Administrative Science Quarterly
57 (3)407–452

Reprints and permissions:
sagepub.com/
journalsPermissions.nav
DOI: 10.1177/0001839212461141
asq.sagepub.com

Organizational Misfits and the Origins of Brokerage in Intrafirm Networks

Adam M. Kleinbaum[1]

Abstract

To extend research on the effects of networks for career outcomes, this paper examines how career processes shape network structure. I hypothesize that brokerage results from two distinct mechanisms: links with former co-workers and with friends of friends accumulated as careers unfold. Furthermore, I hypothesize that ''organizational misfits''—people who followed career trajectories that are atypical in their organization—will have access to more valuable brokerage opportunities than those whose careers followed more conventional paths. I tested this hypothesis with career history data recorded longitudinally for 30,000 employees in a large information technology firm over six years and sequence-analyzed to measure individual-level fit with typical career paths in the organization. Network position was measured using a unique data set of over 250 million electronic mail messages. Empirical results support the hypotheses that diverse, and especially atypical, careers have an effect on brokerage through mechanisms rooted in social capital, even when accounting for endogeneity between networks and mobility. In theorizing about misfit from prototypical patterns, this paper offers a new, theory-driven application of sequence-analytic methods as well as a novel measure of brokerage based on interactions across observable boundaries, a complement to the structural constraint measure based on interactions across holes in social structure.

Keywords: social networks, social capital, career mobility, brokerage, identification

If the social networks literature has taught us anything, it is that brokers do better. In virtually every domain, individual action is embedded in networks of social relations, and the structure of those social relations affects outcomes (Granovetter, 1985). In recent years, one structure in particular has been a focus of theoretical attention: brokerage. A broker is one who connects people and groups that are otherwise disconnected in the informal network structure,

[1] Tuck School of Business, Dartmouth College

one who spans structural holes in the social fabric of an organization (Burt, 1992). Brokerage has been argued to be a source of advantage that accrues from greater visibility: by virtue of their sparse, far-reaching networks, brokers have access to a broader range of information and receive that information earlier than others do. Consequently, they can envision more recombinant potential than others. As a result of this advantage, brokers tend to attain greater career rewards than their otherwise-similar counterparts (Burt, 2005). In recent years, a mountain of empirical evidence has accumulated showing the individual-level benefits of brokerage positions. Brokerage has been linked with myriad positive outcomes, including faster promotion (Brass, 1984; Burt, 1992), more creative output (Fleming, Mingo, and Chen, 2007), larger variable compensation (Burt, 1997), and more favorable performance evaluation (Burt, 2004). But the benefits of brokerage do not accrue only to the broker. Theory and evidence suggest that organizations also benefit from an internal social structure rich in brokerage, as informal networks facilitate the flow of knowledge and, consequently, innovation (e.g., Obstfeld, 2005; Kleinbaum and Tushman, 2007; Alcácer and Zhao, 2012). Simply put, brokerage in intrafirm networks benefits both the broker and the organization.

But although the consequences of brokerage in organizational networks are well established, evidence about its organizational antecedents is scant at best. Several streams of research look to dispositional effects as the source of brokerage. For example, evidence suggests that people with self-monitoring personality types are more likely to be brokers (Oh and Kilduff, 2008) and to remain brokers (Sasovova et al., 2010). Other research posits individual differences in the behavioral predisposition to maintain brokerage positions (Obstfeld, 2005) or, conversely, to promote (and to perceive) closure in social networks (Flynn, Reagans, and Guillory, 2010). Other scholars have focused on differences in brokerage that result from training, rather than from disposition. For example, Burt and Ronchi (2007) argued that people can be taught to network more strategically. Though disposition and training have important effects, however, neither of these approaches addresses the emergent organizational antecedents to network structure. Still other research has shown correlations between rough demographic categories and network structure (Han, 1996; Kleinbaum, Stuart, and Tushman, 2009); while this work describes a firm's population of information brokers, it offers neither insight into the mechanisms of the origins of brokerage nor the ability to account for endogeneity between careers and network structures. A burgeoning list of practitioner-oriented books (e.g., Hoffman and Casnocha, 2012) purports to advise managers on building effective networks, but with little basis in empirical research.

More research on the origins of networks exists at the interfirm level, for which alliance and board interlock data are available. Some of the earliest work in this tradition extends notions of tie formation to the structuring of networks. For example, Gulati (1995) explored the conditions under which a firm will form alliances with new partners versus prior partners. Subsequent work has examined how aspects of the environment, such as market position (Stuart, 1998) or aspects of the social setting that brings partners together (Sorenson and Stuart, 2008), affect tie formation. More recently, research has begun to move beyond the dyad to examine the antecedents of network structures more broadly, including small-world structures (Baum, Shipilov, and Rowley, 2003) or structural holes (Zaheer and Soda, 2009). And without doubt, some of the mechanisms that affect the formation of interorganizational networks apply equally to intrafirm social networks. For example, work by Sorenson and Stuart (2001) on the negative effect of geographic distance on tie formation between venture capital firms directly echoes earlier work showing a propinquity effect on interpersonal tie formation (e.g., Allen, 1977). Similarly, theories of relational and structural embeddedness, developed in a study of alliance formation (Gulati and Gargiulo, 1999), shed light on the formation and evolution of social networks.

What has been overlooked, however, is the way in which social networks form as a consequence of career processes. This is a particularly important theoretical gap because careers—defined as a sequence of jobs occupied by an individual over time (Spilerman, 1977)—are an inherently longitudinal construct (Hall, 2002) and therefore have the potential to significantly inform theory on how networks form over time, a topic of substantial interest recently (Ahuja, Soda, and Zaheer, 2012). This omission is surprising because so much research has examined the effects of networks on career attainment outcomes (e.g., Granovetter, 1974; Brass, 1984; Podolny and Baron, 1997; Burt, 2005). But few would dispute that causality also flows in the opposite direction: as individuals move from position to position through their careers, the task requirements of each new role impose changes on the structure of their networks.

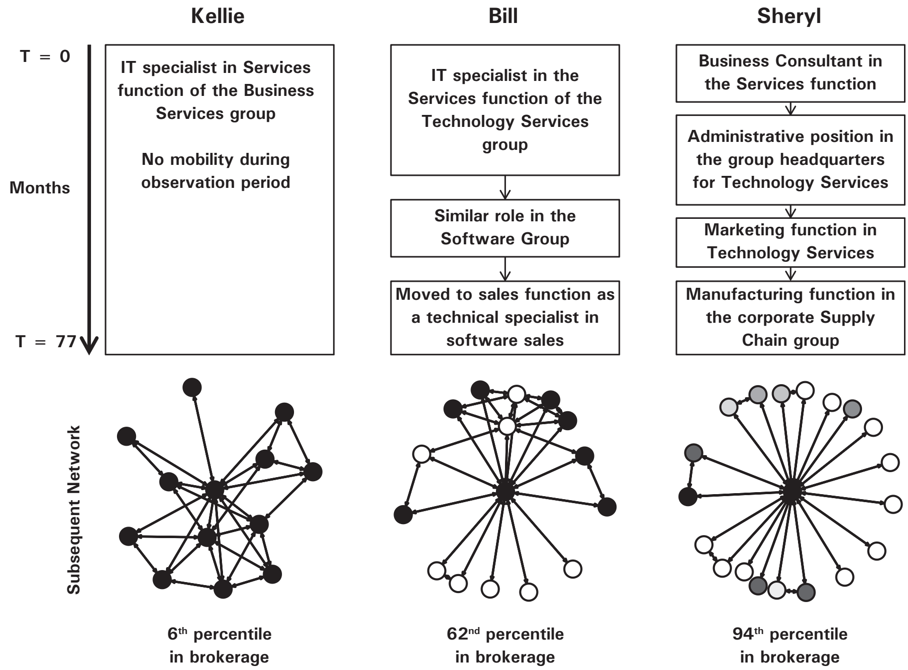
Building on this foundation, the present paper examines how career processes give rise to brokerage in intraorganizational social networks. I theorize that people with diverse career histories are more likely to be brokers because their mobility has facilitated interaction with a larger set of now-distant colleagues. Second, beyond simple measures of career diversity, the specific sequence of positions held will matter: the people most likely to be brokers in the communication network are ''organizational misfits''—those individuals whose career paths are atypical for the organization—because such people are more likely to have networks that connect parts of the organization that are rarely linked. I test the hypotheses with career histories of 30,000 employees from 2000 to 2008 at a large information technology and electronics company that I refer to as BigCo.

MOBILITY AND NETWORK STRUCTURE

The career histories of three employees at BigCo illustrate how variations in career trajectory can influence social network structure. To protect the privacy of its employees, BigCo did not provide me with their real names, nor did it provide some other specifics, such as the precise location of its offices. The names in these examples are pseudonyms, but the descriptions of the people are real and are as detailed as available data and the privacy of BigCo permits. Figure 1 provides a summary. The focal person is at the center of each network diagram; circles, representing his or her contacts, are shaded according to job function.

Kellie has the job title of information technology specialist. She has nearly 20 years of experience at BigCo and held the same position throughout the observation period. She works as a business consultant out of a medium-sized office in a Louisiana city, where a few hundred other employees are based. She occupies a middle manager position and, as of 2006, she had not been promoted since 2000. Although her client mix has changed slightly over the years

Figure 1. The career trajectories of three BigCo employees during the observation period and their subsequent network structures.*



*** The focal person is at the center of each network diagram; shading indicates each person's job function.**

as she has increased her specialization on financial services clients, her role has changed very little. Correspondingly, Kellie's network at the end of the observation period is relatively focused. All of her contacts are in her own job function and nearly all are in her own business unit. She has some contacts in other offices, but mostly within her home state. And most of her contacts are also in direct contact with each other. Kellie's network is highly constrained and largely predicted by her position in the formal structure of BigCo's organization. Compared with other members of the sample, Kellie's network ranks in just the 6th percentile of brokerage.

Whereas Kellie's career at BigCo has been very stable, with no mobility at all in at least six and a half years, Bill, a software salesperson in Georgia, has moved around. In 2000, he held a job similar to Kellie's: he was a consultant of middle manager rank and 20 years of tenure, also with the job title of IT specialist in a consulting business unit. But in 2004, Bill moved to the software business group, where he took an IT specialist position within the business unit producing information management software. Although he changed business units, Bill's job remained quite similar: he was a consultant, helping BigCo's clients with software implementation but focusing specifically now on his new business unit's class of products. Fourteen months later, Bill changed roles again. He stayed in the software group but moved into the sales function as a technical sales specialist. This was a natural transition, given the expertise Bill had developed in earlier roles, and one with ample precedent at BigCo. Compared with Kellie, Bill has a broader, sparser network. He has one cohesive subgroup within his network (his present workgroup), but he is also in touch with numerous people who are outside that group. Most of his contacts reside in the services or sales functions, the job functions he has spent time in recently. Just over half his contacts are in the software business unit, but several are in the consulting business from which he came, and several are in the corporate sales force. By any measure, Bill's network offers numerous brokerage opportunities: he has some access to information that resides elsewhere in the organization. Quantitatively, he is in the 62nd percentile of brokerage, and his network is more favorably structured than those of more than half of his fellow BigCo employees.

Finally, Sheryl, in 2000, also had much in common with Kellie. Sheryl was a business consultant of middle manager rank and over 20 years of tenure with BigCo. But the similarities end shortly thereafter. In 2002, Sheryl was promoted to the executive ranks into an administrative role in the technology consulting unit. A year later, she transitioned into a marketing position in that unit. In 2006, she moved into a new role in the manufacturing function of the corporate supply chain group, working in the corporate headquarters. A sequence of transitions like Sheryl's is highly unusual at BigCo; fewer than 7 percent of the members of the sample have career histories as atypical of BigCo as Sheryl's. Correspondingly, Sheryl has an extremely broad, far-reaching network, more so than that of either Kellie or Bill. Most of her contacts are outside her job function, nearly half are outside her business unit or office, and very few of them are in direct contact with each other. Like Bill, Sheryl has access to a significant amount of information that resides in remote corners of the organization. But unlike Bill, whose mobility spanned a boundary that is frequently traversed within BigCo, Sheryl is well positioned to acquire information from her network that is unlikely to be available to her group from any other source. If Bill is a broker, Sheryl is a super broker: relative to the rest of the sample, Sheryl's network structure places her in the 99th percentile of brokerage.

These anecdotal examples serve to motivate the theoretical development by illustrating two simple points. First, it is hardly surprising that Kellie, whose career was static, had a more focused, less diverse network than that of Bill or Sheryl. The mechanisms that give rise to the network benefits of a diverse career are likely to reside in both social capital effects—the set of direct and indirect ties accumulated over the course of a career—and human capital effects—such as knowledge about the organization, its structure, and its products. But it is perhaps less intuitive that Sheryl, with her unconventional career path, should have a more favorably structured network than Bill. Network advantage may result not merely from a diverse career history—defined here as prior work experience in a diverse set of job functions—but also from an atypical pattern of mobility.

Mobility within Careers

Sociological work on mobility within a career dates back to Spilerman (1977), who observed that most prior research on the sociology of work examined each job in isolation from others (reviewed in Rosenfeld, 1992). He defined a career, very simply, as a sequence of jobs held over time, with implicit dependence of one job on the previous ones. White's (1970) work on vacancy chains conceptualized mobility as an organizational phenomenon in which vacancies flow downward as individuals are promoted upward to fill them. In this sense, White also observed that jobs are fundamentally interdependent. But by focusing on the chain of vacancies, he emphasized the dependence of one actor's mobility on another's, rather than on the time-varying interdependence across an individual's career. Both of these views take classical theorists' assumption that positions logically precede the individuals who happen to occupy them (e.g., Reiley and Mooney, 1939). More recent work challenges this assumption, looking at idiosyncratic jobs (e.g., Miner, 1987; Rousseau, Ho, and Greenberg, 2006), which may be shaped by, or even created for, particular individuals.

More generally, careers have consequences for both organizations and individuals. At the organization level, research has shown, for example, that the career structure of a firm has implications for its ability to attract and retain employees (e.g., Carrell, 2007). At the individual level, there is a long history of research dating back to Doeringer and Piore (1971) showing the effects of careers on financial and non-financial attainment. Although quite a bit is known about the consequences of careers at both levels of analysis, little research has examined how the career process (Hall, 2002) affects an individual's network structure, with consequences for both the individual and the organization. In this paper, I attempt to fill this gap by examining how mobility, which determines how diverse and how typical one's career trajectory is, affects brokerage in an individual's intraorganizational network.

The Benefits of Career Diversity

There are many reasons to expect that a diverse career history would confer certain benefits on the individual. Mintzberg's (1973) study of managerial work suggested that general managers engage in a broad range of highly diverse tasks, requiring diverse experience. Job rotations, or other forms of mobility, might be expected to provide the requisite experience (Campion, Cheraskin, and Stevens, 1994), even if there are impediments to the portability of experience (Dokko, Wilk, and Rothbard, 2009). Consistent with these works, most research on the benefits of career diversity emphasizes human capital explanations rooted in learning or increased motivation. For example, Campion, Cheraski, and Stevens' (1994) study of the finance function of a pharmaceutical firm showed that job rotation is associated with benefits of personal development, job satisfaction, and organizational integration. Mobility—which may but need not necessarily be upward—is seen as career-enhancing because it broadens and deepens one's human capital (Wexley and Latham, 2002). The human capital acquired over the course of a career takes many forms, including learning about the many facets of a business, developing relational skills, and recognition of patterns, which help individuals know where to turn when seeking information of a particular type (e.g., Campion, Cheraskin, and Stevens, 1994). Diverse experience promotes discovery by conferring a greater ability to recombine disparate information (Taylor and Greve, 2006).

Without disputing the human capital perspective on mobility, other scholars have emphasized the benefits of career diversity that are associated with changes in the informal network that are concomitant with mobility (e.g., Dokko, 2004). Perhaps Granovetter (1988: 193) put it best: ''The meaning of individuals' history of mobility is inadequately captured by human capital arguments. As one moves through a sequence of jobs, one acquires not only human capital but also . . . a series of co-workers who necessarily become aware of one's abilities and personality.'' Similarly, Edström and Galbraith (1977) argued that job rotations serve to expand the intraorganizational networks of employees because each new role requires the formation of new task-relevant ties, even as some ties driven by the task structures of prior roles survive. Kleinbaum and Stuart (2012) showed that mobility from an operating unit into the corporate staff has the causal effect of broadening one's network, even after accounting for selection effects. Gulati and Puranam (2009) make a similar argument, albeit on a larger scale, in their case study of reorganizations at Cisco. They suggest that following a reorganization, some of the ties driven by the old structure persist, even as the new structure facilitates the formation of new ties. Corredoira and Rosenkopf (2010) showed that when inventors change firms, their former colleagues become more likely to cite patents owned by their new employer because the interpersonal tie, and its underlying flow of information, may survive the mobility event.

This wide-ranging research implies a hypothesis that is largely taken for granted within the field: that individuals who have experienced more mobility, and therefore have a more diverse career history, should be more likely to bridge otherwise disconnected groups. This argument, while hardly novel, provides an important building block for subsequent theoretical development, so I dub it a baseline hypothesis:

Baseline hypothesis: A diverse intraorganizational career history increases one's brokerage across social and organizational boundaries.

Two mechanisms are likely to drive the effect of a diverse career history on network structure. The first mechanism I hypothesize is a social capital mechanism that relies on preexisting relations between specific pairs of individuals. Working together forges strong ties. These ties originate in the formal task structure of the organization, which creates interaction requirements (Thompson, 1967), and are strengthened by the embeddedness of task-related interactions in the social structure of the work group, which creates normative pressure to expand the multiplexity of relations. Ties that originate as formal work relations tend, over time, to incorporate elements of friendship, advice, social support, and/or instrumental access (Kapferer, 1969; Fischer, 1982). When one or both members of such a strong relationship changes jobs, the communication frequency may drop off significantly, but the underlying trust and emotional connection changes more slowly (Levin and Cross, 2004), as evidenced by research suggesting that ties are severed far more slowly than they are formed (Gulati and Puranam, 2009; Corredoira and Rosenkopf, 2010; Kleinbaum and Stuart, 2012). Such ties may lie dormant for long periods of time, but may be quickly reactivated when needed (Levin, Walter, and Murnighan, 2011). Thus one mechanism by which people with diverse career histories become brokers is by accumulating a diverse rolodex of former colleagues from their prior roles with whom they stay in touch. When they do stay connected, despite being separated by significant organizational and social distance, such ties become bridging ties and form the basis of a ''rolodex mechanism.'' Because bridging ties connect people and groups that are otherwise disconnected, they are the tangible manifestation of brokerage (Burt, 2002; Valente and Fujimoto, 2010).

Hypothesis 1: Prior co-employment increases the likelihood that two organizationally distant people will be linked by a bridging tie.

The second mechanism through which mobility leads to brokerage is indirect ties. As people who change jobs accumulate an ever-larger set of former coworkers, these direct ties also allow them to tap into an even larger population of friends of friends. Burt (1992: 13) described such ''referrals'' as a key informational benefit of a network with structural holes. There are at least three reasons why an individual is more likely to interact with an organizationally distant person if they share a common acquaintance: awareness, normative pressure, and trust. First, much organizational communication is instrumental to the task requirements of the job, so interactions across organizational distance should provide access to information that is not available from more proximate, more accessible sources. A focal person is more likely to become aware that a distant potential contact possesses the needed information if they share an acquaintance in common. In this case, the third party plays the role of a *tertius iungens* (Obstfeld, 2005) or an ''integrator'' (Kleinbaum and Stuart, 2012), brokering an introduction between two contacts who could benefit from interacting. Second, even if the third party is not responsible for introducing the two actors, theory on social closure (Simmel, 1950; Granovetter, 1985; Coleman, 1988) implies that having a common acquaintance will render the two actors more likely to interact and more willing to help one another because of normative pressure resulting from social embeddedness. The mechanism for this effect is reputation: when two people are connected by a third party, they must consider their reputations as good colleagues in the eyes of the observer (Simmel, 1950), in addition to other considerations, such as an intrinsic desire to be helpful or expected benefits in the form of future reciprocity. Consistent with this perspective, indirect ties have been shown to promote the longevity of the dyad (Krackhardt, 1998) and facilitate the usefulness of bridges (Tortoriello and Krackhardt, 2010), thus enabling brokerage. A third, related reason why indirectly tied individuals might be more likely to interact is trust: two people with mutual acquaintances may be quicker to trust one another. Uzzi (1997: 43) characterized trust in embedded dyads as ''a predilection to assume the best when interpreting another's motives and actions.''

There is an irony inherent in the argument that indirect contacts facilitate brokerage.[1] Recent empirical work on brokerage traces its history to Burt (1992) and defines brokerage as a network tie that spans a structural hole; that is, a tie between two individuals who are *not* otherwise connected, either directly or indirectly. This perspective would suggest that sharing a common third party renders a direct tie redundant and undermines, rather than augments, structural brokerage. The structural holes perspective has been remarkably generative because it has enabled theory development and empirical measurement in ways that are more specific and precise than ever before. But it represents a subtle departure from broader classical conceptions of brokerage as a tie between otherwise disconnected groups. For example, Aldrich (1979: 248–259) emphasized the importance of individuals who, broadly, linked the organization with its environment. Tushman's research in R&D labs

[1] I am indebted to Nosh Contractor, Associate Editor Henrich Greve, and an anonymous reviewer for emphasizing this point.

(Tushman, 1978; Tushman and Katz, 1980; Tushman and Scanlan, 1981) focused on individuals who communicated across boundaries: between subgroups within the R&D lab, between the lab and the rest of the organization, or between the organization and the outside world. In this broader conception of brokerage, the question of whether the specific individuals are connected is less important than the question of whether their groups, or the ''thought worlds'' (Dougherty, 1992) they comprise, are connected. For example, Tushman and Katz (1980) argued that research projects perform better when the project team has access to relevant outside knowledge, regardless of whether that access occurs through a ''gatekeeper'' or through individual boundary-spanning ties. Similarly, Fernandez and Gould's (1994) typology defined brokerage in terms of the way individuals facilitate interactions between groups, not between individual people. They too were agnostic about the presence of other ties in the network.

This theoretical perspective is consonant with Burt's structural holes perspective. In his emphasis on efficiency, Burt (2004: 349) argued that information is relatively homogeneous within groups, that is—that thought worlds reside in networks—and so a tie to one group member is ''redundant'' with a tie to another member of the same group insofar as it provides access to the same thought world (Burt, 1992: 20). In a static examination of an organizational network, these two perspectives are likely to coincide. The very purpose of formal organization is to structure interactions, promoting specialization of knowledge and information (Allen, 1977), or what Burt called homogeneity of information within groups. But when we take a dynamic perspective and recognize that people move across intraorganizational boundaries, we find that people who were once part of the same thought world no longer are. As a result, when the focal actor has ties to two alters who are interconnected by virtue of their prior co-employment, they may nevertheless provide access to disparate current information, even as they provide redundant access to older information. The implication of this insight is that ties between the contacts in one's network need not undermine one's ability to be a broker. Sharing one or more mutual acquaintances through an ''embeddedness mechanism'' may motivate a contact to be helpful without necessarily rendering his or her information redundant.

Hypothesis 2: Sharing mutual acquaintances will increase the likelihood that two organizationally distant people will be linked by a bridging tie.

In addition to social-capital-based mechanisms, human capital may also explain some of the effect of career diversity on brokerage. Mobility within an organization promotes learning about the organization and its structure (Krackhardt, 1990) as well as about the different businesses in which the firm competes and what resources enable its competitiveness (Peteraf, 1993). This information constitutes a person's human capital and could potentially contribute to brokerage by providing relevant information about where in the organization to look for particular types of task-relevant information. Additionally, a diversity of prior experiences could promote the development of interpersonal skills that enable one to connect successfully with providers of information.

Despite these arguments, I do not hypothesize a human capital-based mechanism because limitations of the data preclude me from measuring human capital directly to test such a hypothesis in the current empirical setting. But if there is a human capital effect, and if people with diverse career histories have more human capital, I would expect that the baseline hypothesis (that diverse career history leads to brokerage) would continue to hold, even after controlling for social capital effects.

Career Trajectory Effects

To advance theory on the role of career trajectories in shaping social networks over time, it is useful to disaggregate diversity by considering the sequence of job functions that constitutes an individual career and the degree to which it conforms (or does not) to modal patterns in the organization. To begin to explore the relative advantages of a typical versus an atypical career trajectory, we look to theories of categorization. Conformity and non-conformity to categories, and the consequent conferral of legitimacy, are topics that have received substantial theoretical attention in economic sociology dating back at least to classic work in new institutional theory (Meyer and Rowan, 1977). Legitimacy has been defined as ''a generalized perception or assumption that the actions of an entity are desirable, proper, or appropriate within some socially constructed system of norms, values, beliefs, and definitions'' (Suchman, 1995: 574). Based on this definition, legitimacy is a useful lens through which to examine the sequence of job functions that constitutes an individual's career.

Membership in a well-established category gains attention and favorable evaluation from important audience members and is therefore essential to legitimacy. Conversely, deviation from typical behavior is viewed by critical audiences as illegitimate and is therefore penalized. Research has shown, for example, that the stock market discounts publicly traded firms that span multiple sectors of the market and that therefore garner less attention from equity research analysts, an important audience in capital markets (Zuckerman, 1999). French chefs find that their ratings decline when they combine gastronomic elements from different culinary categories (Rao, Monin, and Durand, 2005). Israeli wine producers receive lower product-quality ratings when they cross the boundary between kosher and non-kosher categories (Roberts, Simons, and Swaminathan, 2010). The ''categorical imperative'' to appear legitimate has been demonstrated in myriad empirical settings and, in response, actors will often strive to conform to well-established practices or categories (DiMaggio and Powell, 1983). The result is the widespread ''typecasting'' that occurs in Hollywood and other industries (Zuckerman et al., 2003). When non-conforming actors do persist, they often occupy peripheral niches (Carroll and Swaminathan, 2000). Thus the existence of categories serves to shape the expectations of observers (Hannan, 2007) and, in turn, to enforce category-consistent behavior in the actors themselves. The link between typicality, or fit within categories, and outcomes turns explicitly on the notion of legitimacy in the eyes of observers.

Viewed through the lens of categorization, atypical career trajectories lack legitimacy. In the present examination of the typicality of intraorganizational careers and brokerage, the relevant observers are the actor's potential network contacts, who must decide whether or not to accept the actor into their networks. The logic of categorization would suggest that conformity with typical, well-established career trajectories would garner rewards in the form of greater desirability as a member of others' networks (Reagans and Zuckerman, 2008). By comparison, ''organizational misfits,'' people whose career trajectories deviate from the organization's well-established patterns, would be viewed as undesirable contacts for others and would likely develop less favorable networks.

At the same time, there is a competing argument that could counteract the illegitimacy costs of an atypical career path. Network theory would predict that atypical career trajectories are beneficial to the formation of brokerage in intraorganizational networks because they create opportunities to bridge ''institutional holes'' (Burt, 1992: 148) in the social fabric of an organization that are otherwise rarely bridged. An institutional hole is a structural hole whose existence is induced by the formal structure of the organization. People who change jobs, but along typical career trajectories, may move between distant parts of an organization, but they do so along well-trodden paths, crossing institutional holes with numerous, established bridges. As a result, the boundary-spanning ties that are forged from such mobility are redundant, if not common, thus reducing their value. In contrast, people whose careers depart from the modal patterns in the organization, moving between parts of the organization that are rarely linked, are more likely to have networks that connect otherwise disconnected people and groups with little redundancy. Thus, based on a logic of redundancy, network theory makes the straightforward prediction that atypical careers should lead to more brokerage than typical careers.

Considering the role of uncertainty can help us adjudicate between these competing theoretical predictions. Uncertainty in evaluation is a critical boundary condition that limits the applicability of theories of categorization. As Zuckerman (2005: 173) described: ''[I]nsofar as the quality of work is hard to evaluate, a would-be generalist ('jack of all trades') who *chooses* to work in a wide variety of job categories will closely resemble the candidate who is unskilled in any of the categories ('master of none') and therefore is *compelled* to move from job-type to job-type as a result of failure.'' In market-based settings, such as a mediated market for freelance services (Leung, 2012), such information asymmetries leave outsiders with significant uncertainty about whether an atypical career path signals a ''jack of all trades'' or a ''master of none.'' Individuals with atypical careers have credibility problems in the eyes of external recruiters. As a result, legitimacy—a third-party endorsement that mitigates uncertain quality—becomes extremely important in guiding others' decision making about whether to interact with an individual (Reagans and Zuckerman, 2008). Consistent with this perspective, job candidates with atypical career histories are likely to be viewed unfavorably in mediated labor markets (Leung, 2012).

In contrast, in the internal labor market of a large organization, information asymmetry is reduced substantially (Williamson, 1975). As a result, a more informative indicator of quality is an individual's reputation (Kilduff and Krackhardt, 1994). When an individual's quality is known, either personally or to trusted contacts in one's network, category labels become less informative. Information that circulates readily in a corporate rumor mill is precisely the kind of information that allows one to differentiate between a ''jack of all trades'' and a ''master of none'' without having to rely on noisy indicators such as adherence to broad categories. In a sense, the first two hypotheses reverberate in this argument: when people know you, either directly (''rolodex mechanism'') or indirectly (''embeddedness mechanism''), that knowledge renders categorical evaluations less salient. Thus information asymmetry delineates a boundary condition for the value of legitimacy in theories of categorization (Zuckerman, 2005). And because information asymmetry is reduced in an organization's internal labor market, the effects of an atypical career on network structure are not well-specified based on theories of categorization. With the illegitimacy costs associated with theories of categorization reduced and the informational benefits of the structural holes argument amplified, I formulate the following ''misfit'' hypothesis:

Hypothesis 3: An individual's deviation from prototypical career trajectories in the organization gives rise to brokerage, even when accounting for diversity of experience.

METHODS

Data

The data for this study come from a large information technology and electronics company that I refer to as BigCo. In recent years, the company has pursued a corporate strategy of integration across many of its diverse products and, correspondingly, interdependence among its divisions; as a result, the company's leaders consider informal communication across divisional boundaries to be an important operational priority. The data I analyzed include the complete record, as drawn from the firm's servers, of e-mail communications among 30,328 employees during an observation period of three months in late 2006. The 30,328 employees in the sample are based in 289 different offices scattered across all 50 United States and collectively make up 24 percent of BigCo's U.S. employee population. Privacy laws in some European nations and corresponding company policy precluded my collecting data outside the United States.

I used a snowball sampling procedure, inviting 180 employees to participate in the study, 91 of whom (50.6 percent) agreed.[2] I excluded 25 of these because they were located outside the United States; the remaining 66 formed the core of the snowball sample. Collectively, these 66 people communicated with an additional 30,262 U.S. employees during the three-month period of the e-mail data. The sample comprises these 30,328 employees (66 core members plus their 30,262 direct contacts). Because the sample was collected using a snowballing procedure, it is not a simple random sample; nevertheless, it is compellingly large. Furthermore, the e-mail-based mechanism for the snowball sampling serves to cast a wide net in sweeping people into the sample. It is worth highlighting that 30,262 people were found to have exchanged e-mail

[2] Extensive data are not available on those people who declined to participate to know whether they differed systematically from those who did opt in. I believe that there are no significant sample selection issues due primarily to the large expansion of the sample from the 66-person core to 30,328 total members based on the quasi-randomness of large-scale mass e-mails. Additionally, some people who declined to participate in the core of the snowball sample nevertheless ended up in the broader sample because of their communications with others who did opt into the core of the snowball sample. No statistically significant differences in communication patterns were observed between those who did and those who did not opt in.

directly with the core of just 66 people; these 66 have an average of 3,415 direct contacts, although over 3,200 of those contacts exchanged or co-received only mass e-mails with the focal actor. As this statistic illustrates, the existence of widely distributed bulk e-mails introduces significant randomness. Nevertheless, the sample overrepresents some groups and underrepresents others relative to the employee population as a whole, so the possibility remains that the results could be biased in unknowable ways by this sampling procedure. To avoid any possibility of sampling bias, I exploited the large size of the sample and drew a stratified random subsample of employees designed to match the population demographics along key dimensions; details of the sub-sampling procedure are provided in Appendix A. The final subsample includes 15,116 employees. The analysis presented is based on the more conservative subsample, but the findings do not change substantively in the full sample or in various random draws of the subsample.

I measured intraorganizational networks using e-mail data. Electronic communications are increasingly viewed as a valid source of network data (e.g., Onnela et al., 2007; Kossinets and Watts, 2009), and e-mail data are particularly suitable for this study because of the broad, far-reaching ties that are of interest in the study of brokerage. A long literature suggests that survey respondents in general (Fowler, 1995), and network survey respondents in particular (Bernard, Killworth, and Sailer, 1981), are not always accurate informants. Empirical research comparing e-mail and survey measures of social networks has shown that respondents are especially likely to underreport their ties to physically or organizationally distant alters (Quintane and Kleinbaum, 2011), a group for whom accurate measures are especially important in the present study.

BigCo provided me with all internal e-mail data associated with members of the 30,328-person full sample that were on its servers at the time of data collection. The data came to me as 30,328 text files, each corresponding to the complete e-mail record of one employee. I cleaned and parsed these files; removed duplicates (e.g., a message sent from one member of the sample to another would appear as both a sent message and a received message); and recorded messages with multiple recipients for each sender-recipient pair. The resulting data set consisted of 114 million dyadic e-mail communications exchanged during the fourth quarter of 2006. From this data set, I excluded mass communications and blind carbon copies (BCCs); because mass communications occurred frequently and (by definition) included a disproportionate number of dyadic interactions, this screen reduced the number of communications by an order of magnitude.[3] I also limited the e-mail data to include only messages exchanged by members of the smaller subsample; because this excluded the entire corpus of communications for half the original sample as well as all communications between subsample members and those outside the subsample, this screen further shrank the size of the e-mail data set by nearly an order of magnitude. After cleaning and parsing the communications

[3] I excluded mass e-mails from the network analysis because they are unlikely to indicate socially meaningful interaction. I operationalized a mass e-mail as one with more than four recipients (Kossinets and Watts, 2006; Quintane and Kleinbaum, 2011), although results are robust to alternative thresholds. My interviews at BigCo revealed that blind carbon copies (BCCs) may be associated more with political behavior than with straightforward communication, making their meaning not straightforward to interpret. To be cautious, I therefore excluded BCCs, although the results are substantively unchanged by including them.

data, I collapsed them into a single cross-section and created dyad-level counts of $i \leftrightarrow j$ communications, where i and j index the individuals in the subsample. This approach has been termed the ''volume method'' of inferring a network from e-mail data (Wuchty and Uzzi, 2011). Other ways of inferring the network structure from the e-mail data (for example, treating an $i \rightarrow j$ tie separately from a $j \rightarrow i$ tie) yielded similar results. The final e-mail data set consists of 2.2 million non-mass, non-BCC e-mails exchanged among 198,081 actively communicating dyads drawn from the 15,116-person subsample and is treated as a single cross-section. The overall network density is 0.17 percent.

In addition to communication data, BigCo also provided demographic and human resource (HR) data about each of the employees in the sample, which are linked to the communications data through encrypted employee identifiers. The data include each employee's (time-invariant) gender and grade on the firm's 14-point salary hierarchy as of December 2006. Additionally, longitudinal records consisting of monthly observations over the 77 months co-terminating with the e-mail data—from 2000 through 2006—describe the business unit, major job function, job subfunction, U.S. state, and office location code for each employee.

Individual-level Variables, Measures, and Estimation

The theory depends on both individual-level constructs (e.g., career path typicality) and dyad-level constructs (e.g., prior co-employment); correspondingly, I estimated both individual- and dyad-level models.

Dependent variables. In the individual-level models, the dependent variable is brokerage. I measured brokerage in two different ways. First, to identify the individuals in BigCo who brokered interactions between groups that interacted infrequently, I created a variable, $Improb_i$, which gauges the *improbability* of each individual's overall communication profile, for which high improbabilities occur when an employee frequently communicates across specific group boundaries that are rarely crossed in the company. A theoretical advantage of such a brokerage measure is that it does not assume that indirect ties undermine brokerage. So in a dynamic model of network evolution across a career, in which embedded ties may motivate cooperative behavior without rendering information redundant, the $Improb_i$ measure is a valuable complement to other measures of brokerage. By definition, individuals with high composite improbability scores disproportionately form linkages between groups that communicate infrequently, consistent with theoretical conceptions of brokerage (Gould and Fernandez, 1989; Burt, 2005). To construct $Improb_i$, I began by creating matrices, one for each category (business unit, function, office, salary band), with elements defined as the cumulative proportion of their e-mail that members of group x exchange with those in group y. For example, because there are thirteen job functions at BigCo, I constructed a 13×13 matrix $Pr\mathit{Function}$, in which the xyth cell of the matrix is the proportion of his or her e-mail that the average member of function x exchanges with a member of function y. After producing similar matrices for all four categories, I then calculated, for all actually communicating dyads, the ''improbability'' that employees i and j would communicate as:

$$Improb_{ij} = 1 - \left[PrBU_{ij}{}^{\alpha} \times PrFunction_{ij}{}^{\beta} \times PrOffice_{ij}{}^{\gamma} \times PrBand_{ij}{}^{\delta}\right] \qquad (1)$$

where each of the Pr__ variables reflects the actual incidence of communication at BigCo between two average members of the specific pairs of business units, functions, offices, or salary bands represented in the *ij*th dyad. $Improb_{ij}$ thus takes a Cobb-Douglas form (subtracted from one), in which the exponents α, β, γ and δ are weighting factors designed to correlate with the relative importance of each boundary in structuring communications at BigCo and that sum to one by construction. To determine appropriate weights, I estimated a dyad-level count model of frequency of communication on same-group variables for these four groups (and on group size controls). The resulting four coefficients were rescaled to sum to one and used as exponents in calculating $Improb_{ij}$ in Equation 1. In the model estimates provided below, $\alpha = 0.325$; $\beta = 0.373$; $\gamma = 0.250$; and $\delta = 0.053$. $Improb_{ij}$ is computed for each communicating dyad, and it assumes its greatest value when the *ij*th dyad, given *i* and *j*'s group affiliations and the actual interaction frequencies in BigCo, have group membership profiles that make them least likely to communicate.

Two specific examples from the data, one low and one high, may help to illustrate the intuition of the measure. In one dyad with a very high 60-percent probability (and hence low improbability) of communication, the two members of the pairing were both software engineers in the same business unit; both men worked in the same Virginia office; both ranked as middle managers, and they even joined BigCo in the very same month. Nearly two-thirds of the dyads that share this particular combination of business unit, job function, office, and salary band were live communication links, and this is one of the highest baseline probabilities of interaction among all pairings of group memberships in the data (note that this $Improb_i$ score does not account for the fact that the members of this particular dyad shared a start date or were the same gender). Thus $Improb_{ij}$ is a low 0.40 (= 1 – 0.60) for this dyad. By contrast, a second dyad that actually communicated has just a 0.01 percent baseline probability of interacting, one of the lowest in the dataset. One person in this dyad was in the general executive management function, the other in sales; one was a senior executive (salary band 14), the other a middle manager (salary band 10); they worked in offices on opposite coasts; and they were in different business units. This is an extremely improbable pairing, with a high $Improb_{ij}$ value of 0.9999 (= 1 – 0.0001). These contrasting examples illustrate the intuition of the measure: the first dyad spans only salary bands, and communication across this boundary (especially between adjacent salary bands) is commonplace at BigCo. The second dyad represents a link that jumps four levels in the salary distribution, crosses the geographic expanse of the country, and spans functional and business unit boundaries. It connects two individuals who are highly unlikely to interact and thus represents a bridging tie, the dyadic manifestation of brokerage.

To move from the dyad to the person level, for each employee *i*, I took the weighted (by e-mail volume) average across all alters *j* to get $Improb_i$, the average improbability of *i*'s overall communication profile:

$$Improb_i = \frac{\sum_{j=1}^{n_i}\left(Improb_{ij} \times Freq_{ij}\right)}{Freq_i} \qquad (2)$$

For each focal actor *i, j* indexes *i*'s n_i communication partners; $Freq_{ij}$ measures the number of e-mails exchanged between *i* and *j*; and $Freq_i$ measures *i*'s total communication volume. The higher the value of $Improb_i$, the more of actor *i*'s communication spans boundaries that are infrequently spanned, connecting groups of people who are otherwise relatively inaccessible. Because $Improb_i$ is bounded between 0 and 1, ordinary least squares estimation would be biased and inconsistent; instead, I estimated fractional logit models (Papke and Wooldridge, 1996).

Of course, if results of the hypothesis tests depended critically on any particular choice of exponents in Equation 1, the usefulness of the measure would be in doubt. To assess its robustness, I estimated results using fifteen alternative vectors of exponents, α, β, γ, and δ. In the first, all four factors were equally weighted (0.25), making $Improb_i$ simply one minus the geometric mean of the four individual probabilities. In four vectors, one factor was weighted more heavily (0.4) than the other three (0.2). Conversely, the next four vectors, underweighted one factor (0.1) relative to the other three (0.3). Finally, six vectors weighted two factors heavily (0.3) and two lightly (0.2). Needless to say, each of these vectors was chosen arbitrarily; but the fact that the results across all fifteen vectors were substantively the same as the core results presented here increases confidence in the robustness of the $Improb_i$ measure.

As a second measure of brokerage, I also calculated Burt's (1992) structural constraint, an inverse measure of the presence of structural holes in one's network. An actor *i*'s structural constraint was defined as:

$$Structural\ constraint_i = \sum_{j=1}^{n}\left(P_{ij} + \sum_{q=1}^{n} P_{iq}P_{qj}\right)^2 \tag{3}$$

where P_{ij} represents the proportion of actor *i*'s e-mail volume exchanged with actor *j*. The inner summation in Equation 3 incorporates the indirect constraint imposed on actor *i* through connections among *i*'s direct contacts. I used the igraph package (Csardi and Nepusz, 2006) in the R statistical computing environment (R Development Core Team, 2010) to calculate constraint for each individual in the sample. I subtracted constraint scores from their global maximum to invert the distribution and get a direct measure of *Structural holes*. I followed Burt (1992, 2007) and estimated models using ordinary least squares regression with heteroskedasticity-robust standard errors.

The *Structural holes* measure differs conceptually from the improbability measure of Equation 2 because it is purely network-based: it increases when an individual has many direct contacts and when those contacts are disconnected from one another. The measure is agnostic to the formal group memberships of an individual or of his or her contacts. By contrast, $Improb_i$ assesses the degree to which the focal individual's contacts are separated by organizational and geographic boundaries. There is good reason to expect these measures to be correlated: to the extent that individuals concentrate their interactions within organizational and socio-demographic groups (Han, 1996; Kleinbaum, Stuart, and Tushman, 2009), then those who communicate across groups likely will also span many structural holes.

Independent variables. There are two key independent variables in the individual-level models. The first construct is *Career diversity*, independent of the sequence of positions held. I measured the diversity of a person's career as one minus a Herfindahl concentration index, calculated across the monthly set of positions held during the prior 77 months. Because the theory regarding the effect of career diversity on brokerage is general enough to encompass a variety of different types of career diversity, I measured an index of this form separately for business unit, job function and subfunction, and office location. For example:

$$\textit{Career diversity (location)}_i = 1 - \sum_{l=1}^{L} s_{il}^2 \tag{4}$$

where l indexes the L different office locations (L = 289 in the data set), and s_{il} represents the proportion of the 77-month observation period in which employee i was assigned to office location l. As an illustrative example, during the 77 months of the career history observation period, Jane spent two months in location A and the other 75 months in location B (and 0 months in each of the other of BigCo's 289 U.S. offices). Jane's *Career diversity (location)* measure is calculated as

$$1 - \left(\left(\frac{2}{77}\right)^2 + \left(\frac{75}{77}\right)^2 + \sum_{l=3}^{289}\left(\frac{0}{77}\right)^2\right) = 0.05$$

By contrast, Dick moved to a different office every year and a half, so his *Career diversity (location)* measure is

$$1 - \left(\left(\frac{18}{77}\right)^2 + \left(\frac{18}{77}\right)^2 + \left(\frac{18}{77}\right)^2 + \left(\frac{18}{77}\right)^2 + \left(\frac{5}{77}\right)^2 + \sum_{l=6}^{289}\left(\frac{0}{77}\right)^2\right) = 0.78$$

Consistent with intuition, Dick's career diversity (location) score is much larger than Jane's.

Whereas the career diversity covariates ignore the particular sequence of the focal actor's mobility, focusing instead on its content, the next set of covariates results from a sequence analysis of the job functions occupied by each actor during each month of the 6.5 years co-terminating with the e-mail data. Analysis of social sequences originated in the work of Abbott and collaborators (Abbott and Hrycak, 1990; Abbott, 1995; Abbott and Tsay, 2000), who applied methods developed in biochemistry for the analysis of sequences of nucleotides in DNA or amino acids in proteins (Sankoff and Kruskal, 1983) to the analysis of individuals' careers. Conceptualizing a career as a sequence of positions held, Abbott's early analysis was descriptive. He showed, for example, that the careers of musicians in nineteenth-century Germany tended to fall into one of twenty prototypical patterns (Abbott and Hrycak, 1990). In this analysis, I move beyond description by looking not only at the clustering of individuals' career sequences into prototypical patterns but also at the degree to which a given individual's career sequence conforms to any prototypical pattern and the extent to which such deviations are associated with an outcome of interest,

Table 1. The Nine Prototypical Career Paths at BigCo, Showing the Medoid of each Cluster

Cluster and Description	Medoid Sequence
(1) Services (SV) with a stint in sales (SL)	SV(60)-SL(17)
(2) Services	SV(77)
(3) Research & development (RD)	RD(77)
(4) Corporate staff functions: human resources (HR), administration (AD), supply chain (SC)—typically with little mobility between them	HR(25)-AD(9)-SC(43)
(5) Sales	SL(77)
(6) Marketing (MK)	MK(77)
(7) Research and development, with stints in services (or occasionally in sales)	RD(28)-SL(2)-SV(33)-RD(14)
(8) Finance (FI)	FI(77)
(9) Administration, sales and services	SV(10)-SL(29)-SV(14)-MK(1)-AD(12)-SV(5)-SL(6)

namely, brokerage. Methodological details about the sequence analysis are provided in Appendix B.

The result of the sequence analysis was a set of nine prototypical career paths at BigCo, summarized in table 1. I describe them here briefly and provide more details in the Online Appendix (http://asq.sagepub.com/supplemental), which graphs the most frequently occurring variations within all nine clusters in color. Five of the nine prototypical career paths involve no mobility: stable careers in the services, sales, marketing, or R&D functions are typical at BigCo. The remaining four prototypical career paths consist of typical patterns of mobility at BigCo. Cluster 1 includes consultants (services function) who have spent significant amounts of time in sales; for example, the medoid (analogous to a median, but in multidimensional space) career path in cluster 1 consists of 60 months in the services function, followed by 17 months in sales. This stint in sales is enough to significantly differentiate a cluster 1 career from a cluster 2 career as a functionally immobile consultant. Cluster 4 consists of various corporate staff functions (HR, administration, supply chain), most typically with little mobility between them.

By construction, every individual in the sample was assigned to one of these nine prototypical career paths. But, consistent with the theoretical framework, some individuals had careers that were more typical—that corresponded more closely to a BigCo career prototype—than others. To quantify career typicality, I calculated each actor's *Misfit*, a continuous measure of the Euclidian distance between his or her own sequence and the medoid of the cluster to which he or she was assigned. People with large *Misfit* scores are relatively far from any cluster; their careers deviate significantly from all of the prototypical career patterns in BigCo. For example, Sheryl (profiled in figure 1 above), the manufacturing executive in the corporate supply chain group with an extremely diverse and far-reaching network, has a *Misfit* score more than two standard deviations above the sample mean. I included in the individual-level models dummy variables indicating the career path cluster to which the focal actor belonged (with cluster 3 as the omitted category) as well as a continuous measure of the actor's *Misfit* with his or her cluster. Because the distribution of *Misfit* has a lower bound of zero, models were estimated on the natural logarithm of *Misfit* to improve model fit and avoid problems of skewness.

Control variables. I controlled for a range of individual-level sociodemographic variables that may affect an actor's propensity to engage in brokerage. Most importantly, I controlled for diversity of prior experience to show the effect of misfit on brokerage, net of pure diversity effects. A dummy variable *Female* was included to control for gender. I included dummy variables for assignment to the *Corporate headquarters* and for key job functions (*Marketing* and *Sales*) to account for task-related differences in communications patterns. I included dummy variables for pay grade—one for *Middle managers* and one for each executive rank, relative to the omitted category of rank and file—to account for the effect of seniority on communication patterns. To account for individual-level differences in communication patterns, I controlled for the natural logarithm of the focal actor's total e-mail volume, exchanged with other members of the sample and with BigCo employees not included in the sample, respectively. Finally, I included a series of size controls to account for size differences in the focal actor's business unit, job function, office location, and salary band (all log-scaled).

Accounting for endogeneity. To make a *causal* argument about how and why career mobility affects the structure of social networks when mobility is endogenously related to network structure creates difficulties for causal inference. The above analysis addresses endogeneity only through a simple lag structure—mobility events precede the observed network structure in time—but unobservable intermediate network structure may vary endogenously with individual mobility choices, making the identification of mobility effects on network structure problematic. Better solutions to this identification problem would be to use random assignment to different mobility conditions (Winship and Morgan, 1999), or an instrumental variable or natural experiment that is exogenously associated with individual mobility but that does not affect network structure (e.g., Angrist, 1990). Unfortunately, these approaches are rarely possible when studying careers, so the best remaining solution to the identification problem is to use a propensity score estimator (Rosenbaum and Rubin, 1984).

A propensity score here is the probability that an individual experiences atypical mobility, conditional on observable covariates. Propensity scores can eliminate bias by comparing outcomes (network structures, in the present case) between people with a similar ex ante probability of mobility, as estimated from their pre-treatment covariates. The propensity score is reliable and yields an unbiased estimate of the effect of an atypical career sequence on network structure, if we can assume that outcomes are independent of assignment to treatment, conditional on the observed covariates. The intuition is simple: if assignment to treatment covaries with the observed variables, then the propensity score can be used to create a weighted or matched sample (Rubin, 1977) in which assignment to treatment is effectively random, conditional on the observable covariates. In this way, propensity score estimators allow us to approximate a controlled experiment using observational data. I did this by augmenting the data set with an additional wave of e-mail network data at BigCo—a second cross-section, spanning the first quarter of 2008 and including an additional 157 million dyadic communications—and observing the sequence of mobility that occurs in the intervening 15-month period. Network structure during the earlier tranche of e-mail data (and other covariates) is used to construct the propensity score; the inverse of the propensity score is then used as a weighting factor in estimating the effects of mobility during the intervening period on network structure in the later tranche of e-mail data.

Although the propensity score approach better identifies the mobility effect than the simple lag structure, I was limited in this analysis by the relatively short observation period for mobility between the two panels of communication data—just 15 months long, compared with the 77 months of observable career history prior to the first panel of communication data. This shorter interval of sequence analysis is less likely to reveal stable patterns in the career sequences, but convergent results across both analyses would lend credence to the causal argument.

To do this analysis, I performed a new sequence analysis over the 15-month observation window that spanned the two panels of e-mail network data, producing a new set of cluster dummy variables as well as a new individual *Misfit* score. The clustering algorithm yielded a 13-cluster solution in which 10 clusters were centered on no-mobility sequences and which bears a strong resemblance to the 9-cluster solution of the earlier, 77-month observation period. Given the short interval of this sequence analysis and the overall rare occurrence of mobility events, the distribution of *Misfit* is highly skewed, with a large number of people (89 percent of the sample) having a *Misfit* value of zero, mostly because they experienced no mobility; therefore in models with the propensity score estimator, hypothesis 3 was tested using a dichotomous covariate, defined to be one for any individual with *Misfit* greater than zero and to be zero otherwise. Individuals with positive values of *Misfit* in the later period of career observations also had significantly higher values of *Misfit* in the earlier period ($p < .05$), suggesting that mobility during the brief second observation window was not merely idiosyncratic noise. Next, I ran probit models to estimate the effect of initial (i.e., pre-mobility) network structure (and other observable covariates) on the probability of having a nonzero value of *Misfit*. These estimated probabilities of the conditional likelihood that $Misfit > 0$ are propensity scores that can be used to construct matching estimators of the causal effect of having $Misfit > 0$ on network structure in the second tranche of e-mail data. In the second-stage models, observations could be weighted by the inverse of their propensity scores (inverse probability of treatment weights, or IPTW) to create a pseudo-population that would give consistent, unbiased estimates of the mobility effect on brokerage (Rubin, 1977). If the propensity scores are highly variable, however, extreme outlying values of the weighting factor could contribute heavily to the pseudo-population, resulting in an estimator with a large variance. This potential problem is averted by the use of a stabilized weight (Azoulay, Ding, and Stuart, 2009). The stabilized weight is calculated as the propensity score estimated on the full first-stage model divided by the propensity score estimated when excluding the covariate believed to be endogenous (i.e., pre-mobility network structure); stabilizing the weighting factor in the second-stage models increases their efficiency but does not affect the consistency of the estimator (Hernán, Brumback, and Robins, 2000). The results reported are no different from those obtained when unstabilized weighting factors are used. Finally, because results based on the primary sequence analysis excluded truncated sequences, I dropped from the propensity score analysis individuals with less than 77 months of tenure with BigCo; the results were substantively unchanged if these people were included.

Dyad-level Variables, Measures, and Estimation

Dependent variable. The dyad-level analog of brokerage is the existence of a bridging tie (Burt, 2002; Valente and Fujimoto, 2010). I operationalized a bridging tie as communication within a dyad that spanned significant organizational distance that made their communication improbable. Organizational distance was measured as $Improb_{ij}$ in Equation 1 above. Only dyads with $Improb_{ij}$ above the 75th percentile were included in the analysis (the analysis was robust to alternative thresholds); that is, every dyad in the analysis was, by construction, separated by significant organizational distance and was therefore unlikely to communicate. Thus the dyad-level sample was constructed to consist exclusively of dyads who were at risk of bridging. The dependent variable is a binary indicator for the presence (1) or absence (0) of communication between i and j during the three-month e-mail observation window in the fourth quarter of 2006. Although data were available on the count of e-mail interactions, I used a binary dependent variable because the presence of any communication reveals that an underlying relationship exists despite the organizational distance. In the context of such a long-distance tie, more frequent communication may not be a reliable indication of tie strength, as it might not be for a long-time friend who is no longer proximate, such as a college roommate. To test this assertion empirically, I estimated zero-inflated Poisson models (unreported results, available from the author). Coefficients of key covariates were significant in the inflation models but not in the count models; this analysis supports the choice to treat the dependent variable in dyad-level models as binary. Because the dependent variable is binary, models were estimated using logistic regression. To account for common person effects (Kenny, Kashy, and Cook, 2006), or non-independence between dyadic observations that included the same individual, standard errors were clustered simultaneously on *both* dyad members using the clus_nway.ado routine in Stata (Kleinbaum, Stuart, and Tushman, 2012); this approach is similar to adjusting standard errors with the quadratic assignment procedure or to estimating exponential random graph models but can be implemented in larger data sets (Cameron, Gelbach, and Miller, 2011).

I made one additional adjustment to the sample of dyads before estimating regression models. The matrix of dyadic communication was extremely large (over 114 million dyads, before imposing the organizational distance screen), so it was not expeditious to work with the full matrix. One potential solution to this problem would be to sample from the matrix randomly, but this approach ignores the fact that most of the variance in the estimation is provided by the realized ties (i.e., the non-zero cells) (Cosslett, 1981; Imbens, 1992; Lancaster and Imbens, 1996). Because of the sparsity of the matrix (over 99.8 percent of cells were zero), most of the variance would be lost. Instead, I constructed a case cohort data set (King and Zeng, 2001) that included all communicating dyads and a random sample of non-communicating dyads, weighted according to the inverse of their probability of being sampled. Results are extremely robust to the number of zeros included, as long as they are properly weighted; the present analyses included a 1:1 ratio of zeros to non-zeros.

Independent variables. I measured career diversity at the dyad level as the geometric mean of the Herfindahl indices of job functions (or, separately, of subfunctions, office locations, or business units) that the dyad members had spent time in during the observation window. The resulting variable was inverted to provide a measure of diversity, rather than concentration. Of course, career diversity is a fundamentally individual (not dyadic) construct; these models were intended to support the individual-level career diversity results, to lend credence to the dyadic approach, and to examine the effect of career diversity while controlling for dyad-level social capital constructs (i.e., prior co-employment and common third parties). Nevertheless, because this measure lacks face validity, two alternative specifications were checked for robustness. In one, I calculated the arithmetic, rather than the geometric, mean of the dyad members' Herfindahl indices. In another, I dichotomized each actor's career diversity (''mover'' versus ''stayer'') and entered dummy variables for ''one dyad member moved'' and ''both dyad members moved'' into regressions. Results were substantively identical across all specifications, increasing confidence in the findings.

To measure prior co-employment, I created a series of variables of the form *Length of prior co-employment in the same function*, the number of months prior to the terminal period in which *i* and *j* were both assigned to the same job function (or, separately, subfunction, office location, or business unit). By construction, the distribution is truncated at zero and is highly skewed, so I estimated models using the natural logarithm of one plus the variable. Unreported results showed substantively similar effects when the co-employment variables were specified as simple binary indicators of whether *i* and *j* had ever been members of the same group (as opposed to the length of such periods) or, alternatively, as counts of the number of distinct spells, where a spell was defined as a continuous period of any duration; this increases confidence that the findings reflect meaningful effects of prior co-employment and are not an artifact of variable specification.

To measure embeddedness with mutual acquaintances, I created a variable *Common third parties*, a count of the number of unique individuals who communicated with both *i* and *j* in the full sample. The distribution of the count of *Common third parties* was also truncated at 0 and, not surprisingly, was also highly skewed: nearly two-thirds of dyads in the sample had no common third parties and fully 90 percent had seven or fewer. The right tail was extremely long, however, with a maximum of 116 common third parties in one dyad. Because of this skewed distribution and because the effect of the marginal third party should diminish with the count of common third parties, I tested hypothesis 2 on the natural logarithm of one plus *Common third parties*.

Control variables. In estimating models, I controlled for a range of social and organizational variables that affect the likelihood of interaction between dyad members. Actors sharing a social focus (Feld, 1981) are more likely to interact, and organizational research shows that formal structure is a highly relevant social focus in intrafirm networks (Han, 1996; Kleinbaum, Stuart, and Tushman, 2009). For this reason, I controlled for *Same business unit, Same function,* and *Same subfunction*, dummy variables set to 1 if and only if the two members of the dyad were assigned to the same business unit, job function, and subfunction, respectively. Physical proximity is known to affect the propensity for interaction (Festinger, Schachter, and Back, 1950; Allen, 1977), so I included control variables for *Same office*, a dummy variable set to 1 if and only if the two dyad members were assigned to the same office building and *Distance in miles*, the natural logarithm of one mile plus the door-to-door driving distance between their office buildings. It is well known that the propensity of group members to interact tends to diminish with group size (Wasserman and Faust, 1994). To account for the effects of group size, I controlled for five group size variables corresponding to business unit, function, subfunction, office, and salary band. Consistent with an extensive econometric literature that applies models of gravitational attraction to network models of world trade (Carrère, 2006), these group size controls were specified as the natural logarithm of the product of the size of *i*'s group by the size of *j*'s group.

I included six control variables to absorb individual-level heterogeneity. I calculated *Within sample volume controls* for each dyad member as the natural logarithm of one plus the number of e-mails the actor exchanged with all other (non-*i*-*j*) partners in the sample. By including them, I conditioned on the total count of individual *i*'s and individual *j*'s e-mails. After conditioning on their total e-mail volume, the variance remaining to identify the other regression parameters relates to the distribution of communications across potential partners, rather than being driven by the overall communications volume of the two actors in a dyad. Likewise, I included *Beyond sample volume controls* for each dyad member, the natural logarithm of one plus the number of e-mails the two actors exchanged with other employees of BigCo who were not in the sample. These covariates adjust for the fact that the individuals within the sample may vary in their propensity to communicate beyond it. Finally, to ensure that the results are not driven by the size of a person's network, I included two control variables for *Degree*, the log of one plus each dyad member's count of unique communication partners.

RESULTS

Descriptive statistics and correlations for the individual-level data are shown in table 2. I begin by establishing that a diverse career history is associated with a broad present-day network in individual-level models. Table 3 presents regressions of the brokerage measures, $Improb_i$ or *Structural holes*, on a series of variables measuring the diversity of the focal actor's career experience across business units, job functions, subfunctions, or offices, and on control variables. Across all four group specifications, the coefficients of *Career diversity* on $Improb_i$ is positive: the more diverse an actor's career history across groups, the more likely that actor is to engage in improbable category-spanning communication. When the effect on *Structural holes* is examined, the results are similar: diverse experience across a variety of job subfunctions or offices is associated with more brokerage across structural holes; no significant effect is found for business unit or job function on *Structural holes*. These results are consistent with the baseline hypothesis, that career diversity is associated with brokerage.

Table 2. Summary Statistics and Correlations of Variables in Individual-level Analysis

Variable	Mean	S.D.	1	2	3	4	5	6	7	8	9	10	11
1. Improb_i	.846	.092											
2. Structural holes	1.375	.197	.24										
3. Misfit score 2000–2006 (logged)	1.241	1.818	.17	.10									
4. Career diversity (function)	.108	.176	.16	.08	.77								
5. Career diversity (subfunction)	.205	.229	.22	.13	.47	.67							
6. Career diversity (location)	.257	.247	.10	.00	.05	.04	.06						
7. Career diversity (business unit)	.167	.214	.15	−.05	.19	.17	.19	.14					
8. Corporate headquarters	.256	.436	−.12	.00	.07	−.08	−.11	−.02	.31				
9. Marketing	.023	.148	.13	.03	.04	.05	.05	.01	.00	−.02			
10. Sales	.159	.366	.22	.16	.07	.23	.36	−.06	−.08	−.24	−.08		
11. Middle manager	.911	.284	−.04	.07	−.07	−.02	−.07	.01	−.02	−.01	−.01	.07	
12. Executive (band 11)	.023	.151	.03	.08	.03	.02	.14	.01	.02	−.01	.03	−.03	−.62
13. Executive (band 12)	.009	.095	.04	.04	.02	.02	.06	.02	.03	.00	.02	−.01	−.31
14. Executive (band 13)	.002	.047	.03	.03	.03	.03	.05	.01	.02	.02	.04	−.01	−.20
15. Executive (band 14)	.001	.033	.01	.00	.06	.01	.01	.01	.02	.04	.01	−.01	−.11
16. Female	.301	.459	.07	.06	.14	.05	.02	.02	.07	.10	.09	−.01	.02
17. E-mail volume within sample (logged)	7.251	1.916	.04	.30	.08	.06	.07	.01	−.06	−.01	.04	.10	.02
18. E-mail volume beyond sample (logged)	6.902	1.878	−.04	.23	.07	.05	.07	.00	.00	.08	.00	.04	.13
19. Business unit size (logged)	8.131	1.170	−.03	−.03	−.01	−.09	−.03	.00	.12	.42	−.09	.06	.01
20. Job function size (logged)	8.629	1.107	.10	−.11	−.40	−.07	.07	.02	.02	−.32	−.17	.14	.05
21. Office location size (logged)	8.081	2.136	−.08	.03	−.01	−.01	−.01	−.05	−.06	−.03	.02	−.02	.00
22. Salary band size (logged)	8.610	.690	−.04	.18	−.05	.01	.00	.00	−.03	−.05	.00	.10	.81

Variable	12	13	14	15	16	17	18	19	20	21
13. Executive (band 12)	−.01									
14. Executive (band 13)	−.01	.00								
15. Executive (band 14)	−.01	.00	.00							
16. Female	−.03	.00	−.02	−.01						
17. E-mail volume within sample (logged)	.05	.05	.03	.03	.06					
18. E-mail volume beyond sample (logged)	.04	.02	.02	.02	.07	.07				
19. Business unit size (logged)	−.06	−.02	−.01	.01	.05	.00	−.07			
20. Job function size (logged)	−.05	.04	−.09	−.10	−.13	−.10	−.12	.01		
21. Office location size (logged)	.03	.01	.02	.02	.00	.03	.05	−.14	−.07	
22. Salary band size (logged)	−.29	−.28	−.29	−.21	−.01	.12	.18	−.04	.04	.02

To further establish a baseline, especially in light of the need for dyad-level variables, I also tested the baseline hypothesis using dyad-level models. Results appear in table 4. Model 1 contains only control variables. Models 2–5 show that across three of four group variables, organizationally distant dyads in which members have had diverse career histories are significantly more likely to communicate across organizational distance than less mobile dyads. Thus, consistent with the baseline hypothesis, both individual-level models and dyad-level models point to a significant effect of prior career diversity on present-day brokerage.

To tease apart empirically the mechanisms for the diversity effect on brokerage, I added to the dyad-level models in table 4 a set of covariates that measure the extent of prior co-employment between members of the dyad. In models 6–9, I added to the baseline model a series of variables indicating the number of months of prior co-employment between members of the dyad. Across all four measures, I find a significant, positive association between the amount of prior co-employment and the propensity of dyad members to bridge the present-day organizational distance between them.[4] These results support hypothesis 1, that there is a rolodex mechanism.

Using prior co-employment to test the rolodex mechanism assumes that prior co-employment is an indication that two people are likely to know each other. I refined the analysis to further test this assumption in two ways. First, network density in a group tends to decrease with group size, so this condition is particularly likely to be met when the group in which two people are co-employed is small. To increase confidence that the co-employment effect is driven by a rolodex mechanism, I split the sample of previously co-employed dyads at the median size of the group in which they were co-employed and ran a series of dyad-level regressions that included covariates of the type *Prior co-employment in a small function* (separately for function, subfunction, office, and business unit), dummy variables indicating that the dyad members were previously co-employed in a function, subfunction, office, or business unit that was smaller than the median group of its type. Because I was testing the effect of group size, only dyads previously co-employed in the relevant group (e.g., job function in model 1 of table 5) were included in these regressions; dyads previously co-employed in a different category of group or not at all were excluded. For this reason, I also did not run the full model. Results in table 5 show that two people who were previously co-employed in a small job subfunction were more likely to stay connected than two people previously co-employed in a large job subfunction. Similar results obtain for office location and business unit; the effect for job function is insignificant, perhaps because even the smallest job functions are very large. In a second, independent analysis, I respecified the co-employment variables to focus on smaller groups by combining group types (unreported results). For example, instead of counting the number of periods during which two people were in the same office (but possibly in different functions) or the same job function (but possibly in different office locations), I respecified the co-employment variables to count the number of periods during which the dyad members were in the same office *and* the same job function at the same time. Members of such dyads are much more likely to actually know one another. And consistent with the intuition of the rolodex mechanism, such dyads were much more likely to be connected in the present-day network, when they are no longer co-employed. These results lend credence to the assumption that co-employment is a reasonable proxy for the presence of an interpersonal relationship. They therefore buttress support for hypothesis 1, that a rolodex mechanism facilitates brokerage ties between people who have previously been co-employed.

There is at least one plausible alternative explanation that is also consistent with the results presented thus far. Rather than indicating that people are likely to know one another, co-employment might indicate the existence of a common affiliation that enables two people to interact more effectively. One could imagine that having experiences in common, even if not contemporaneously (e.g., ''remember that great taco stand across the street from

[4] Although specifying the functional form of the prior co-employment effect on brokerage was not the primary goal of this paper, unreported results indicate a curvilinear, inverted-U-shaped effect, in which more coemployment is associated with first increases, then decreases in the propensity to bridge the gap, exhibiting positive effects of the *Length of prior co-employment* variables and negative effects of (*Length of prior co-employment*)^2.

Table 3. Regressions of Individual-level Brokerage on Career Diversity and Control Variables*

	Improb$_i$					
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Career diversity (function)		0.588				0.480
		(0.029)••				(0.030)••
Career diversity (subfunction)			0.386			0.112
			(0.033)••			(0.040)••
Career diversity (location)				0.374		0.166
				(0.026)••		(0.032)••
Career diversity (business unit)					0.297	0.207
					(0.022)••	(0.022)••
Corporate headquarters	–0.026	–0.144	–0.026	–0.036	–0.024	–0.125
	(0.015)	(0.017)••	(0.015)	(0.015)•	(0.015)	(0.017)••
Marketing	0.841	0.806	0.821	0.798	0.832	0.782
	(0.031)••	(0.030)••	(0.031)••	(0.031)••	(0.031)••	(0.031)••
Sales	0.462	0.449	0.420	0.374	0.474	0.408
	(0.012)••	(0.012)••	(0.012)••	(0.013)••	(0.012)••	(0.013)••
Middle manager	0.079	0.072	0.087	0.088	0.065	0.069
	(0.038)•	(0.038)	(0.038)•	(0.038)•	(0.038)	(0.038)
Executive (band 11)	0.196	0.158	0.186	0.113	0.174	0.110
	(0.044)••	(0.044)••	(0.044)••	(0.045)•	(0.044)••	(0.044)•
Executive (band 12)	0.395	0.367	0.330	0.308	0.359	0.289
	(0.047)••	(0.046)••	(0.047)••	(0.046)••	(0.046)••	(0.045)••
Executive (band 13)	0.598	0.480	0.559	0.483	0.561	0.414
	(0.122)••	(0.121)••	(0.109)••	(0.111)••	(0.123)••	(0.115)••
Executive (band 14)	0.430	0.275	0.416	0.334	0.412	0.244
	(0.124)••	(0.116)•	(0.123)••	(0.122)••	(0.130)••	(0.121)•
Female	0.103	0.091	0.096	0.098	0.098	0.085
	(0.012)••	(0.012)••	(0.012)••	(0.012)••	(0.012)••	(0.012)••
E-mail volume within sample (logged)	0.006	0.009	0.005	0.005	0.005	0.007
	(0.003)•	(0.003)••	(0.003)	(0.003)	(0.003)	(0.003)•
E-mail volume beyond sample (logged)	–0.011	–0.010	–0.012	–0.013	–0.011	–0.011
	(0.003)••	(0.003)••	(0.003)••	(0.003)••	(0.003)••	(0.003)••
Business unit size (logged)	–0.026	–0.017	–0.020	–0.021	–0.026	–0.014
	(0.006)••	(0.006)••	(0.006)••	(0.006)••	(0.006)••	(0.006)•
Job function size (logged)	0.047	0.029	0.052	0.042	0.046	0.031
	(0.005)••	(0.005)••	(0.005)••	(0.005)••	(0.005)••	(0.005)••
Location size (logged)	–0.023	–0.021	–0.022	–0.022	–0.022	–0.020
	(0.003)••	(0.003)••	(0.003)••	(0.003)••	(0.003)••	(0.003)••
Salary band size (logged)	–0.030	–0.038	–0.033	–0.039	–0.033	–0.043
	(0.015)•	(0.015)•	(0.015)•	(0.015)••	(0.015)•	(0.015)••
Constant	1.807	1.861	1.702	1.836	1.765	1.805
	(0.130)••	(0.129)••	(0.130)••	(0.129)••	(0.130)••	(0.129)••
Observations	15,116	15,116	15,116	15,116	15,116	15,116
Log pseudolikelihood / R-squared	–4648.39	–4635.49	–4644.58	–4642.79	–4643.20	–4630.68

(*continued*)

the office?''), might nevertheless facilitate interaction between people who had never previously met (Feld, 1981). Research on alumni networks, for example, shows that sharing an alma mater will tend to make people more likely to interact and help one another, even if they have never met (e.g., Rider, 2012). To distinguish between a true rolodex effect and an effect of common affiliation, I added to the dyad models covariates for *Shared*

Table 3. (continued)

Variable	Structural Holes (7)	(8)	(9)	(10)	(11)	(12)
Career diversity (function)		–0.012				–0.024
		(0.007)				(0.007)••
Career diversity (subfunction)			0.024			–0.007
			(0.008)••			(0.010)
Career diversity (location)				0.035		0.044
				(0.007)••		(0.009)••
Career diversity (business unit)					0.008	0.007
					(0.006)	(0.006)
Corporate headquarters	–0.016	–0.014	–0.016	–0.017	–0.016	–0.012
	(0.005)••	(0.005)••	(0.005)••	(0.005)••	(0.005)••	(0.005)•
Marketing	0.016	0.017	0.014	0.011	0.015	0.012
	(0.008)•	(0.008)•	–0.008	–0.008	(0.008)•	–0.008
Sales	0.095	0.097	0.092	0.085	0.095	0.086
	(0.006)••	(0.006)••	(0.006)••	(0.007)••	(0.006)••	(0.007)••
Middle manager	–0.065	–0.065	–0.064	–0.064	–0.065	–0.064
	(0.014)••	(0.014)••	(0.014)••	(0.014)••	(0.014)••	(0.014)••
Executive (band 11)	0.114	0.115	0.114	0.107	0.113	0.106
	(0.012)••	(0.012)••	(0.012)••	(0.012)••	(0.012)••	(0.012)••
Executive (band 12)	0.205	0.206	0.202	0.198	0.204	0.197
	(0.013)••	(0.013)••	(0.013)••	(0.013)••	(0.013)••	(0.013)••
Executive (band 13)	0.258	0.26	0.256	0.248	0.258	0.25
	(0.018)••	(0.018)••	(0.018)••	(0.018)••	(0.018)••	(0.018)••
Executive (band 14)	0.182	0.184	0.182	0.175	0.182	0.178
	(0.041)••	(0.041)••	(0.041)••	(0.042)••	(0.041)••	(0.042)••
Female	0.010	0.010	0.010	0.010	0.010	0.010
	(0.003)••	(0.003)••	(0.003)••	(0.003)••	(0.003)••	(0.003)••
E-mail volume within sample (logged)	0.025	0.025	0.025	0.025	0.025	0.025
	(0.001)••	(0.001)••	(0.001)••	(0.001)••	(0.001)••	(0.001)••
E-mail volume beyond sample (logged)	0.011	0.011	0.01	0.01	0.011	0.01
	(0.002)••	(0.002)••	(0.002)••	(0.002)••	(0.002)••	(0.002)••
Business unit size (logged)	0.001	0.000	0.001	0.001	0.001	0.001
	–0.001	–0.001	–0.001	–0.001	–0.001	–0.001
Job function size (logged)	–0.040	–0.040	–0.039	–0.039	–0.039	–0.039
	(0.003)••	(0.003)••	(0.003)••	(0.003)••	(0.003)••	(0.003)••
Location size (logged)	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)•	(0.000)•	(0.000)•	(0.000)•	(0.000)•	(0.000)•
Salary band size (logged)	0.078	0.078	0.077	0.077	0.078	0.077
	(0.005)••	(0.005)••	(0.005)••	(0.005)••	(0.005)••	(0.005)••
Constant	–0.847	–0.858	–0.831	–0.812	–0.841	–0.825
	(0.046)••	(0.047)••	(0.046)••	(0.047)••	(0.047)••	(0.047)••
Observations	14,445	14,445	14,445	14,445	14,445	14,445
Log pseudolikelihood / R-squared	0.18	0.18	0.18	0.18	0.18	0.18

•$p < .05$; ••$p < .01$.

* Robust standard errors are in parentheses. $Improb_i$ (models 1–6) measures communication across organizational boundaries; constraint (models 7–12) is a reverse-scored measure of the communication across structural holes.

function, *Shared subfunction*, *Shared location*, and *Shared business unit*. Each of these covariates is a binary indicator that the two members of the dyad had some experience working in the same job function, subfunction,

Table 4. Dyad-level Models of the Probability That a Bridging Tie Will Occur between Organizationally Distant Actors*

	Baseline	Career Diversity				Length of Prior Co-employment			
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Same office	0.773	0.768	0.768	0.773	0.782	0.758	0.803	–0.269	0.667
	(0.197)••	(0.199)••	(0.198)••	(0.197)••	(0.195)••	(0.200)••	(0.198)••	(0.214)	(0.244)••
Same business unit	–0.447	–0.441	–0.441	–0.447	–0.458	–0.478	–0.491	–0.454	–1.052
	(0.110)••	(0.110)••	(0.110)••	(0.110)••	(0.111)••	(0.110)••	(0.110)••	(0.111)••	(0.112)••
Same function	0.433	0.430	0.435	0.433	0.425	–0.086	0.323	0.429	0.126
	(0.050)••	(0.050)••	(0.050)••	(0.050)••	(0.049)••	(0.056)	(0.050)••	(0.050)••	(0.057)•
Same subfunction	0.882	0.883	0.878	0.882	0.873	0.861	0.050	0.879	0.701
	(0.081)••	(0.081)••	(0.081)••	(0.081)••	(0.081)••	(0.082)••	(0.089)	(0.082)••	(0.087)••
Distance in miles (logged)	–0.104	–0.104	–0.104	–0.104	–0.103	–0.105	–0.103	–0.062	–0.102
	(0.013)••	(0.013)••	(0.013)••	(0.013)••	(0.013)••	(0.012)••	(0.013)••	(0.012)••	(0.014)••
Same band	0.245	0.246	0.244	0.245	0.242	0.239	0.223	0.248	0.231
	(0.029)••	(0.029)••	(0.029)••	(0.029)••	(0.029)••	(0.029)••	(0.029)••	(0.029)••	(0.031)••
Average tenure (logged)	–0.023	–0.018	–0.021	–0.023	–0.028	–0.047	–0.039	–0.035	–0.083
	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)	(0.028)••
Function diversity		0.318							
		(0.101)••							
Subfunction diversity			0.227						
			(0.088)•						
Office diversity				–0.042					
				(0.083)					
Business unit diversity					0.687				
					(0.098)••				
Length prior co-employment in the same function (logged)						0.189			
						(0.012)••			
Length prior co-employment in the same subfunction (logged)							0.309		
							(0.016)••		
Length prior co-employment in the same office (logged)								0.415	
								(0.031)••	
Length prior co-employment in the same business unit (logged)									0.441
									(0.011)••
Number of common third parties (logged)									
Constant	–7.236	–7.391	–7.346	–7.215	–7.455	–7.287	–7.289	–7.709	–9.434
	(0.401)••	(0.401)••	(0.402)••	(0.404)••	(0.400)••	(0.408)••	(0.399)••	(0.403)••	(0.452)••
Observations	138,262	138,262	138,262	138,262	138,262	138,262	138,262	138,262	138,262

(*continued*)

office location, or business unit, respectively, even if not contemporaneously. When entered together with the corresponding prior co-employment covariate, their coefficients are driven by the non-contemporaneous common-affiliation effect. Unreported results indicate that having shared prior experiences is significantly associated with bridging but that it absorbs a minority of the variance. Net of the shared affiliation effect, the rolodex effect of prior co-employment remains strong and significant.

To test hypothesis 2, that having ties to common third parties raises the likelihood of a bridging tie linking organizationally distant dyads, I added to the

Table 4. (continued)

	3rd Parties	Length Co-Employed and Common 3rd Parties				Co-employment, Common 3rd parties and Career Diversity			
Variable	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Same office	–0.572	–0.517	–0.545	–1.233	–0.590	–0.538	–0.550	–1.242	–0.599
	(0.746)	(0.736)	(0.729)	(0.780)	(0.751)	(0.740)	(0.729)	(0.779)	(0.754)
Same business unit	–1.372	–1.419	–1.428	–1.339	–1.446	–1.449	–1.433	–1.323	–1.434
	(0.314)••	(0.315)••	(0.323)••	(0.308)••	(0.306)••	(0.316)••	(0.323)••	(0.304)••	(0.305)••
Same function	0.205	–0.184	0.146	0.189	0.180	–0.247	0.143	0.190	0.182
	(0.122)	(0.143)	(0.115)	(0.121)	(0.117)	(0.152)	(0.115)	(0.123)	(0.116)
Same subfunction	0.208	0.196	–0.499	0.217	0.168	0.199	–0.507	0.222	0.168
	(0.278)	(0.265)	(0.298)	(0.278)	(0.272)	(0.262)	(0.297)	(0.276)	(0.271)
Distance in miles (logged)	–0.114	–0.109	–0.118	–0.086	–0.114	–0.111	–0.119	–0.088	–0.112
	(0.043)••	(0.043)•	(0.042)••	(0.043)•	(0.044)••	(0.043)••	(0.042)••	(0.043)•	(0.043)••
Same band	0.057	0.067	0.054	0.061	0.060	0.069	0.055	0.066	0.062
	(0.140)	(0.134)	(0.136)	(0.139)	(0.142)	(0.133)	(0.135)	(0.141)	(0.140)
Average tenure (logged)	–0.027	–0.044	–0.047	–0.045	–0.039	–0.054	–0.049	–0.045	–0.039
	(0.073)	(0.073)	(0.073)	(0.072)	(0.075)	(0.073)	(0.073)	(0.071)	(0.075)
Function diversity						–0.488			
						(0.337)			
Subfunction diversity							–0.075		
							(0.248)		
Office diversity								–0.338	
								(0.296)	
Business unit diversity									0.137
									(0.340)
Length prior co-employment in the same function (logged)		0.147				0.165			
		(0.033)••				(0.035)••			
Length prior co-employment in the same subfunction (logged)			0.253				0.256		
			(0.041)••				(0.041)••		
Length prior co-employment in the same office (logged)				0.293				0.300	
				(0.087)••				(0.088)••	
Length prior co-employment in the same business unit (logged)					0.079				0.076
					(0.037)•				(0.040)
Number of common third parties (logged)	3.909	3.901	3.890	3.893	3.837	3.911	3.891	3.895	3.837
	(0.107)••	(0.103)••	(0.104)••	(0.107)••	(0.107)••	(0.102)••	(0.103)••	(0.106)••	(0.107)••
Constant	–9.007	–9.036	–8.966	–9.283	–9.491	–8.735	–8.918	–9.077	–9.513
	(1.369)••	(1.340)••	(1.327)••	(1.372)••	(1.320)••	(1.320)••	(1.301)••	(1.431)••	(1.327)••
Observations	138,262	138,262	138,262	138,262	138,262	138,262	138,262	138,262	138,262

•$p < .05$; ••$p < .01$.

* Robust standard errors are in parentheses. Models testing career diversity's association with brokerage include controls for group size, within-sample e-mail volume, beyond-sample e-mail volume, and network degree.

baseline model the covariate *Common third parties* in model 10 of table 4. Consistent with hypothesis 2, the coefficient of *Common third parties* is positive and significant. Furthermore, its magnitude is extremely large, and its addition improves the pseudo-R^2 of the model dramatically, from 10 percent to 37 percent. Models 11–14 include both the *Length of prior co-employment* and the *Common third parties* covariates together. Results indicate that the effects

Table 5. Dyad-level Models of the Probability That a Bridging Tie Will Occur between Organizationally Distant Actors in Small vs. Large Groups*

Variable	(1)	(2)	(3)	(4)
Prior co-employment in a small (versus a large) job function	–0.026			
	(0.063)			
Prior co-employment in a small (versus a large) job subfunction		0.449		
		(0.076)••		
Prior co-employment in a small (versus a large) office location			1.032	
			(0.148)••	
Prior co-employment in a small (versus a large) business unit				0.162
				(0.061)••
Same office	0.153	–0.259	–0.101	0.557
	(0.387)	(0.475)	(0.259)	(0.308)
Same business unit	–1.145	–1.417	–1.020	–0.760
	(0.204)••	(0.246)••	(0.469)•	(0.134)••
Same function	0.043	0.233	0.489	0.070
	(0.060)	(0.099)•	(0.233)•	(0.072)
Same subfunction	0.883	0.323	0.334	0.639
	(0.084)••	(0.110)••	(0.402)	(0.108)••
Distance in miles (logged)	–0.086	–0.088	–0.081	–0.079
	(0.023)••	(0.031)••	(0.034)•	(0.021)••
Same band	0.317	0.269	0.405	0.256
	(0.049)••	(0.068)••	(0.138)••	(0.046)••
Average tenure (logged)	–0.045	–0.104	–0.210	–0.095
	(0.044)	(0.061)	(0.128)	(0.041)•
Constant	–6.437	–6.267	–3.819	–7.576
	(0.663)••	(0.972)••	(1.676)•	(0.652)••
Observations	37,043	12,262	2,453	27,473

•$p < .05$; ••$p < .01$.

* Robust standard errors are in parentheses. Models include controls for group size, within-sample e-mail volume, beyond-sample e-mail volume, and network degree.

of prior co-employment and Common third parties on brokerage are largely independent; both sets of covariates are positive, significant, and little diminished in magnitude relative to the models that enter them separately. The final set of results in table 4, taken from models 15–18, includes covariates for prior co-employment, embeddedness in common third parties, and career diversity. The surprising finding is that although career diversity is a significant predictor of the existence of bridging ties when entered into a model with only control variables (table 4, models 2–5) and when entered together with co-employment variables (not shown), when accounting for the effects of common third parties (hypothesis 2), career diversity has no remaining effect.[5] It thus appears that interactions across great organizational distance are facilitated by networks of both direct ties, resulting from prior co-employment, and indirect ties; net of

[5] To fully test a mediating relationship, I ran two additional sets of models: First, net of other controls, career diversity is a significant predictor of the number of common third parties. Second, the number of common third parties is a significant predictor of the existence of a bridging tie. Additionally, models in table 3 show that diversity alone predicts bridging, but diversity while controlling for common third parties does not. Together, these results point to common third parties as a mediating variable of the effect of career diversity on bridging ties.

Table 6. Individual-level Models of Brokerage on Sequence-analytic Covariates*

	Simple, One-Stage Models				Two-Stage IPTW Models			
	Improb$_i$		Structural Holes		Improb$_i$		Structural Holes	
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Misfit, 2000–2006	0.141	0.044	0.007	0.001				
(logged)	(0.007)•••	(0.009)•••	(0.002)•••	(0.002)				
Misfit, 2006–2008					0.108	0.101	0.003	0.006
(greater than zero)					(0.019)•••	(0.042)••	(0.002)•	(0.002)•••
Career trajectory cluster dummies	No	Yes	No	Yes	No	Yes	No	Yes
Two-stage IPTW model	No	No	No	No	Yes	Yes	Yes	Yes
Career diversity	–0.605	0.093	–0.039	–0.020	0.592	0.660	0.008	0.009
(job function)	(0.061)•••	(0.087)	(0.015)••	(0.022)	(0.060)•••	(0.067)•••	(0.005)•	(0.005)
Corporate	–0.025	–0.089	0.004	–0.005	–0.064	–0.073	–0.003	–0.001
headquarters	(0.017)	(0.016)•••	(0.005)	(0.005)	(0.031)••	(0.033)••	(0.003)	(0.003)
Marketing	0.941	0.648	0.013	0.010	1.396	1.369	0.001	0.000
	(0.037)•••	(0.056)•••	(0.009)	(0.013)	(0.263)•••	(0.266)•••	(0.003)	(0.005)
Sales	0.457	0.309	0.063	0.039	0.468	0.438	0.007	0.005
	(0.015)•••	(0.028)•••	(0.004)•••	(0.007)•••	(0.048)•••	(0.063)•••	(0.003)••	(0.006)
Middle manager	0.204	0.216	–0.013	–0.010	0.080	0.044	–0.031	–0.030
	(0.062)•••	(0.060)•••	(0.024)	(0.024)	(0.084)	(0.095)	(0.006)•••	(0.006)•••
Executive (band 11)	0.307	0.448	0.132	0.143	0.066	0.165	–0.010	–0.004
	(0.060)•••	(0.061)•••	(0.019)•••	(0.019)•••	(0.091)	(0.121)	(0.005)••	(0.004)
Executive (band 12)	0.375	0.325	0.180	0.183	0.310	0.348	0.014	0.016
	(0.069)•••	(0.069)•••	(0.022)•••	(0.022)•••	(0.074)•••	(0.097)•••	(0.005)•••	(0.005)•••
Executive (band 13)	0.685	0.593	0.273	0.285	0.307	0.666	0.016	0.020
	(0.132)•••	(0.117)•••	(0.020)•••	(0.020)•••	(0.159)•	(0.159)•••	(0.007)••	(0.011)•
Executive (band 14)	0.275	0.070	0.181	0.193	0.174	0.302	0.029	0.027
	(0.164)•	(0.137)	(0.057)•••	(0.056)•••	(0.197)	(0.216)	(0.009)•••	(0.011)••
Female	0.088	0.021	0.015	0.011	0.177	0.188	0.010	0.011
	(0.015)•••	(0.013)	(0.004)•••	(0.004)•••	(0.020)•••	(0.022)•••	(0.002)•••	(0.002)•••
Tenure in years	–0.025	–0.002	0.004	0.003	–0.009	–0.022	0.001	0.000
(logged)	(0.012)••	(0.011)	(0.003)	(0.003)	(0.019)	(0.022)	(0.002)	(0.002)
E-mail volume	0.006	–0.007	0.023	0.022	0.012	0.014	0.018	0.017
within sample	(0.004)	(0.004)•	(0.001)•••	(0.001)•••	(0.005)•••	(0.005)•••	(0.002)•••	(0.002)•••
E-mail volume	–0.018	–0.005	0.015	0.016	–0.085	–0.097	0.008	0.008
beyond sample	(0.004)•••	(0.004)	(0.001)•••	(0.001)•••	(0.014)•••	(0.016)•••	(0.002)•••	(0.002)•••
Business unit	–0.028	–0.141	–0.003	–0.007	–0.041	–0.019	0.001	0.001
size (logged)	(0.007)•••	(0.007)•••	(0.002)	(0.002)•••	(0.009)•••	(0.011)•	(0.001)	(0.001)
Job function	0.142	–0.011	–0.008	0.003	0.014	0.014	–0.002	–0.001
size (logged)	(0.007)•••	(0.012)	(0.002)•••	(0.003)	(0.009)	(0.010)	(0.001)••	(0.001)
Office location	–0.023	–0.021	0.000	0.000	0.010	0.011	0.000	0.000
size (logged)	(0.003)•••	(0.003)•••	(0.001)	(0.001)	(0.001)•••	(0.001)•••	(0.000)	(0.000)
Salary band	–0.059	–0.037	0.062	0.063	–0.050	–0.043	0.017	0.017
size (logged)	(0.020)•••	(0.020)•	(0.007)•••	(0.007)•••	(0.026)•	(0.029)	(0.003)•••	(0.003)•••
Constant	1.161	2.536	0.636	0.551	2.785	2.634	0.654	0.656
	(0.167)•••	(0.178)•••	(0.052)•••	(0.055)•••	(0.242)•••	(0.272)•••	(0.031)•••	(0.035)•••
Log pseudolikelihood / R-squared	–3119.12	–3056.26	0.19	0.20	–1648.17	–1346.71	0.20	0.19

•$p < .10$; ••$p < .05$; •••$p < .01$.

* Robust standard errors are in parentheses. Models 1–4 are simple, one-stage models estimated on the e-mail network of the fourth quarter of 2006 and of mobility during the prior 77 months; Models 5–8 are two-stage inverse probability of treatment weighted models estimated on the e-mail network of the first quarter of 2008 and of mobility during the prior 15 months, weighted by the inverse of the propensity score.

these social capital effects, career diversity offers little additional benefit for brokerage.

Finally, we turn to the sequence analytic test of hypothesis 3. Table 6 shows results of regressions of the brokerage measures on dummy variables for each of the prototypical career paths at BigCo on *Misfit*, the measure of career atypicality, and on control variables. The omitted category for the career trajectory dummy variables is cluster 3, continuous assignment to the R&D function. To conserve space, the cluster dummies are not shown in the table, but two results merit mention because they are consistent with conventional wisdom, and therefore lend face validity to the sequence analysis. First, consistent with anecdotal accounts of R&D as being inward-facing and socially isolated, each of the other career paths is associated with significantly higher values of both $Improb_i$ and *Structural holes* than the omitted category of R&D (although for two clusters, the differences in *Structural holes* are statistically insignificant). Second, the clusters representing job functions that are concentrated in the corporate headquarters—namely, finance (cluster 8) and other staff functions (cluster 4) —are associated with the most brokerage (see also Kleinbaum and Stuart, 2012), though not all differences are significant.

To examine the degree of fit with prototypical career trajectories, we consider the coefficients of *Misfit*, as estimated in the simple models (models 1–4) and in two-stage, inverse probability of treatment weighted models (models 5–8) of $Improb_i$ and of *Structural holes*. Across both measures of brokerage and across both model specifications, the results indicate that, consistent with hypothesis 3, the greater the misfit—the more the focal person's career trajectory deviates from the prototypical patterns at BigCo, even accounting for the overall diversity of his or her prior experience—the richer that person's communication network is in brokerage. By the $Improb_i$ measure of brokerage, this result is robust to whether or not the simple models include career path dummy variables, which account for the fact that some clusters are more cohesive than others. In the two-stage models that account for endogenous mobility, all four specifications support hypothesis 3. Table 6 shows the misfit effect while controlling for prior functional *Career diversity*. In unreported models, variables for career diversity with respect to office location, subfunction, and business unit were substituted, all yielding substantively identical results.

DISCUSSION AND CONCLUSION

It is well known that brokers in social networks gain benefits through their role in connecting otherwise disconnected actors, but there has been little theory or empirical evidence about the origins of brokerage in intraorganizational networks. In this paper, I examined the role of career processes in forging the bridging ties across organizational and social space that lie at the heart of brokerage. Empirical results from both individual-level and dyad-level models indicate that a diverse career history is associated with brokerage. I hypothesized and found empirical support for two distinct types of ties that enable brokerage in mobile individuals: ties to direct, personal contacts with whom one has worked previously and ties to indirect contacts, the friends of one's friends. After accounting for these direct and indirect ties, a diverse career history has no additional effect on brokerage. I also attempted to move beyond aggregate measures of diversity to examine specific career trajectories. I found evidence that organizational misfits, whose career paths are atypical in the organization, are especially likely to be brokers. Thus the evidence marshaled here suggests that the role of career trajectories in the origins of brokerage in intraorganizational networks is to modify the opportunity structure for mobile individuals by bringing them into direct or indirect contact with other, potentially valuable contacts. And the more atypical the career path, the less redundant, and thus more valuable, the ties to those contacts are likely to be.

Of course, these findings beg the question of what enables some people to interact more productively with their former colleagues and friends of friends than other people. Answering this question definitively is beyond the scope of the present research, but I surmise that overlaid on the structure of interactions are individual differences in ability to interact productively with dissimilar others. This conclusion echoes Burt (2010: 224), who suggested that ''brokerage seems not to be beneficial for the information it provides so much as it is beneficial as a forcing function for the cognitive and emotional skills required to manage communication between colleagues who do not agree in their opinion or behavior.'' Combining Burt's insight with the present results raises the possibility that people who have changed jobs more frequently may have better honed these cognitive and emotional skills than people who have stayed in the same role. They may be more cosmopolitan, in the sense that ''experience in many and diverse social worlds confers upon an actor a facility with interacting and exchanging productively in new social worlds'' (Reagans and Zuckerman, 2008: 936). If cosmopolitanism can be learned, it is a form of ''human capital in the creation of social capital'' (cf. Coleman, 1988). As a result of this propensity, cosmopolitans may be better brokers of social interactions, even after accounting for their richer stock of social capital.

Surprisingly, however, the career diversity effect on brokerage disappears when the rolodex and embeddedness mechanisms are entered into the model. One interpretation of this finding is that bridging ties are more likely to persist when they are embedded in common third parties (Krackhardt, 1998). Social capital thus seems to explain all the variance in the effect of career diversity on brokerage, leaving no significant variance left to be explained by human capital. This interpretation of the result should be regarded as merely tentative pending future research that explicitly measures human capital as well as social capital. Unfortunately, BigCo did not allow me to collect a good proxy for human capital. One might argue that *Function diversity* is a measure of human capital insofar as it captures the breadth of one's functional experience and knowledge (e.g., Boxman, De Graaf, and Flap, 1991). Because it also incorporates aspects of social capital, I do not make this claim. But to the extent that such an argument is convincing, it reinforces the suggestion that social capital may be a more important benefit of career mobility than human capital. Nevertheless, this tentative finding is surprising in light of the voluminous literature that emphasizes the human capital benefits of career mobility. Although the present research does not dispute the human capital perspective, it suggests that social capital may be at least as important a benefit of mobility. However surprising, this result is not entirely without precedent: empirical work that has examined the effect of job rotation on performance has shown little evidence of any human capital effect at all (Cappelli and Neumark, 2001).

This research has several limitations. First and foremost, despite the massive volume of data, this is nevertheless a case study of a single organization, and I thus make no claim of generalizability beyond the empirical setting. The company I studied is, in many respects, fairly typical of large, diversified, American firms, and I expect that the present findings would apply quite broadly. But because the data are limited to BigCo, I cannot know for sure. Relatedly, I can only examine mobility within BigCo itself. BigCo's employees undoubtedly have both networks and opportunities outside of BigCo that I could not observe. As such, I bound my claims to the domain of intraorganizational careers within the large enterprise.

Second, there is the possibility that the results presented are affected by sample selection bias because the sample was defined at a single point in time: December 2006. Career histories are captured retrospectively, over the preceding 6.5 years for all members of the sample. But missing from the sample are people who left the organization prior to December 2006. It is possible that departure from the organization is associated with career history and/or network structure. In particular, one concern is that people with atypical career histories might be more likely to become brokers, conditional on staying in the organization, but may also be more likely to leave. To assess this possibility empirically, I again exploited the second tranche of e-mail data by observing people who were in the sample during December 2006 but who left prior to the end of data collection in March 2008. Importantly, although departure from the organization prior to December 2006 was not observable, departures between then and March 2008 were observable. To assess the extent of survivorship bias, I estimated (unreported) logistic regressions of the probability of leaving the company during the interval between waves of e-mail data on *Misfit* scores during the primary period of career history observation, from 2000 to 2006. Results indicate no significant effect of *Misfit*, any of the dummy variables for prototypical career patterns, or the measures of network structure on the probability of departure from BigCo during the 15-month interval from the fourth quarter of 2006 to the first quarter of 2008, even though a significant number of people (nearly 8 percent of the subsample) left the firm. This result provides some reassurance that the results are not biased by a systematic association between *Misfit* and departure from or, conversely, retention in the sample.

Third, the construction of the $Improb_{ij}$ measure (Equation 1) makes the simplifying assumption that the four components it comprises vary independently. Although correlations among the components are low, they are not zero. Nevertheless, this simplifying assumption is necessary. In principle, a better approach would be to use a single matrix in which each cell corresponds to communication between members of *i*'s particular attribute vector and members of *j*'s particular attribute vector. Unfortunately, this approach is computationally infeasible: BigCo's 31 business units, 13 job functions, 289 office locations, and 15 salary bands imply 1.75 million unique combinations for each of *i* and *j*, or a matrix with over 3 trillion cells. The simplifying assumption of independence enables empirical estimation without dropping any of the important variables that affect interaction frequency.

Despite these limitations, this research makes several significant theoretical contributions. First, I identified organizational misfits, whose career trajectories deviate from the prototypical career paths in their organization, and demonstrate a causal effect of misfit with prototypical career sequences on brokerage in the communication network. Given the veritable mountain of empirical evidence that shows the benefits of brokerage, the implication of this research is that being an organizational misfit might be a valuable role for people to play and a valuable asset for organizations to possess. Yet we know little about what consequences might result from such a position in other aspects of organizational life. For example, misfits might be more inclined to suffer from the anomie (e.g., Merton, 1938) that undermines morale, productivity, and advancement opportunities, even as they accrue benefits to their networks. We also do not know to what extent the present results generalize to careers across organizations. On the one hand, we might anticipate that the external labor market would be marked by higher levels of uncertainty in evaluation, making theories of categorization, and their attendant concern for legitimacy, more salient (Leung, 2012). On the other hand, the network that a new employee brings to the organization is increasingly important to individual and organizational performance (Corredoira and Rosenkopf, 2010), making networks that span institutional holes in industry structure particularly valuable, for the same reasons that networks spanning institutional holes in organizational structure are valuable. Future research should thoroughly explore other consequences of atypical careers, both within and across firms.

By showing the network benefits of an atypical career, this research makes a contribution to theories of categories and categorization processes. Extant research has taught us the benefits of conforming to existing categories. This research reinforces the boundary conditions for those benefits: when category memberships are less salient because other information is available, the costs of illegitimacy are mitigated and may be offset by other benefits. In this case, because an individual's career history is a less salient determinant of whether a potential contact will accept that person into his or her network, the benefits of ties across infrequently traversed boundaries—ties that tend to be possessed by organizational misfits—exceed the costs of having an illegitimate career path.

The present research also makes two contributions to the literature on careers and career diversity. First, it brings a careers-as-process perspective (Hall, 2002) to the network literature by examining the effects of careers on network structure, a valuable complement to prior research documenting the effects of networks on career outcomes. Second, it raises questions about the role of human-capital-based explanations for the effect of career diversity on network structure. This finding has implications for researchers as well as for both firms and their human resource strategies and for individuals and their career management strategies. For researchers, the implication is that the extant research on the human capital benefits of career diversity should be complemented by a greater attention to the social capital benefits. More generally, this study points to the need to integrate research on careers as processes with social network research.

For firms, this result has implications for the design of rotational management programs and of career paths more generally. Most executives readily agree that rotational programs serve to build one's network; but in many firms, the social capital benefits of job rotation are implicitly viewed as accidental byproducts of the primary goal of building human capital. Rather than designing programs to increase human capital by providing a set of experiences that are functionally diverse, firms should design programs to provide a set of experiences that jointly optimize the breadth and depth of human capital and social capital. The practical differences between these two approaches are minimal and virtually costless; rather, the differences lie around the margins in the way programs are framed and implemented. Insofar as the career diversity effect really does lie in social capital, more than human capital, such subtle shifts in emphasis have the potential to significantly increase the value of rotational programs. Additionally, this work suggests that when firms are screening internal candidates to fill particular roles, they would do well to consider not only the relevance of the candidates' human capital—their accumulated knowledge and experience—but also their social capital—whether the network they've built across their career gives them ready access to parts of the organization that are interdependent with the role in question.

For individuals, this research has two significant implications. First, though it is practically a truism to say it, the finding that social capital mechanisms explain the career diversity effect means that networking matters. Perhaps more provocatively, the finding that organizational misfits become brokers suggests that the emphasis that many job-seekers place on having a coherent ''story'' is perhaps overblown, at least in an intraorganizational context. Rather, there is a fundamental trade-off to be considered: atypical career transitions may undermine perceived legitimacy, but they also create opportunities for forging rare and valuable bridges.

This research complements existing work on the dispositional antecedents of brokerage in two ways. First, I make no claim that variation in career history explains all variation in brokerage; rather, I assert that attention to career history is a valuable complement to other perspectives on the origins of brokerage. Second, dispositional traits are not measured in this study, so the possibility remains that the antecedents of brokerage described here are associated with dispositional antecedents to brokerage, such as self-monitoring (Sasovova et al., 2010; Kleinbaum, Jordan, and Audia, 2012). Although it seems unlikely that high self-monitors—people who tend to monitor their behavior in order to fit in with different audiences—would sort into careers that are atypical for the organization, it is perhaps more plausible that organizational decision makers would sort high self-monitors into atypical careers, because such people might be perceived as being more adaptable and therefore more likely to successfully navigate unusual transitions. This is an empirical question left to future research. Depending on the answer, this study makes a contribution either by describing an independent effect on brokerage or by providing a more granular account of the role of sequences of career mobility as one mechanism by which disposition affects network structure.

Acknowledgments

The author would like to thank Associate Editor Henrich Greve, Pino Audia, Bruce Harreld, Matissa Hollister, Steve Kahl, Andy King, Chris Marquis, Kathleen McGinn, Misiek Piskorski, Elena Obukhova, Jeff Polzer, Brian Rubineau, Beppe Soda, Toby Stuart, Denis Trapido, Mike Tushman, and Paul Wolfson; seminar participants at Harvard University; conference attendees at the ASQ/OMT Conference on Coordination, the Wharton People in Organizations conference, the Organization Science Winter Conference; members of the Tuck Research Workshop and the Dartmouth Interdisciplinary Network Research Group for useful comments and feedback; and of course, BigCo and its many employees who spent time with me. Per the usual disclaimer, any remaining errors are exclusively my own. Financial support from the Ewing Marion Kauffman Foundation is gratefully acknowledged.

REFERENCES

Abbott, A.
1995 ''Sequence analysis: New methods for old ideas.'' Annual Review of Sociology, 21: 93–113.
Abbott, A., and A. Hrycak
1990 ''Measuring resemblance in sequence data: An optimal matching analysis of musicians' careers.'' American Journal of Sociology, 96: 144–185.
Abbott, A., and A. Tsay
2000 ''Sequence analysis and optimal matching methods in sociology: Review and prospect.'' Sociological Methods Research, 29: 3–33.
Ahuja, G., G. Soda, and A. Zaheer
2012 ''The genesis and dynamics of organizational networks.'' Organization Science, 23: 434–448.
Alcácer, J., and M. Zhao
2012 ''Local R&D strategies and multilocation firms: The role of internal linkages.'' Management Science, 58: 734–753.
Aldenderfer, M. S., and R. K. Blashfield
1984 Cluster Analysis. Beverly Hills, CA: Sage.
Aldrich, H.
1979 Organizations and Environments. Englewood Cliffs, NJ: Prentice-Hall.
Allen, T. J.
1977 Managing the Flow of Technology: Technology Transfer and the Dissemination of Technological Information within the R&D Organization. Cambridge, MA: MIT Press.
Angrist, J. D.
1990 ''Lifetime earnings and the Vietnam era draft lottery: Evidence from social security administrative records.'' American Economic Review, 80: 313–336.
Azoulay, P., W. Ding, and T. E. Stuart
2009 ''The impact of academic patenting on the rate, quality, and direction of (public) research output.'' Journal of Industrial Economics, 57: 637–676.
Baum, J. A. C., A. V. Shipilov, and T. J. Rowley
2003 ''Where do small worlds come from?'' Industrial and Corporate Change, 12: 697–725.
Bernard, H. R., P. D. Killworth, and L. Sailer
1981 ''Summary of research on informant accuracy in network data and the reverse small world problem.'' Connections, 4: 11–25.
Boxman, E. A. W., P. M. De Graaf, and H. D. Flap
1991 ''The impact of social and human capital on the income attainment of Dutch managers.'' Social Networks, 13: 51–73.
Brass, D. J.
1984 ''Being in the right place: A structural analysis of individual influence in an organization.'' Administrative Science Quarterly, 29: 518–539.
Burt, R. S.
1992 Structural holes: The Social Structure of Competition. Cambridge, MA: Harvard University Press.
Burt, R. S.
1997 ''The contingent value of social capital.'' Administrative Science Quarterly, 42: 339–365.
Burt, R. S.
2002 ''Bridge decay.'' Social Networks, 24: 333–363.

Burt, R. S.
2004 ''Structural holes and good ideas.'' American Journal of Sociology, 110: 349–399.
Burt, R. S.
2005 Brokerage and Closure: An Introduction to Social Capital. New York: Oxford University Press.
Burt, R. S.
2007 ''Secondhand brokerage: Evidence on the importance of local structure for managers, bankers, and analysts.'' Academy of Management Journal, 50: 119–148.
Burt, R. S.
2010 Neighbor Networks: Competitive Advantage Local and Personal. Oxford: Oxford University Press.
Burt, R. S., and D. Ronchi
2007 ''Teaching executives to see social capital: Results from a field experiment.'' Social Science Research, 36: 1156–1183.
Cameron, A. C., J. B. Gelbach, and D. L. Miller
2011 ''Robust inference with multi-way clustering.'' Journal of Business and Economic Statistics, 29: 238–249.
Campion, M. A., L. Cheraskin, and M. J. Stevens
1994 ''Career-related antecedents and outcomes of job rotation.'' Academy of Management Journal, 37: 1518–1542.
Cappelli, P., and D. Neumark
2001 ''Do 'high-performance' work practices improve establishment-level outcomes?'' Industrial and Labor Relations Review, 54: 737–775.
Carrell, S. E.
2007 ''The national internal labor market encounters the local labor market: Effects on employee retention.'' Labour Economics, 14: 774–787.
Carrère, C.
2006 ''Revisiting the effects of regional trade agreements on trade flows with proper specification of the gravity model.'' European Economic Review, 50: 223–247.
Carroll, G. R., and A. Swaminathan
2000 ''Why the microbrewery movement? Organizational dynamics of resource partitioning in the U.S. brewing industry.'' American Journal of Sociology, 106: 715–762.
Coleman, J. S.
1988 ''Social capital in the creation of human capital.'' American Journal of Sociology, 94: S95-S120.
Corredoira, R. A., and L. Rosenkopf
2010 ''Should auld acquaintance be forgot? The reverse transfer of knowledge through mobility ties.'' Strategic Management Journal, 31: 159–181.
Cosslett, S. R.
1981 ''Maximum likelihood estimator for choice-based samples.'' Econometrica, 49: 1289–1316.
Csardi, G., and T. Nepusz
2006 ''The igraph software package for complex network research.'' InterJournal, Ms. no. 1695.
DiMaggio, P. J., and W. W. Powell
1983 ''The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields.'' American Sociological Review, 48: 147–160.
Doeringer, P. B., and M. J. Piore
1971 Internal Labor Markets and Manpower Analysis. Lexington, MA: D.C. Heath.
Dokko, G.
2004 ''What you know or who you know? Human capital and social capital as determinants of individual performance.'' Unpublished doctoral dissertation, University of Pennsylvania.

Dokko, G., S. L. Wilk, and N. P. Rothbard
2009 "Unpacking prior experience: How career history affects job performance." Organization Science, 20: 51–68.
Dougherty, D.
1992 "Interpretive barriers to successful product innovation in large firms." Organization Science, 3: 179–202.
Edström, A., and J. R. Galbraith
1977 "Transfer of managers as a coordination and control strategy in multinational organizations." Administrative Science Quarterly, 22: 248–263.
Feld, S. L.
1981 "The focused organization of social ties." American Journal of Sociology, 86: 1015–1035.
Fernandez, R. M., and R. V. Gould
1994 "A dilemma of state power: Brokerage and influence in the national health policy domain." American Journal of Sociology, 99: 1455–1491.
Festinger, L., S. Schachter, and K. Back
1950 Social Pressures in Informal Groups: A Study of Human Factors in Housing. New York: Harper.
Fischer, C. S.
1982 To Dwell among Friends: Personal Networks in Town and City. Chicago: University of Chicago Press.
Fleming, L., S. Mingo, and D. Chen
2007 "Collaborative brokerage, generative creativity, and creative success." Administrative Science Quarterly, 52: 443–475.
Flynn, F. J., R. E. Reagans, and L. Guillory
2010 "Do you two know each other? Transitivity, homophily, and the need for (network) closure." Journal of Personality and Social Psychology, 99: 855–869.
Fowler, F. J.
1995 Improving Survey Questions: Design and Evaluation. Thousand Oaks, CA: Sage.
Gould, R. V., and R. M. Fernandez
1989 "Structures of mediation: A formal approach to brokerage in transaction networks." Sociological Methodology, 19: 89–126.
Granovetter, M. S.
1974 Getting a Job: A Study of Contacts and Careers. Cambridge, MA: Harvard University Press.
Granovetter, M. S.
1985 "Economic action and social structure: The problem of embeddedness." American Journal of Sociology, 91: 481–510.
Granovetter, M. S.
1988 "The sociological and economic approaches to labor market analysis." In G. Farkas and P. England (eds.), Industries, Firms, and Jobs: Sociological and Economic Approaches: 187–216. New York: Plenum Press.
Gulati, R.
1995 "Does familiarity breed trust? The implications of repeated ties for contractual choice in alliances." Academy of Management Journal, 38: 85–112.
Gulati, R., and M. Gargiulo
1999 "Where do interorganizational networks come from?" American Journal of Sociology, 104: 1439–1493.
Gulati, R., and P. Puranam
2009 "Renewal through reorganization: The value of inconsistencies between formal and informal organization." Organization Science, 20: 422–440.
Hall, D. T.
2002 Careers in and out of Organizations. Thousand Oaks, CA: Sage.

Han, S.-K.
1996 ''Structuring relations in on-the-job networks.'' Social Networks, 18: 47–67.
Hannan, M. T.
2007 Logics of Organization Theory: Audiences, Codes, and Ecologies. Princeton, NJ: Princeton University Press.
Hernán, M. Á., B. Brumback, and J. M. Robins
2000 ''Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men.'' Epidemiology, 11: 561–570.
Hoffman, R., and B. Casnocha
2012 The Start-up of You: Adapt to the Future, Invest in Yourself, and Transform Your Career. New York: Crown Business.
Imbens, G. W.
1992 ''An efficient method of moments estimator for discrete choice models with choice-based sampling.'' Econometrica, 60: 1187–1214.
Kapferer, B.
1969 ''Norms and the manipulation of relationships in a work context.'' In J. C. Mitchell (ed.), Social Networks in Urban Situations: 181–244. Manchester: Manchester University Press.
Kaufman, L., and P. J. Rousseeuw
1990 Finding Groups in Data: An Introduction to Cluster Analysis. New York: Wiley.
Kenny, D. A., D. A. Kashy, and W. L. Cook
2006 Dyadic Data Analysis. New York: Guilford Press.
Kilduff, M., and D. Krackhardt
1994 ''Bringing the individual back in: A structural analysis of the internal market for reputation in organizations.'' Academy of Management Journal, 37: 87–108.
King, G., and L. Zeng
2001 ''Logistic regression in rare events data.'' Political Analysis, 9: 137–163.
Kleinbaum, A. M., A. H. Jordan, and P. Audia
2012 '''Do you read me'? How empathy mitigates the social costs of high self-monitoring.'' Tuck Working Paper, Dartmouth College: http://ssrn.com/abstract=2134566.
Kleinbaum, A. M., and T. E. Stuart
2012 ''Inside the black box of the corporate staff: Social networks and the implementation of corporate strategy.'' Tuck Working Paper, Dartmouth College: http://ssrn.com/abstract=1804220.
Kleinbaum, A. M., T. E. Stuart, and M. L. Tushman
2009 ''Communication (and coordination?) in a modern, complex organization.'' Tuck Working Paper, Dartmouth College: http://ssrn.com/abstract=1162171.
Kleinbaum, A. M., T. E. Stuart, and M. L. Tushman
2012 ''Discretion within constraint: Homophily and structure in a formal organization.'' Tuck Working Paper, Dartmouth College: http://ssrn.com/abstract=1749512.
Kleinbaum, A. M., and M. L. Tushman
2007 ''Building bridges: The social structure of interdependent innovation.'' Strategic Entrepreneurship Journal, 1: 103–122.
Kossinets, G., and D. J. Watts
2006 ''Empirical analysis of an evolving social network.'' Science, 311: 88–90.
Kossinets, G., and D. J. Watts
2009 ''Origins of homophily in an evolving social network.'' American Journal of Sociology, 115: 405–450.
Krackhardt, D.
1990 ''Assessing the political landscape: Structure, cognition, and power in organizations.'' Administrative Science Quarterly, 35: 342–369.

Krackhardt, D.
1998 ''Simmelian ties: Super strong and sticky.'' In R. M. Kramer and M. A. Neale (eds.), Power and Influence in Organizations: 21–38. Thousand Oaks, CA: Sage.
Lancaster, T., and G. Imbens
1996 ''Case-control studies with contaminated controls.'' Journal of Econometrics, 71: 145–160.
Lesnard, L.
2008 ''Off-scheduling within dual-earner couples: An unequal and negative externality for family time.'' American Journal of Sociology, 114: 447–490.
Leung, M. D.
2012 ''Dilettante or Renaissance man? How the sequence of category memberships affects credibility in an online market for services.'' Haas Working Paper, University of California, Berkeley.
Levin, D. Z., and R. Cross
2004 ''The strength of weak ties you can trust: The mediating role of trust in effective knowledge transfer.'' Management Science, 50: 1477–1490.
Levin, D. Z., J. Walter, and J. K. Murnighan
2011 ''Dormant ties: The value of reconnecting.'' Organization Science, 22: 923–939.
Maechler, M., P. J. Rousseeuw, A. Struyf, and M. Hubert
2005 ''Cluster: Cluster analysis basics and extensions.'' R package version 1.14.12.
Merton, R. K.
1938 ''Social structure and anomie.'' American Sociological Review, 3: 672–682.
Meyer, J. W., and B. Rowan
1977 ''Institutionalized organizations: Formal structure as myth and ceremony.'' American Journal of Sociology, 83: 340–363.
Miner, A. S.
1987 ''Idiosyncratic jobs in formalized organizations.'' Administrative Science Quarterly, 32: 327–351.
Mintzberg, H.
1973 The Nature of Managerial Work. New York: Harper & Row.
Needleman, S. B., and C. D. Wunsch
1970 ''A general method applicable to the search for similarities in the amino acid sequence of two proteins.'' Journal of Molecular Biology, 48: 443–453.
Obstfeld, D.
2005 ''Social networks, the *tertius iungens* orientation, and involvement in innovation.'' Administrative Science Quarterly, 50: 100–130.
Oh, H., and M. Kilduff
2008 ''The ripple effect of personality on social structure: Self-monitoring origins of network brokerage.'' Journal of Applied Psychology, 95: 1155–1164.
Onnela, J. P., J. Saramaki, J. Hyvonen, G. Szabo, D. Lazer, K. Kaski, J. Kertesz, and A. L. Barabasi
2007 ''Structure and tie strengths in mobile communication networks.'' Proceedings of the National Academy of Sciences, 104: 7332–7336.
Papke, L. E., and J. M. Wooldridge
1996 ''Econometric methods for fractional response variables with an application to 401 (k) plan participation rates.'' Journal of Applied Econometrics, 11: 619–632.
Peteraf, M. A.
1993 ''The cornerstones of competitive advantage: A resource-based view.'' Strategic Management Journal, 14: 179–191.
Podolny, J. M., and J. N. Baron
1997 ''Resources and relationships: Social networks and mobility in the workplace.'' American Sociological Review, 62: 673–693.

Quintane, E., and A. M. Kleinbaum
2011 ''Matter over mind? E-mail data and the measurement of social networks.'' Connections, 31: 22–46.
R Development Core Team
2010 ''R: A language and environment for statistical computing.'' 2.11.1 ed. Vienna: R Foundation for Statistical Computing.
Rao, H., P. Monin, and R. Durand
2005 ''Border crossing: Bricolage and the erosion of categorical boundaries in French gastronomy.'' American Sociological Review, 70: 968–991.
Reagans, R. E., and E. W. Zuckerman
2008 ''Why knowledge does not equal power: The network redundancy trade-off.'' Industrial and Corporate Change, 17: 903–944.
Reiley, A. C., and J. D. Mooney
1939 The Principles of Organisation. New York: Harper.
Rider, C. I.
2012 ''How employees' prior affiliations constrain organizational network positions: A study of U.S. venture capital and private equity.'' Administrative Science Quarterly, 57: 453–483.
Roberts, P. W., T. Simons, and A. Swaminathan
2010 ''Crossing a categorical boundary: The implications of switching from non-kosher wine production in the Israeli wine market.'' In G. Hsu, G. Negro, and Ö. Koçak (eds.), Research in the Sociology of Organizations, vol. 31: Categories in Markets: Origins and Evolution: 153–173. Bingley, UK: Emerald Books.
Rosch, E. H.
1973 ''On the internal structure of perceptual and semantic categories.'' In T. E. Moore (ed.), Cognitive Development and the Acquisition of Language: 111–144. Oxford: Academic Press.
Rosenbaum, P. R., and D. B. Rubin
1984 ''Reducing bias in observational studies using subclassification on the propensity score.'' Journal of the American Statistical Association, 79: 516–524.
Rosenfeld, R. A.
1992 ''Job mobility and career processes.'' Annual Review of Sociology, 18: 39–61.
Rousseau, D. M., V. T. Ho, and J. Greenberg
2006 ''I-deals: Idiosyncratic terms in employment relationships.'' Academy of Management Review, 31: 977–994.
Rubin, D. B.
1977 ''Assignment to treatment group on the basis of a covariate.'' Journal of Educational Statistics, 2: 1–26.
Sankoff, D., and J. B. Kruskal
1983 Time Warps, String Edits, and Macromolecules: The Theory and Practice of Sequence Comparison. Reading, MA: Addison-Wesley.
Sasovova, Z., A. Mehra, S. P. Borgatti, and M. C. Schippers
2010 ''Network churn: The effects of self-monitoring personality on brokerage dynamics.'' Administrative Science Quarterly, 55: 639–670.
Simmel, G.
1950 The Sociology of Georg Simmel. Glencoe, IL: Free Press.
Sorenson, O., and T. E. Stuart
2001 ''Syndication networks and the spatial distribution of venture capital investments.'' American Journal of Sociology, 106: 1546–1588.
Sorenson, O., and T. E. 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.

Spilerman, S.
1977 ''Careers, labor market structure, and socioeconomic achievement.'' American Journal of Sociology, 83: 551–593.
Stuart, T. E.
1998 ''Network positions and propensities to collaborate: An investigation of strategic alliance formation in a high-technology industry.'' Administrative Science Quarterly, 43: 668–698.
Studer, M., G. Ritschard, A. Gabadinho, and N. Müller
2010 ''Discrepancy analysis of complex objects using dissimilarities.'' In F. Guillet, G. Ritschard, D. Zighed, and H. Briand (eds.), Advances in Knowledge Discovery and Management: 3–19. Berlin: Springer.
Suchman, M. C.
1995 ''Managing legitimacy: Strategic and institutional approaches.'' Academy of Management Review, 20: 571–610.
Taylor, A., and H. R. Greve
2006 ''Superman or the Fantastic Four? Knowledge combination and experience in innovative teams.'' Academy of Management Journal, 49: 723–740.
Thompson, J. D.
1967 Organizations in Action: Social Science Bases of Administrative Theory, 2d ed. New Brunswick, NJ: Transaction.
Tortoriello, M., and D. Krackhardt
2010 ''Activating cross-boundary knowledge: The role of Simmelian ties in the generation of innovations.'' Academy of Management Journal, 53: 167–181.
Tushman, M. L.
1978 ''Technical communication in R&D laboratories: The impact of project work characteristics.'' Academy of Management Journal, 21: 624–645.
Tushman, M. L., and R. Katz
1980 ''External communication and project performance: An investigation into the role of gatekeepers.'' Management Science, 26: 1071–1085.
Tushman, M. L., and T. J. Scanlan
1981 ''Characteristics and external orientations of boundary spanning individuals.'' Academy of Management Journal, 24: 83–98.
Uzzi, B.
1997 ''Social structure and competition in interfirm networks: The paradox of embeddedness.'' Administrative Science Quarterly, 42: 35–67.
Valente, T. W., and K. Fujimoto
2010 ''Bridging: Locating critical connectors in a network.'' Social Networks, 32: 212–220.
Wasserman, S., and K. Faust
1994 Social Network Analysis: Methods and Applications. New York: Cambridge University Press.
Wexley, K. N., and G. P. Latham
2002 Developing and Training Human Resources in Organizations, 3rd ed. Upper Saddle River, NJ: Pearson/Prentice Hall.
White, H. C.
1970 Chains of Opportunity: System Models of Mobility in Organizations. Cambridge, MA: Harvard University Press.
Williamson, O. E.
1975 Markets and Hierarchies, Analysis and Antitrust Implications: A Study in the Economics of Internal Organization. New York: Free Press.
Winship, C., and S. L. Morgan
1999 ''The estimation of causal effects from observational data.'' Annual Review of Sociology, 25: 659–706.

Wuchty, S., and B. Uzzi
2011 ''Human communication dynamics in digital footsteps: A study of the agreement between self-reported ties and email networks.'' PLoS ONE, 6: e26972.
Zaheer, A., and G. Soda
2009 ''Network evolution: The origins of structural holes.'' Administrative Science Quarterly, 54: 1–31.
Zuckerman, E. W.
1999 ''The categorical imperative: Securities analysts and the illegitimacy discount.'' American Journal of Sociology, 104: 1398–1438.
Zuckerman, E. W.
2005 ''Typecasting and generalism in firm and market: Genre-based career concentration in the feature film industry, 1933–1995.'' In C. Jones and P. H. Thornton (eds.), Research in the Sociology of Organizations, vol. 23: Transformation in Cultural Industries: 171–214.New York: Elsevier/JAI Press.
Zuckerman, E. W., T.-Y. Kim, K. Ukanwa, and J. von Rittmann
2003 ''Robust identities or nonentities? Typecasting in the feature-film labor market.'' American Journal of Sociology, 108: 1018–1074.

APPENDIX A: Subsampling Procedure

To assemble the representative subsample, I created a three-dimensional matrix of salary band (middle managers, 11, 12, 13, 14, everyone else), function (general executive management, marketing, sales, services, everyone else; I excluded administrative staff altogether) and business unit (corporate headquarters, everyone else). For each of the 60 cells of this $6 \times 5 \times 2$ matrix, I calculated the sampling probability that would be needed to achieve a subsample rate of 15.9 percent of the U.S. employee population (compared with 23.8 percent of the U.S. employee population of the firm in the original sample). I chose to make the subsample representative of only selected groups in order to maintain a large sample size. Making the subsample representative across the board would have diminished the sample to just 2.9 percent of the U.S. employee population of the firm. Once I had these probabilities, I used a random number generator to determine whether each person in the overall sample, given his or her salary band, function, and business unit, would be included in the subsample. Several different random draws of the subsample all produced identical results.

APPENDIX B: Sequence Analysis Methodology

The sequence analysis proceeds through two steps; a comprehensive description of sequence analytic methods in the social sciences (including optimal matching) is available in the work of Abbott and collaborators. Consistent with the observations of other scholars (Abbott and Tsay, 2000; Lesnard, 2008), truncated sequences in my data tended to cluster together, creating a large number of clusters that were highly similar, except in length. Though the substantive conclusions are unchanged in the full sample, for the sake of parsimony I limited the sequence analysis to those individuals whose careers at BigCo spanned the entire 77-month period ending in 2006 (N = 9,797).

In the first step, I used optimal matching to estimate an $N \times N$ matrix for the dyadic distance in ''career space'' between the career paths of each pair of actors. The distance measure relies on the Needleman-Wunsch algorithm, which employs a series of insertions, deletions, and substitutions of job functions to find the least costly way to convert one person's sequence of job functions into the sequence of another person (Needleman and Wunsch, 1970; Sankoff and Kruskal, 1983); by assumption, insertions and deletions have a fixed cost of one, while substitution costs vary based on the observed frequency of transition between the job functions at BigCo, such that a substitution is always preferable to the equivalent insertion and deletion. The dyadic career distance between *i* and *j* is defined as the sum of the costs of all insertions, deletions, and substitutions needed to convert *i*'s sequence of job functions into *j*'s sequence of job functions. I implemented sequence analysis using the R package TraMineR (Studer et al., 2010).

Once the complete N $\times$ N career distance matrix was calculated, I used a clustering algorithm to group together those individuals with similar career sequences (i.e., actors separated by a small ''career distance'') and, in so doing, to induce from the data the set of prototypical career paths at BigCo. I used the partitioning around medoids (PAM) algorithm (Kaufman and Rousseeuw, 1990), as implemented in the R package *Cluster* (Maechler et al., 2005). A cluster's medoid is defined as the sequence in the data that is closest to the center of that cluster; it is analogous to a median but is defined in multidimensional space. The PAM algorithm determines the optimal partitioning of the data into predefined *k* clusters by randomly choosing *k* sequences as medoids, assigning each sequence in the data set to one cluster, then iteratively optimizing the choice of medoids to find the best-fitting solution of *k* clusters. I ran the PAM algorithm for *k* values of two through 50.

I chose this approach because it is consistent with research on categorization in cognitive psychology, which suggests that ''categories are composed of a 'core meaning' which consists of the 'clearest cases' (best examples) of the category, 'surrounded' by other category members of decreasing similarity to the core meaning'' (Rosch, 1973: 112). Empirically, Rosch's ''clearest cases'' correspond to the medoids of each cluster and describe the prototypical career paths within BigCo. An observation's *silhouette width* is a measure of how much better it fits with its assigned cluster, compared with the next-nearest cluster; the average silhouette width across all observations in the data set gives a summary measure of how well the clustering solution fits the data. The nine-cluster solution fit the data best, with an average silhouette width of 0.792; an average silhouette width above 0.70 is evidence that ''a strong structure has been found'' (Kaufman and Rousseeuw, 1990: 88), so I chose the nine-cluster solution. The resulting nine prototypical career paths at BigCo are described in table 1 and in the Online Appendix (http://asq.sagepub.com/supplemental).

Both quantitative measures of fit and concerns about theoretical parsimony suggest that the nine-cluster solution is a suitable choice. No clustering solution is perfect (Kaufman and Rousseeuw, 1990), however, and the problem with this approach is that a small number of people with unvarying, infrequently occurring career trajectories are assigned to clusters in which they fit poorly. For example, 53 people in the sample spent the entire observation period in the legal function. They were assigned to Cluster 4, consisting of various corporate staff functions. But unlike the other functions in Cluster 4, between which interfunctional mobility is fairly commonplace, legal is far more isolated, with few people moving between legal and other functions. As a result, these 53 people had very high misfit scores, reflecting not an atypically diverse career but simply a job function that was relatively isolated but was too small to warrant its own cluster. One possible solution to this problem would be to increase the number of clusters until a legal cluster is created. But because legal was so small, a cluster dedicated to legal did not emerge until the 38-cluster solution, which had many clusters that were highly similar to one another. As a result, this is not a very parsimonious solution. Instead, I dropped from the analysis reported in table 5 any individual who had stayed in the same function throughout the observation period *and* had a misfit score more than two standard deviations above the mean. Compared with unreported models that include these observations, the effect sizes in table 5 are slightly (but statistically significantly) larger, and their direction and significance were unchanged, lending confidence to the robustness of these results.

Additionally, because different clustering algorithms can sometimes yield different results (Aldenderfer and Blashfield, 1984), I also tried several methods of agglomerative and divisive clustering; alternative approaches yielded solutions with very similar sets of prototypical career sequences, lending further credence to these findings.

Author's Biography

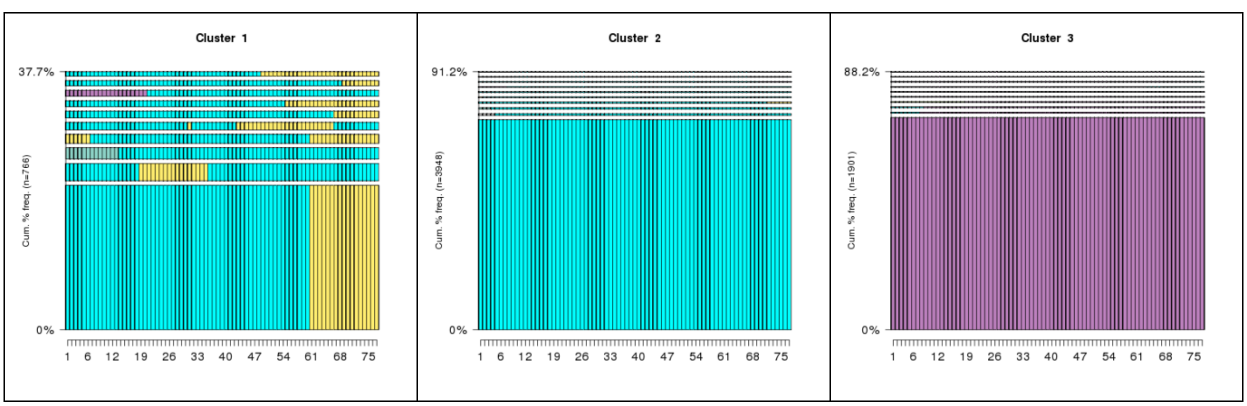
Adam M. Kleinbaum is an assistant professor in the Strategy and Management Group at the Tuck School of Business at Dartmouth College, 100 Tuck Hall, Hanover, NH 03755 (e-mail: Adam.M.Kleinbaum@tuck.dartmouth.edu). His research explores the origins and consequences of network structure in organizations; in doing so, he has been at the forefront of electronic communication and ''big data'' to measure intraorganizational networks. He holds a D.B.A. in Management from the Harvard Business School.

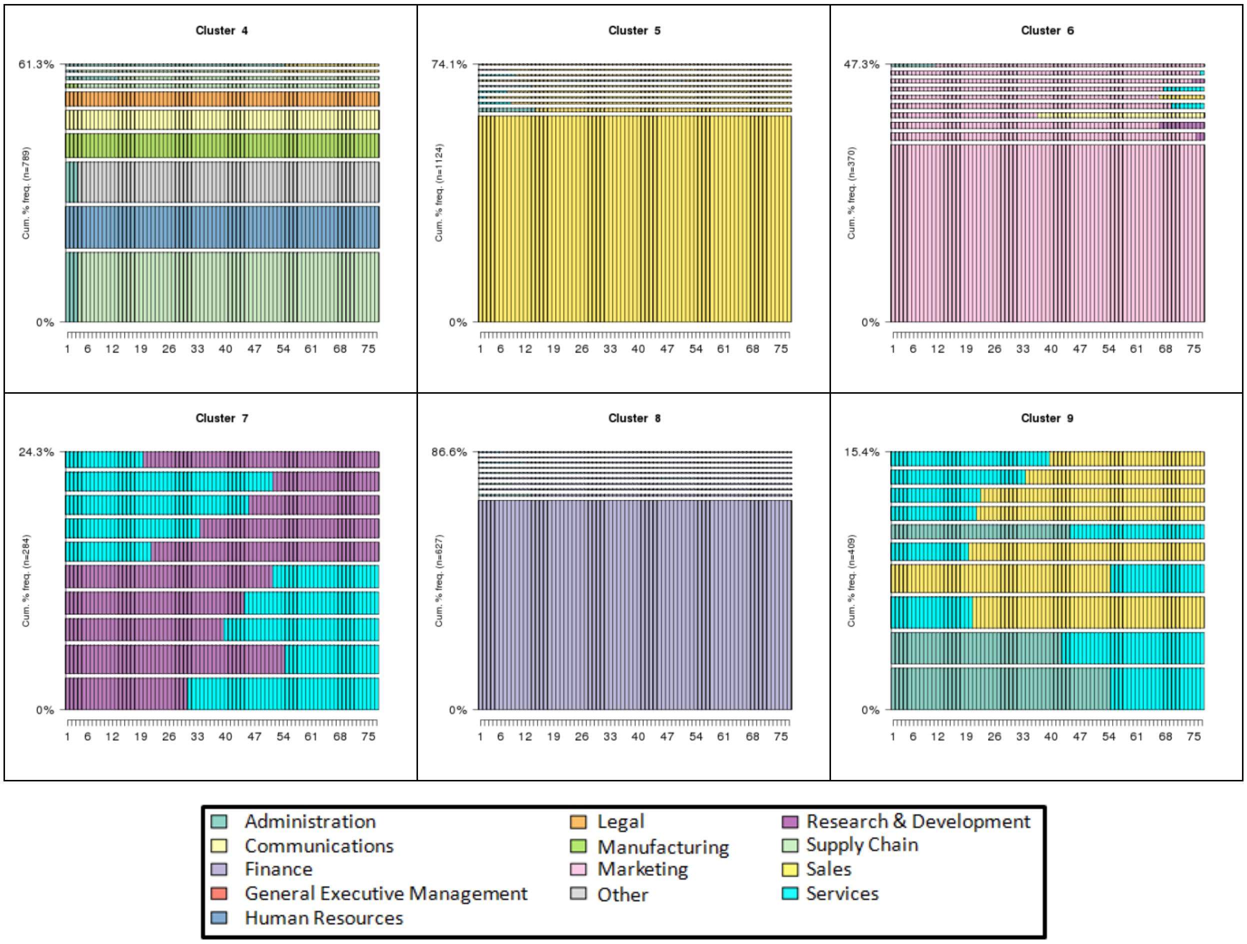
ONLINE APPENDIX: The Nine Prototypical Career Paths at BigCo

Each of the nine charts below conveys a great deal of information about one of the nine prototypical career trajectories at BigCo. Each horizontal row represents the sequence of job functions that make up a career; the 77 vertical slices of each row correspond to the 77 months of the observation period; the color of each slice indicates the job function held during that month; the height of the row corresponds to the number of people whose careers match that precise sequence. Within a chart, the patterns of color differ from row to row, corresponding to slight variations in career sequence. The Y-axis indicates the total number of people belonging to the cluster. Each chart shows the ten most frequently-occurring sequences within the cluster; some clusters are more homogeneous than others, so the ten sequences shown correspond to a maximum of 91% of people in the cluster (Cluster 2) and a minimum of 15% of people in the cluster (Cluster 9) (shown atop the Y-axis of each chart).

The Nine Prototypical Career Paths at BigCo, Showing the Medoid of each Cluster

Cluster and Description	Medoid Sequence
(1) Services (SV) with a stint in sales (SL)	SV(60)-SL(17)
(2) Services	SV(77)
(3) Research & development (RD)	RD(77)
(4) Corporate staff functions: human resources (HR), administration (AD), supply chain (SC)—typically with little mobility between them	HR(25)-AD(9)-SC(43)
(5) Sales	SL(77)
(6) Marketing (MK)	MK(77)
(7) Research and development, with stints in services (or occasionally in sales)	RD(28)-SL(2)-SV(33)-RD(14)
(8) Finance (FI)	FI(77)
(9) Administration, sales and services	SV(10)-SL(29)-SV(14)-MK(1)-AD(12)-SV(5)-SL(6)




Cluster 4
61.3%
0%
Cum. % freq. (n=789)
Cluster 5
74.1%
0%
Cum. % freq. (n=1124)
Cluster 6
47.3%
0%
Cum. % freq. (n=370)
Cluster 7
24.3%
0%
Cum. % freq. (n=284)
Cluster 8
86.6%
0%
Cum. % freq. (n=627)
Cluster 9
15.4%
0%
Cum. % freq. (n=409)
1 6 12 19 26 33 40 47 54 61 68 75
Administration
Communications
Finance
General Executive Management
Human Resources
Legal
Manufacturing
Marketing
Other
Research & Development
Supply Chain
Sales
Services