique sender. I kept only emails with unique (i.e., not multiple or distributions list) recipients, which comprised 61.2% of all sent emails. For descriptive statistics, I used email from the second week of November 2008, concurrent with my office-loading data. For regression analysis, I used email from the second week of February 2009. Each correspondence was considered a "tie" between unique dyads, and I generated a symmetric adjacency matrix to capture communication patterns.

Furthermore, I generate measures of network density, which reflect the proportion of potential communication patterns that are observed (equation 1.).

ND = $\frac{L_{ij}}{n(n-1)/2}$; n = # of individuals; L_{ij} = Total # of observed communication dyads; (eqn.1)

Network density is a simple, descriptive measure that reflects the probability of communication between any two individuals (Wasserman and Faust 1994). To illustrate spatial ecological effects on communication patterns, I generate network density measures for discrete subsets of BTCO employees: across spatial partitions (i.e., same floor, same building, etc.), expertise partitions (i.e., same lab, same department, same division, etc.) and both (see Results).

V. Variable Construction.

Va. Dependent Variables. This paper examines an individual's ability to form relationships that cross an organization's internal structure (i.e., a brokered network rich in structural holes), observed using electronic mail logs from February 2009. I adopt Burt's (1992) measure of network constraint, which quantifies the extent to which an individual's network does *not* span structural holes. Network constraint measures the extent to which an individual *i* is invested (both directly and indirectly) in communication partner *j*. A network constraint variable depends upon three qualities: size, density, and hierarchy. Network size is the number of an individual's communication partners. Density is the extent to which communication partners communication with one another, and hierarchy measures if the individual's share information indirectly through a central (hierarchical) contact. Network constraint is generated as in equation 2:

$$\text{Network constraint}_i = \sum_j \left(p_{ij} + \sum_q p_{iq}p_{qj}\right)^2, q{\neq}i, j; \qquad \text{(eqn.2)}$$

Where p_{ij} captures individual i's investment in individual j, and $p_{iq}p_{qj}$ captures the relationship between q (another of i's contacts that is not j) and j. This measure is summed across all of individual i's contacts j (Burt 1992, pp. 55). Network constraint is widely distributed, ranging from .051 to 1. I present the distribution of network constraint in Figure 5, which is positively skewed for the dataset. I use the log of network constraint as my key dependent variable.

Appendix 1 examines correlations between network constraint and performance evaluations, as measured through year-end discretionary bonus. A strong negative correlation between network constraint and discretionary bonus illustrates the relevance of brokered networks for this dataset, consistent with a number of prior studies (Burt 1992).

Vb. Explanatory Variables. To generate measures that reflect an individual's position in BTCO's spatial topography is a more complex task. As illustrated in detail above, BTCO's spatial distributions closely mirror the expertise overlap of the organization. Laboratories, which have the closest expertise, are co-located to the greatest extent possible. Furthermore, members of the same division are also co- located, both due to infrastructure constraints as well as to foster communication and collaboration within the division. Lastly, BTCO's research division is spread across 8 buildings and 15 building-floors. Ex ante, it is not immediately clear how to account for a vertical dimension (within-building, across floors). In recognition of a fundamental difference between vertical distance and horizontal distance, Festinger, Schachter, Back (1950) decomposed the two dimensions into separate variable. The use of separate same-building and same-floor variables is mirrored in Marmaros & Sacerdote's (2006) exploration of friendship ties in Dartmouth College residence halls.

To reflect an individual's position in BTCO's spatial topography, I consider only each individual's co-located peers on the same building-floor for three reasons. First, laboratories rarely span building-floors. Consistent with BTCO's policies of co-locating laboratories, 93% (65 out of 70) laboratories in the sample have all members on one building-floor. Second, building-floors are primarily composed of employees from one division (i.e. immunology, neurobiology, chemistry, etc.) Ten out of fifteen (66%) building-floors have employees from one sub-discipline of biology.

Third, analysis of BTCO communication flows suggests that the bulk of electronic communication is between correspondents on the same building-floor. Furthermore, the likelihood of communication between correspondents in the *same* building, but *different* floors, is no higher than the likelihood between individuals in different buildings (see Results section). Thus, it is unlikely that local spatial topography has large effects beyond the environment of each individual's building-floor. Lastly, this assumption greatly simplifies the construction of spatial ecological variables.

I measure inter-personal distance using the Pythagorean Theorem ($x^2 + y^2 = z^2$), disregarding the role of spatial partitions (i.e., walls and doors). Following Sorenson and Audia (2000), I sum across the inverse of all interpersonal distances to generate a within-floor measure of spatial density for each individual *i* (eqn. 3). The distance from individual *i* and *j* is in the denominator. As a result, the measure of spatial density for each individual *i* increases as the individual has a greater number of spatially proximate employees. I first generate a global spatial density measure, with *j* being all BTCO employees on *i*'s building-floor. I make no predictions about the association of this measure with network constraint.

Global Spatial Density$_i$ = $\sum_i^j \frac{1}{\omega_{ij}}$; ω_{ij} = meters between *i* and *j*; $i \neq j$; *i* & *j* on same building-floor; (eqn 3)

In contrast to this global spatial density measure, spatial ecology emphasizes that spatial density must be considered in concert with social dimensions. To examine the effects of expertise overlap between co-located individuals, I first defer to the organization itself. I restrict the population of co-located peers, *j*, along formal organizational lines. Members of the same laboratory have significant overlap in expertise. To measure an individual's spatial proximity to their workgroup, I restrict the construction of an individual's spatial density measure to employees j in the *same* laboratory as individual i (eqn. 4) . I call this measure laboratory density, or LD_i:

Laboratory Spatial Density$_i$ = $\sum_i^j \frac{1}{\omega_{ij}}$; ω_{ij} = meters between i and j; $i \neq j$; i & j on same building-floor; i & j in the same laboratory; (eqn 4)

As LD_i rises, the spatial concentration of functional-equivalents also increases. Hypothesis 1 suggests that increasing LD_i will constrain a network.

To examine the effects of spatial density for non-laboratory members, I further restrict an individual's spatial density measure to employees j who are in *different* laboratories to individual i. Following the global spatial density and laboratory spatial density measures described above, I first generate a non-laboratory density measure, where i & j are in different laboratories. I further parse this non-laboratory density measure to a) i & j are in different laboratories but in the same division, and b) i & j are in different laboratories and different divisions. As individuals within the same division share a common disciplinary background, I propose that there is a greater likelihood of communication between members of the same division. Hypothesis 2 suggests that increasing non-lab spatial density will foster a network rich in structural holes.

A second, corroborating method to examine the expertise overlap between co-located individuals is to defer to the nature of the scientific output of each laboratory. For each laboratory (head), I have collected their published scientific output by hand. Associated with these publications are MeSH keywords, a "controlled vocabulary thesaurus" that is assigned to each publication by the National Library of Medicine, not the authors (Azoulay, Ding and Stuart 2009). In 2008, there were 25,186 keywords to index journal articles in Medline. Between each laboratory i and j, I generate a symmetric measure of scientific overlap (α): the count of keywords in common between labs i and j divided by the sum of total keywords for lab i and j over the past 5 years. I then use this overlap score to parse the population into those individuals j which are scientifically proximate to individual i (α overlap is in the 50^{th}-100^{th} percentile) and scientifically distant (α overlap is in the 0-50^{th} percentile), where i and j are in different laboratories. I generate spatial density for each of these populations following eqn. 3 above. As communication is more likely to occur between scientifically proximate populations than distant ones, I suggest that a high spatial density of scientifically proximate increases the prevalence of structural holes.

Vc. Control Variables. For all models, I include variables to account for laboratory size, as well as the number of employees on each building-floor. I also include an indicator for research associates in laboratories headed by senior executives at BTCO. To control for heterogeneity across divisions and salary-grades (i.e., research associate, senior research associate, etc.), I include dummies for these variables in all models. Lastly, I include a number of controls for each research associate's human capital. An individual's human capital can be separated into general- and organization-specific human capital. To reflect an individual's general knowledge, including the ability to access technical literature, I use indicator variables for each individual's highest educational degree (BA, MA, and PhD). I did not separate individuals who had post-doctoral appointments from those who did not. To proxy for an individual's organization-specific human capital, I include tenure quartiles indicators for model flexibility.

VI. Methodology.

VIa. Sample. To examine spatial ecological effects within BTCO's internal research organization, it is important to analyze effects across comparable individuals: a significant empirical challenge when individuals are recruited for their specific skill sets and expertise. To overcome this issue, I restrict empirical analysis to research associates within BTCO. These individuals are front-line scientific workers carrying out experiments on a daily basis. They are hired through the human resources division, are not managers, and have limited strategic autonomy. However, they are not without discretion. Consistent with Woodward's (1958) notion of a fluid organizational form, BTCO research associates commonly have the ability to search through the organization for the advice, resources, and information that they feel is necessary to carry on their work. This discretion is consistent with BTCO's history and culture of maintaining an academic setting in a for-profit firm. Although research associates are comparable to one another, they are not automatons on an assembly line.

VIb. Regressions. To examine correlations between spatial ecological factors and network constraint, I run regressions in log-linear form with robust standard errors clustered at the laboratory level. All regression analysis is in the cross-section. To the extent possible, I construct the dependent variable using data from subsequent time periods relative to the explanatory variables. Although I continue to collect email network data over time, the network constraint measure is largely time-invariant. Measuring network constraint over 6 consecutive 4-week windows (total of 6 months), I observed that there was significantly greater variation between (SD = 0.52) individuals than within (SD = .29). Furthermore, BTCO's internal ecology is relatively static. From March through November, 2008, the vast majority of individuals did not move offices, consistent with the bureaucratic, path-dependent process of office-allocation within a building-floor. Thus, I begin my analysis of spatial ecological effects in cross-section.

VII. Results.

VIIa. Descriptive Statistics. I begin my discussion of the results with a set of descriptive statistics. Table 1 reports characteristics of technicians in my dataset. First, the population is diverse along ethnic and gender lines. Almost 60% of the research associates are female and over 50% are of Asian (both Far and Near) descent. The typical research associate is 39 years of age, married, and has been with the firm for 6 years. Organizational tenure is skewed, with a long right-hand tail. The median research associate has been with the organization for 4 years, reflecting significant organizational growth in recent years. The longest-tenured research associate has been with the firm for 29 years. BTCO research associates are highly educated. Thirty percent of the individuals in the dataset hold a doctorate, and over half have a master's level education of some kind.

Research associates in the same laboratory are co-located. Figure 4, panel A presents a histogram to illustrate the spatial distance between technicians in the same laboratory. By comparison, research associates in different laboratories have a much broader spatial distribution within a building-floor (Figure 4 Panel B). However, a significant number of research associates have office locations spatially distant from their laboratory. Before I turn to regression results, I first present an overview of BTCO's communication patterns across spatial and social dimensions.

VIIb. Communication Patterns. I illustrate the joint effects of spatial and social distributions on communication patterns using a week's worth of electronic mail in the month of November, 2008. I generated all pair-wise potential dyads between i and j technicians at BTCO and constructed a dichotomous (i.e., 0 or 1), symmetrized adjancency matrix for all dataset members. Overall, there are 58,684 potential relationships across the full dataset, only 1,581 of which are observed. Network density is a simple descriptive statistic that represents the "saturation" of a network. I simply divide the number of observed relationships (i.e., 1,581 dyads) by the number of potential relationships (i.e., 58,684 dyads). Thus, the network density of the full dataset is 2.7%, consistent with a population of this size (see Table 2, Column 4, Row 5). To examine the association of spatial and expertise proximity with communication flows, I present network densities restricted to subsets of the population (Table 2).

The probability of communication drops off monotonically as spatial distance increases, suggesting that co-location increases the flow of communication (Table 2, column 4). Individuals co-located on the same building-floor have much higher network densities than across the overall network. Furthermore, communication flows within a building-floor cluster of office/cubicles are higher than between clusters, suggesting that micro-spatial proximity within a building-floor also structures communication flows.

There is a sharp drop-off in network density beyond a building-floor. Somewhat surprisingly, same-building/different-floor network density is not much higher than baseline. Thus, these descriptive statistics suggest that communication flows drop off precipitously once individuals are no-longer co-localized along a planar axis, consistent with prior findings (Allen 1977; Festinger, Schachter and Back 1950). However, as the expertise dimensions of BTCO Research is structured in parallel with spatial distributions, these patterns may not be solely attributable to spatial agglomeration. Instead of chance opportunities driving patterns of interaction, these results are also consistent with increases in communication patterns due to task interdependence and homophily along expertise and divisional dimensions (Hackman 2002).

When network density is mapped along expertise dimensions, the likelihood of communication also decreases monotonically as formal distance increases (Table 2, Row 5). Overall network density within a laboratory is over 50%, consistent with their characterization as cohesive workgroups. In contrast, communication patterns within a division match the baseline communication flows, while flows between divisions are well below baseline. These descriptive statistics are consistent with the idea that both formal organizational structure and expertise strongly shape communication flows and must be considered in concert with spatial proximity.

Lastly, I examine network densities across both spatial and expertise dimensions. Communication patterns within a laboratory are relatively insensitive to spatial distance (Table 2, Column 1). Although descriptive statistics across the entire dataset suggested that communication flows dropped off precipitously across building-floors (column 4), I see no evidence for a decreased probability of within-laboratory communication between building-floors, or even across buildings (column 1). Furthermore, network densities within laboratories are very high (~0.5). Although surprising, these results are consistent with the presence of other formal mechanisms, such as laboratory meetings and a common supervisor, that foster communication flows (Owen-Smith 2001) as well as strong interdependence within the laboratory.

In contrast to intra-laboratory network densities, non-laboratory communication flows are highly dependent on co-location. For same division members localized on a building-floor, I observe a slight decrease in network density across cubicle-clusters (Table 2, Column 2). I document almost a 3-fold decrease in communication for non-lab members located in different building-floors. I observe no further decrease in communication for individuals located in different buildings. Taken together, these patterns of communication suggest (1) that the spatial topography beyond a discrete building-floor does not significantly contribute to the probability of communication patterns; (2) individuals within the same laboratory/workgroup are fundamentally different from non-laboratory members; and (3) both spatial and social dimensions must be considered in concert.

VIIc. Spatial Ecology. I examine the role of spatial ecological factors in Table 3. I remind the reader that the dependent variable, network constraint, captures the *absence* of structural holes. I present a baseline regression in column 1. A negative coefficient for laboratory size suggests that larger laboratories foster broader networks, or an individual's ability to develop structural holes. The number of floor employees does not correlate with network constraint. All regressions include three quantiles of organizational tenure and indicators for each individual's highest educational degree. Alternative specifications on the human capital determinants of network constraint are presented in Appendix 2.

I explore the *overall* effect of employee density in Table 3, column 2, taking no account of a co-located employee's dimension of expertise. Spatial density is a measure of each individual's proximity to other employees. Thus, as the measure increases, individuals are positioned in more spatially "crowded" environments. In other words, global spatial density captures the extent to which an individual is proximate to the "centroid" of each building-floor. Global spatial density has no effect on network constraint (Column 2). This null result is in direct contrast to current theories which emphasize the positive effects of agglomeration. However, when spatial density is broken down across dimensions of expertise, I find strong evidence for spatial ecological effects.

Table 4, Column 3 separates an individual's spatial density into the two components that comprise Hypotheses 1 & 2: a) proximity to same workgroup members as captured by Laboratory spatial density and b) proximity to non-workgroup members as captured by non-Lab spatial density. For each individual i, the total number of alters j does not change between column 2 and column 3.

Laboratory spatial density is positively correlated with a constrained network, providing support for hypothesis 1. As individuals move from the spatial periphery towards the "centroid" of their workgroup, individuals have networks with fewer structural holes. By contrast, I observe the opposite effects for non-Lab spatial density. As individuals move towards a higher density of non-workgroup members, they have a less constrained network. Thus, I find support for hypothesis 2. Moreover, separating co-located peers into workgroup and non-workgroup members increases the model fit (column 2 & 3).

To further examine the role of expertise overlap between non-workgroup members, for each individual *i*, I separate all non-lab alters *j* into those in the same division and those in different divisions. As BTCO divisions are organized along disciplinary lines, individuals in the same division share a common language and norms, facilitating relationships. I then construct separate measures of within-floor spatial density: Same Division spatial density and Different Division spatial density (column 4). Co-location with non-workgroup members who are in the same division facilitates a network rich in structural holes. However, co-location with non-workgroup members whose expertise is farther afield (i.e., immunologists co-located with chemists) has no effect on an individual's ability to occupy brokerage positions. Taken together, column 4 suggests that some degree of overlap in expertise is necessarily to initiate and maintain relationships.

Although straightforward, using cut-points in the organization chart to represent the expertise overlap of BTCO employees may be problematic. To overcome these potential difficulties, I turn outside the organization to publication associated keywords as a measure of scientific overlap. If laboratories have more publication keywords in common, they are more proximate in scientific space and are more likely to have expertise that is relevant to one another. Importantly, these measures of expertise overlap are generated by the National Library of Medicine, not by the authors or BTCO. Paralleling the methodology above, for each individual *i*, I separate all alters *j* into those who are have high (50^{th}-100^{th} percentile) or low (0-50^{th} percentile) scientific overlap. I then construct separate measures of within-floor spatial density: High Science Overlap spatial density and Low Science Overlap spatial density. Consistent with spatial density measures derived from the organizational chart (column 4), I find that co-location with individuals who have a high degree of scientific overlap fosters a network rich in structural holes (column 5), providing support for hypothesis 2. The segregation of Non-Lab spatial density into more fine-grained measures of scientific overlap does not change the coefficient of Laboratory spatial density.

To examine the economic significance of environmental factors, I use Table 4 column 2 to predict an individual's network constraint along spatial ecological dimensions, holding the control variables at the population mean. The predicted mean network constraint is .346, comparable to an observed network constraint of .313. At different percentiles of Laboratory spatial density (while holding Non-Lab spatial density at the median), an individual's network constraint can vary between .323 & .405, or 24% of the median individual's network constraint. As Non-Lab spatial density changes (while holding Laboratory spatial density at the median), an individual's network can vary between .275 & .381, or 30% of the median individual's network constraint. I do not observe an an interaction between Laboratory and Non-Lab spatial density (data not shown).

Lastly, these effects are on par with predicted human capital effects on network constraint. Individuals with greater human capital are more desirable relationship partners, and thus have greater flexibility in structuring their network. I account for human capital using both organizational tenure (i.e., specific human capital) and highest educational degree (i.e., general human capital). Individuals with a Ph.D. have a network constraint of .294, considerably lower than the network constraint of non-Ph.D. holders .369. Individuals who have <=2 years of organizational tenure have a predicted network constraint of .400. By comparison, individuals with 3-4 years of organizational tenure have a predicted network constraint of .343, and those individuals with >4 years of organizational tenure have a predicted network constraint of .315. Thus, spatial ecological factors play as prominent a role as human capital in allowing an individual to adopt networks rich in structural holes.

VIII. Discussion and Conclusion (abridged due to space limitations).

This paper considers both individual and environmental determinants of a focal actor's ability to form expansive networks rich in structural holes. Through an ecological lens that considers geography in concert with the social attributes of co-located peers, I provide evidence that environmental factors play as prominent a role in enabling unconstrained networks as individual-level attributes. Furthermore, I argue that considerations of spatial dimensions alone are incomplete and potentially misleading. In concert with the expertise overlap of co-located peers, spatial effects are contingent: co-located peers may enable or constrain a focal actors' ability to form relationships that span an organization's internal communication structure. Consistent with an ecological perspective, extreme similarities between co-located individuals heighten competitive pressures, inhibiting an individual's ability to grow his or her network (Hannan and Freeman 1989).

The findings in this paper speak to organizational ecologists and network theorists alike (Burt 1992; Hannan and Freeman 1989). First, competition and cooperation between co-located individuals have clear parallels in niche theory, a central component of organizational ecology. Considerations of spatial and social distributions as niche dimensions is mirrored in ecological work on voluntary associations (McPherson, Popielarz and Drobnic 1992), newspapers (Carroll 1985), as well as day care organizations (Baum and Singh 1994). My focus on spatial distributions returns attention to the foundations of biological ecology through which population ecologists have drawn their inspiration (i.e. Evelyn Hutchinson). Second, this paper closely mirrors the graph theoretical concepts of network scholars. In particular, the idea of spatial co-location as a constraint, due to a high density of socially equivalent alters dovetails with the concept of structural equivalence (Burt 1992). The idea of spatial co-location as a facilitator of broad networks directly explores the antecedents of brokered network positions.

[Due to space limitations, References, Appendix 1 (Network Constraint and Discretionary Bonus), and Appendix 2 (Human Capital Determinants of Network Constraint) are available upon request]

Figure 1: Spatial Ecological Effects with Four Combinations of Social and Spatial Overlap.

<table>
<tr><td></td><td></td><td colspan="2">Social Density</td></tr>
<tr><td></td><td></td><td>Low Overlap</td><td>High Overlap</td></tr>
<tr><td rowspan="2">Spatial Density</td><td>Low Overlap</td><td>Undeveloped Networks

Low Likelihood of forming relationships.</td><td>Distinctive Networks

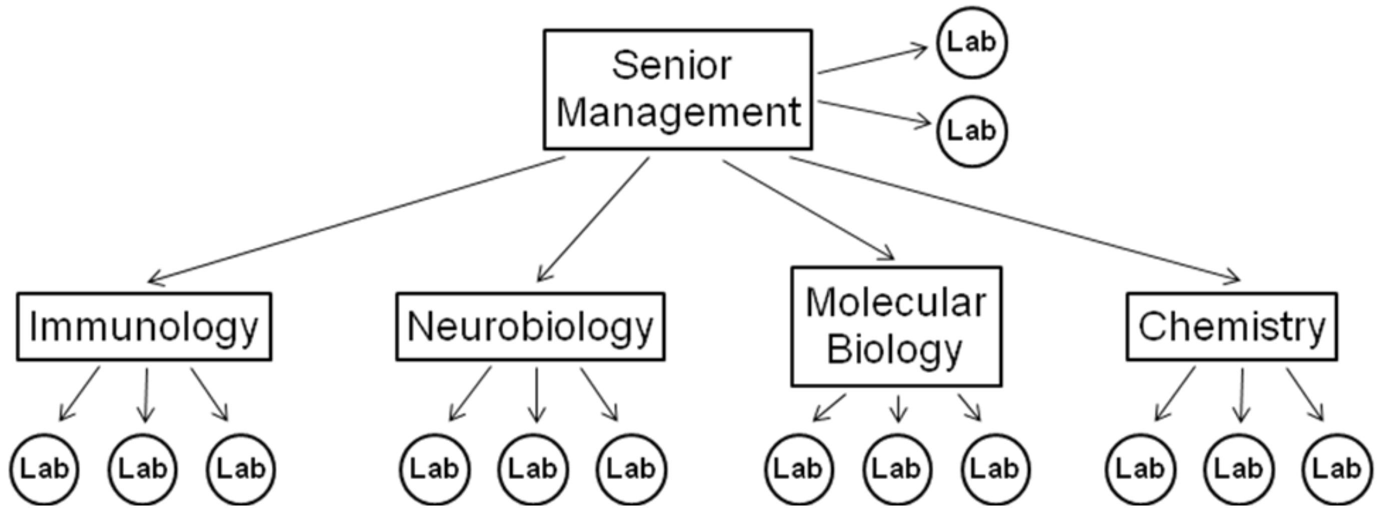
Increased share of external attention to facilitate bridging relationships</td></tr>
<tr><td>High Overlap</td><td>Varied Networks

High likelihood of forming non-redundant relationships, facilitating reach across the organization.</td><td>Cohesive Networks
Decreased share of external attention.

High likelihood of forming redundant relationships, inhibiting reach across the organization.</td></tr>
</table>

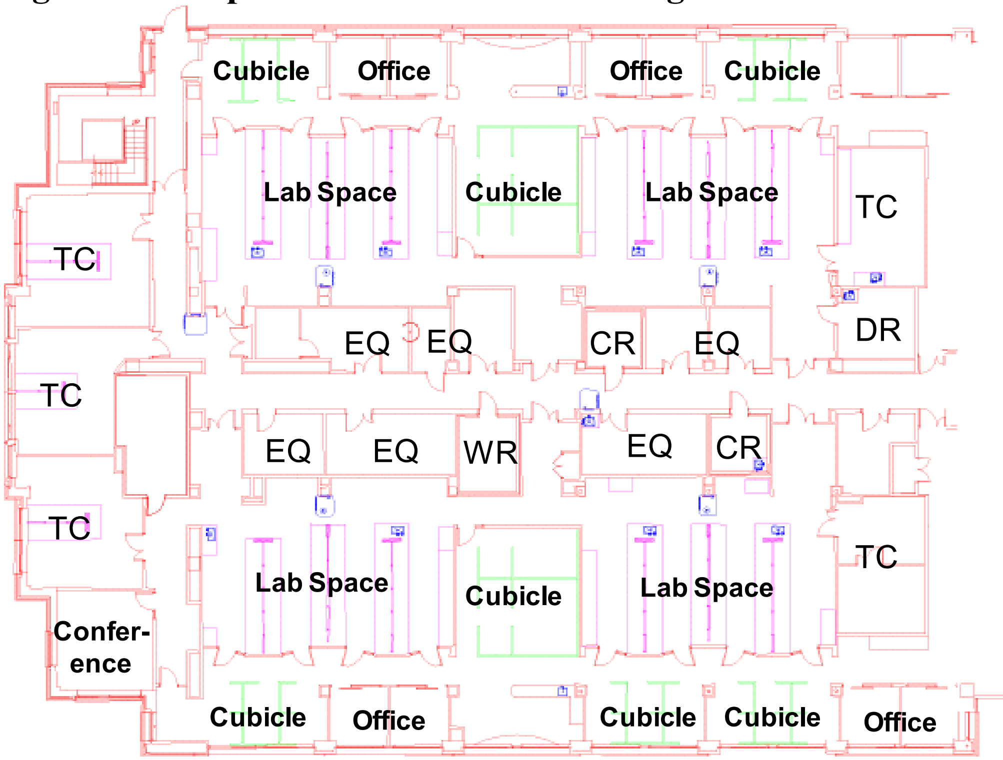
Note: Hypothesis 1 suggests that individuals more spatially distant from their social equivalents have broader, less constrained networks. In the diagram above, moving from the lower-right quadrant to the upper-right quadrant is beneficial. Hypothesis 2 suggests that individuals more spatially proximate to non-equivalent peers have broader, less constrained networks. In the diagram above, moving from the upper-left quadrant to the lower-left quadrant is beneficial.

Figure 2: Schematic of Formal Organizational Chart



Note: Schematic of the Formal Organization of BTCO. Division are indicated in squares. Laboratories are indicated with circles. The number of laboratories is representative, and does not reflect the distribution of laboratories across the organization.

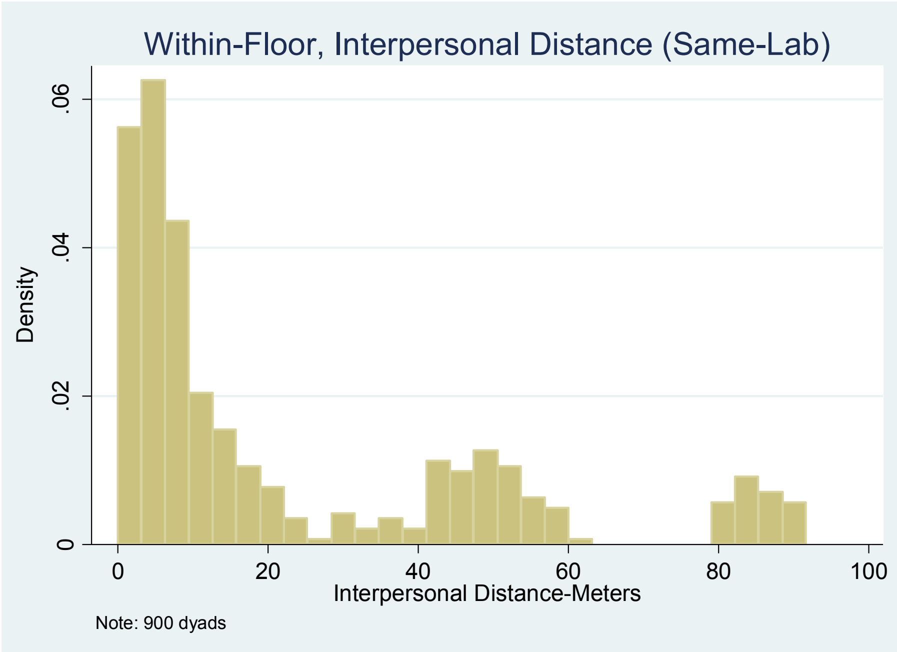
Figure 3: Blueprint of One BTCO Building-Floor



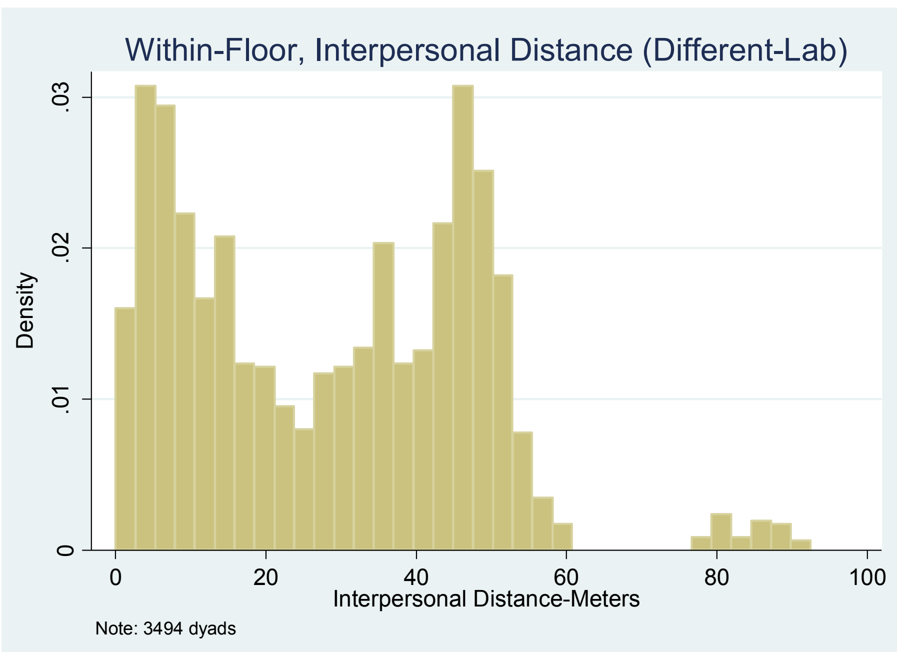
Note: Blueprint and functional schematic of one building-floor. Cubicles, offices, laboratory space and a conference room are indicated. Specialized equipment rooms, for this building-floor: tissue culture (TC) facilities, are shown. Common equipment rooms such as warm-rooms (WR), cold-rooms (CR), dar-rooms (DR), and large equipment such as centrifuges (EQ) are also indicated. This building-floor connects to other buildings through a walk-way to the right.

Figure 4: Histogram of Within-Floor Interpersonal distances

Panel A: Same-Lab distance

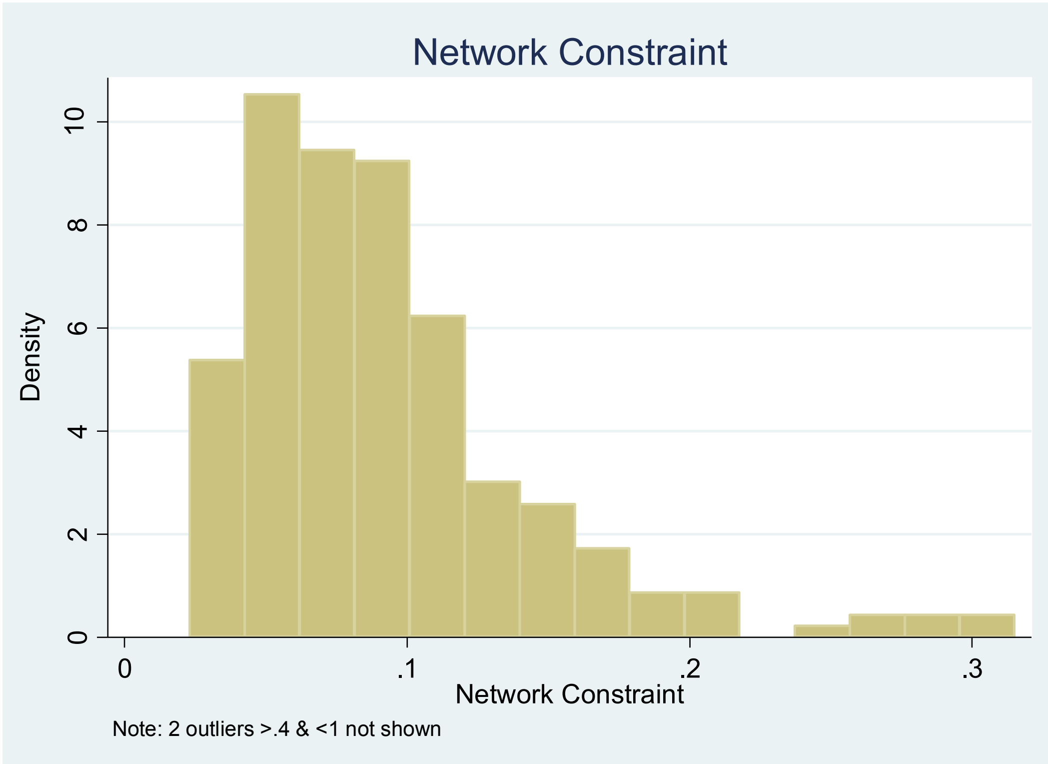


Panel B: Different-Lab Distance



Note: Individual office locations were computed from blueprints of each building-floor. Office locations were transposed into east-west and north-south coordinates and interpersonal distances were calculated using the Pythagorean Theorem, with no consideration of intervening building walls. Bi-modal distributions are consistent with the segregation of offices into localized clusters within building-floors (see Figure 2).

Figure 5: Network Constraint Distribution



Note: Network Constraint distribution was determined using electronic mail communication from January, 2009 electronic mail logs in UCInet following Burt's (1992) measure of network constraint.

Table 1: Descriptive Statistics- Research Associate Characteristics (N = 239)

	Mean	Std. Dev.	Min	Max
Male	.401	.491	0	1
Asian	.545	.499	0	1
Caucasian	.405	.492	0	1
Age	39.355	8.531	24	70
Married	.550	.499	0	1
Firm tenure	6.041	5.540	1	29
BA-highest degree	.459	.499	1	1
MA-highest degree	.236	.425	0	1
PhD-highest degree	.306	.462	0	1
Lab Head is an Executive	.062	.242	0	1
Laboratory size	7.136	5.678	1	22
Floor size (non-Lab)	23.643	7.960	4	34
Global spatial sensity	2.728	1.084	.686	6.375
Laboratory spatial density	.857	.711	.020	3.514
Non-Lab spatial density	1.871	1.175	.189	6.022
Same Division spatial density (non-Lab)	1.630	1.299	0	6.022
Different Division spatial density (non-Lab)	.242	.505	0	3.282
High Science Overlap spatial density (non-Lab)	1.266	1.040	0	4.719
Low Science Overlap spatial density (non-Lab)	.602	.810	0	4.120
Network Constraint	.341	.216	.051	1.010
Share of Discretionary Bonus	1.013	.182	0	178

Table 2: The Effect of Spatial and Social Distance on Communication
(Network Density Above/# of Potential Ties in Parentheses Below)

		(1) **Same Lab**	(2) **Different Lab/Same Division**	(3) **Different Division**	(4) **All Social Dyads**
(1)	**Same Cluster**	0.564 (621)	0.098 (1445)	0.018 (218)	0.212 (2284)
(2)	**Different Cluster/Same Floor**	0.478 (274)	0.077 (1384)	0.027 (442)	0.119 (2100)
(3)	**Different Floor/Same Building**	0.590 (222)	0.029 (1210)	0.006 (4178)	0.034 (5610)
(4)	**Different Building**	0.429 (56)	0.026 (13421)	0.007 (35093)	0.013 (48570)
(5)	**All Spatial Dyads**	0.533 (1173)	0.036 (17460)	0.007 (39931)	0.027 (58564)

Note 1: 242 unique individuals in 70 laboratories.
Note 2: Seven days of electronic mail communication were used to generate network data. One or more email correspondences, regardless of directionality were considered a network tie. Network density is the proportion of observed network ties, divided by all potential network ties. The number of potential network ties is presented in parentheses. Individuals were assigned to "clusters" on a case-by-case basis.

Table 3: Spatial Ecological Determinants of Log Network Constraint-(OLS)

	(1)	(2)	(3)	(4)	(5)
Global spatial density		-0.037			
		(0.039)			
Laboratory spatial density			0.112*	0.110*	0.106+
			(0.054)	(0.054)	(0.055)
Non-Lab spatial density			-0.080*		
			(0.040)		
Same Division spatial density (non-Lab)				-0.080*	
				(0.039)	
Different Division spatial density (non-Lab)				0.028	
				(0.062)	
High Science Overlap spatial density (non-Lab)					-.090*
					(0.041)
Low Science Overlap spatial density (non-Lab)					-.058
					(0.058)
male	-0.071	-0.073	-0.061	-0.061	-0.057
	(0.069)	(0.069)	(0.068)	(0.068)	(0.068)
lab size	-0.025**	-0.025**	-0.035**	-0.034**	-0.034**
	(0.006)	(0.006)	(0.007)	(0.007)	(0.007)
# of floor employees	-0.002	-0.000	0.002	0.002	0.002
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Constant	0.830**	-0.796**	-0.740**	-0.763**	-0.762**
	(0.210)	(0.229)	(0.218)	(0.213)	(0.212)
Observations	239	239	239	239	239
R-squared	0.13	0.14	0.17	0.18	0.17
Log-likelihood	-179	-178	-173	-172	-173
# of lab clusters	70	70	70	70	70

Note: Estimates are displayed as raw coefficients. All models include, but do not show, indicators for functional-area and salary-band, as well as an indicator if the laboratory head is a high-level executive. Three quartiles of organizational tenure, as well as highest educational degree indicators are included, but not shown. All spatial density measures are computed for individuals on the same building-floor. Spatial density is calculated using the equation $\sum_i^j \frac{1}{\omega_{ij}}$; where ω_{ij} is the interpersonal distance (in meters) between *i* and *j*. Global spatial density considers the proximity of all alters *j*. Laboratory spatial density only includes *j* if $lab_j = lab_i$. Non-lab spatial density only includes *j* if $lab_j \neq lab_i$. Same division only includes *j* if $lab_j \neq lab_i$ and $division_j = division_i$. Different division only includes *j* if $lab_j \neq lab_i$ and $division_j \neq division_i$. High Science Overlap spatial density only includes *j* if $lab_j \neq lab_i$ and *i* and *j* are in the 50-100th percentile of scientific (MeSH keyword) overlap within the dataset. Low Science Overlap spatial density only includes *j* if $lab_j \neq lab_i$ and *i* and *j* are in the 0-50th percentile of scientific (MeSH keyword) overlap within the dataset. Robust standard errors, clustered by laboratory, are in parentheses below; + significant at 10%, *significant at 5%; **significant at 1%.