Organization Science
Vol. 18, No. 2, March–April 2007, pp. 165–180
ISSN 1047-7039 | EISSN 1526-5455 | 07 | 1802 | 0165
informs®
DOI 10.1287/orsc.1060.0242

Brokerage, Boundary Spanning, and Leadership in Open Innovation Communities

Lee Fleming
Morgan Hall 485, Harvard Business School, Boston, Massachusetts 02163, lfleming@hbs.edu

David M. Waguespack
Robert H. Smith School of Business, 4515 Van Munching Hall, University of Maryland, College Park, Maryland 20742-1815, dwaguesp@rhsmith.umd.edu

What types of human and social capital identify the emergence of leaders of open innovation communities? Consistent with the norms of an engineering culture, we find that future leaders must first make strong technical contributions. Beyond technical contributions, they must then integrate their communities in order to mobilize volunteers and avoid the ever-present danger of forking and balkanization. This is enabled by two correlated but distinct social positions: social brokerage and boundary spanning between technological areas. An inherent lack of trust associated with brokerage positions can be overcome through physical interaction. Boundary spanners do not suffer this handicap and are much more likely than brokers to advance to leadership. The research separates the influence of human and social capital on promotion, and highlights previously unexamined differences between brokerage- and boundary-spanning positions. Longitudinal analyses of careers within the Internet Engineering Task Force community from 1986–2002 support the arguments.

Key words: brokerage; leadership; open source; careers; social networks

Open innovation communities typically lack financial or corporate backing, forgo personal ownership rights to their members' work, rely on volunteers, and eschew formal planning and management structures. Despite these apparent handicaps, they have dramatically changed our conceptions of how innovation can and should be managed and have prompted calls for new theories of innovation (von Hippel and von Krogh 2003). Open community development methods appear superior to proprietary efforts on some measures (Kogut and Metiu 2001, Mockus et al. 2002). Open source operating systems challenge the world's most powerful software firms, and proponents of community innovation are extending the model into new contexts such as gene transfer technology (Broothaerts et al. 2005), medical innovation, crime solving, textbook and encyclopedia publishing, education, space exploration (Goetz 2003), and communities in developing countries, for which software customization in local languages remains cost prohibitive (*Economist* 2003). Open communities have spawned some of the most important technological breakthroughs of our era, including Web browsers, e-mail, and the Web itself. The very protocols that enable different Internet technologies to work together emerged from innovation within a voluntary, nonproprietary, and open innovation community.

Although most communities take full advantage of electronic communication media, their members innovate—not anonymously and randomly in cyberspace, but with reference to identity, reputation, technologically derived status, collegial networks, and physical interaction (Raymond 2000, Lakhani and von Hippel 2003, O'Mahony and Ferraro, forthcoming). Despite their bazaarlike, egalitarian, argumentative, unplanned, chaotic appearance, open innovation communities rely heavily on strong leadership to function effectively and to resist splintering, forking, and balkanization (DiBona et al. 1999, Lerner and Tirole 2001, Kogut and Metiu 2001, Lee and Cole 2003, von Hippel and von Krogh 2003, Lakhani 2004). Reputation counts: although it remains informal, leadership must be constantly earned through technical acumen and managerial skill. The implication is that in order to understand the success of open innovation communities, we must understand the emergence of their leaders (von Hippel and von Krogh 2003).

This understanding can also yield insights into how human and social capital influence career mobility. The importance of social capital is illustrated by Burt (1992), who demonstrates that *brokers*, individuals who connect otherwise disconnected actors, can exploit structural holes to advance more quickly in their careers. An individual who works with others who do not otherwise interact can control information and shape collegial and managerial perceptions. Predating Burt's structural hole theory, Allen (1977) and Tushman (1977) illustrate a widespread correlation between ability, ties across multiple organizations, and leadership. They describe how boundary spanners usually contribute the best engineering; identify, translate, and relay information within and across engineering firms; and often assume managerial authority. These two literatures, within sociological and management of technology traditions, have developed with surprisingly little mutual awareness or interaction, given the similarity of research questions. Furthermore, neither has estimated the relative importance of human versus social capital, or controlled for an individual's desire for promotion.

Although the measures of brokerage and boundary spanning correlate empirically, the concepts remain theoretically distinct. Brokers can span boundaries, but not all boundary spanners broker. This unexamined difference leaves a variety of unanswered questions. For example, what are the different mechanisms by which the occupants of the two positions attain leadership positions? Burt's early descriptions characterize brokers as calculating and politically savvy operators, while Allen and Tushman characterize boundary spanners as well respected guardians who redirect crucial information, both within and outside the firm. If these characterizations hold any validity, then colleagues of the broker and boundary spanner will surely hold different perceptions about individuals in each position. Colleagues will be less likely to trust a broker (Coleman 1988, Burt 2001b), and this lack of trust will be exacerbated within open innovation communities, which are inherently wary of balkanization and cooptation by commercial interests. The most effective strategies and behaviors of aspiring leaders will therefore be different for brokerage- and for boundary-spanning roles.

More generally, open innovation communities provide an opportunity to develop theories of human and social capital in a novel context that lacks pecuniary incentives, hierarchical authority, and formal structure. Leadership in such communities depends more on the trust and mobilization of peers than on approval of superiors. To wit, members cannot be fired or forced to participate in any activity, nor can they be compelled to pay attention to any other member. Ascendancy in such relationships relies purely, to borrow a phrase from politics, on "the power to persuade" (Neustadt 1990, p. 11). In addition to providing field settings in which to research a novel social phenomenon, open innovation communities often electronically archive their interactions. These public records afford social scientists an unprecedented opportunity to construct longitudinal databases of human and social capital, social and political processes, and a host of important outcomes, including the emergence of leadership.

To explain the emergence of leadership within open innovation communities, we induct theory from interviews, archival research, and field observations at community events. Estimating rate models for appointment as a working group leader in the Internet Engineering Task Force (IETF), arguably the world's first open innovation community (Bradner 1999), we first demonstrate the importance of technical contribution for future leaders. Individuals who broker work collaborations are more likely to assume leadership, but the effect is strongly contingent on physical presence within the community, a consequence of the diminished trust inherent in brokered social contexts. Consistent with the argument that they must overcome lack of trust, brokers also encounter difficulties when they attempt to span technological boundaries within the community. Boundary spanners, in contrast, do not suffer from a lack of trust and are more likely than brokers to assume leadership positions. In summary, future leaders are most likely to be individuals that make a strong technical contribution from a structural position that can bind the community together.

Open Innovation Communities and the IETF

The institution of science might be considered the first open innovation community (Dalle and David 2003), but its technological equivalent began in academic computer departments in the 1960s (DiBona et al. 1999). Consistent with commonly espoused norms of science (Merton 1973), department members made their software readily available to others, forgoing financial compensation, and, in return, earned reputation and status. The advent of the Internet and World Wide Web enlarged the potential scale of these efforts, giving rise to such extensive communities as Linux, Perl, Apache, Debian, and the IETF. We define an *open innovation community* as a group of unpaid volunteers who work informally, attempt to keep their processes of innovation public and available to any qualified contributor, and seek to distribute their work at no charge. While we induct theory from many open innovation communities, we test our hypotheses with a specific data set culled from the archives of the IETF, with the dependent variable as time to appointment as a working group leader within the IETF. We first briefly describe the IETF and process of appointment.

Although strictly speaking the IETF is an open standards rather than an open source community, it nevertheless exhibits critical open innovation community features in that any individual can volunteer to participate, proceedings remain transparent, and all technology originated therein is made available for free (Bradner 1999). The IETF is also the most long lived of the well-known open innovation communities and arguably exerts the greatest social and economic impact because of its association with the Internet. Of its potential as a model for open innovation communities Bradner (1999, pp. 47 and 52) reports that

> IETF standards are developed in an open, all-inclusive process in which any interested individual can participate. All IETF documents are freely available over the Internet and can be reproduced at will. In fact the IETF's open document process is a case study in the potential of the Open Source movement.... The IETF supported the concept of open sources long before the Open Source movement was formed.

The IETF emerged in 1986 from an amalgam of ad hoc Defense Advanced Research Projects Administration (DARPA) committees and has no official mandate to govern Internet technology. Challenged by more traditional standards-developing organizations and even government bodies, it has nonetheless emerged as the de facto standards-developing organization for the Internet (Abbate 1999, Mowery and Simcoe 2002; see also Harris 2001; Bradner 1996, 1998; and Hoffman and Bradner 2002 for insiders' descriptions). The IETF develops and maintains the core Internet standard, TCP/IP (Transmission Control Protocol/Internet Protocol), as well as many other standards that are pervasive in modern computing and networking but largely invisible to the average user. Although much of the communication and work of the IETF is conducted via electronic media, members meet three times a year, carrying forward a tradition begun in January 1986 when 21 IETFers met for the first time in San Diego. Membership remains open to all comers and has continued to grow, with as many as 2,800 individuals attending meetings and thousands more interacting online. Members participate in the IETF at least nominally as individuals (Bradner 1999), although they typically work for firms, universities, or governments. There being no dues or membership lists, in principle any person can join the IETF.

The IETF accomplishes most of its work within aptly named working groups (WGs) organized under larger functional areas. Individuals can freely associate, virtually or physically, with any of the extant technical working groups. Groups and their leaders seek and, if successful, are granted charters to address specific technical problems within a delimited time and domain. The Secure Shell Group, for example, was chartered to update and standardize the popular SSH protocol (Secure Shell secsh 2003). Working groups have chairs as well as individuals or design teams charged to produce documents that detail proposed standards. Area directors (ADs) appoint working group chairs. A nominating committee (NOMCON), randomly selected from community members who have attended at least two IETF meetings in the past two years, appoints ADs (generally two for each area).

Within the period of study, 344 working groups exist for varying periods of time, and some 480 individuals are first appointed to chair one of those groups. Approximately 32% of first-time working group chairs are appointed to lead existing working groups, either replacing an incumbent chair or working alongside the incumbents in a cochair capacity. The other 68% of first-time appointments are for new working groups. New working groups emerge mainly from grassroots interest in a topical area and are typically preceded by a Birds of a Feather (BOF) meeting convened during a conference.[1] For example, the IETF has a series of working groups related to different aspects of the domain-naming system (DNS). The first stage in group formation entails soliciting BOF participation via electronic invitations disseminated to the IETF mailing list. Organizers who generate favorable sign-up response receive physical meeting space at the next IETF conference. If a meeting is well attended and elicits broad interest, an AD will charter a group and appoint a chair. The chair is generally, but not always, the BOF organizer. Appointment can thus be construed as confirmation of initially successful leadership and a vote of confidence by community members and current leaders in the individual's leadership potential.

Individuals aspire to working group leadership in the IETF for a number of reasons: First, it enables someone to influence the direction of the technology development. This can happen if an individual inventor seeks legitimacy for her personal inventions (Bradner 2006). Second, it provides an opportunity to gain management experience without promotion within a private firm. Third, it gives an individual visibility in a community that can provide professional opportunity and career mobility. Finally, firms often support their employees' aspirations for leadership because the firm gains esteem in the larger technological and business community. Such esteem is particularly important for a startup that wishes to undergo a liquidity event.

Human Capital, Social Capital, and Leadership in Open Innovation Communities

Open innovation communities, like most engineering cultures, highly value technical contributions, eschewing titles and even democracy in favor of proven technology (Wasserman 2003). Despite a strong aversion to nontechnical sources of prestige and authority, community members readily recognize technical contribution. Successful contributors, according to Raymond (1998, p.), gain "good reputation among one's peers, attention and cooperation from others... and higher status in the... exchange economy." Rivlin (2003) illustrates how Linus Torvalds (the original author of LINUX) realizes that his authority is technically derived, tenuous, and constantly in need of collective reaffirmation:

> His hold over Linux is based more on loyalty than legalities. He owns the rights to the name and nothing else. Theoretically, someone could appropriate every last line

> of his OS [operating system] and rename it Sally. "I can't afford to make too many stupid mistakes," Torvalds says, "because then people watching will say, hey, maybe we can find someone better. I don't have any authority over Linux other than this notion that I know what I'm doing." He jokingly refers to himself as "Linux's hood ornament," and he's anything but an autocrat. His power is based on nothing more than the collective respect of his cohorts.

Status accrues to past contributors and translates into a higher probability of future leadership (Lee and Cole 2003). As Hamm describes (2005), "In a world where everybody can look at every bit of code that is submitted, only the A+ stuff gets in and only the best programmers rise to become Torvalds's top aides." Individuals who can solve difficult problems gain reputation and esteem in the opinions of their colleagues. Such individuals will become leaders whose opinions are sought out, respected, and that will influence the community's future.

Hypothesis 1 (H1). *Technical contribution will increase a member's likelihood of becoming an open innovation community leader.*

While technical contribution remains the primary prerequisite for aspiring leaders, the most important role—and challenge—for an open innovation community leader is to integrate and bind the community together. This occurs because open innovation communities remain voluntary; members can always leave and mobilize a new effort, a process known as *forking*. (DiBona et al. 1999, Kogut and Metiu 2001, Lerner and Tirole 2001). Of Torvalds, Rivlin (2003, p. 157) writes, "More than anything he seeks to avoid taking sides in a way that might splinter his followers. 'I'd much rather have 15 people arguing about something than 15 people splitting into two camps, each side convinced it's right and not talking to the other,' he says."

Leaders can forestall this process by creating and occupying network positions, such as social brokerage and boundary spanning, that integrate and bind the community together (Perrone et al. 2003). Burt (1992, 1997, 2000, 2001a) defines a *broker* as the only social connection or bridge among a group of actors and argues that brokering enhances both career mobility and promotion. In the current empirical context, brokering occurs when coauthors do not collaborate on another project in the absence of the focal engineer. The focal actor who brokers among colleagues thus occupies and exploits a structural hole. Contrast this with constraint, which occurs when an actor's alters know each other well, and when there exists redundant, dense, and cohesive interaction among the actor's contacts (Coleman 1988). Burt's classic information-control arguments for the strategic superiority of brokering are that brokers (1) gain first access to information and control of its diffusion, (2) can present different strategies to different groups (because unconnected observers will lack the opportunity to compare the strategies; see also Padgett and Ansell 1993), and (3) will be considered for an expanded set of opportunities because they will be known to a wider set of groups.

If open innovation communities were truly open there would be no opportunity to control information, and the context would remain outside the boundary conditions of Burt's arguments. Despite the strong functional and normative pressures for openness, it remains impossible for all information within open communities to be shared. Intense working relationships such as collaboration on technical publications will involve dyadic or small-group communication that will not be widely shared, nor are personal e-mails, or telephone and face-to-face conversations likely to be shared. This limitation is recognized in the working group guidelines for the IETF: "It is often useful, and perhaps inevitable, for a sub-group of a working group to develop a proposal to solve a particular problem. Such a sub-group is called a design team. In order for a design team to remain small and agile, it is acceptable to have closed membership and private meetings" (Bradner 1998, p. 18; Bradner 2004). For these reasons, open innovation communities remain within the boundary conditions of Burt's arguments.

Despite the advantages of brokering, the role can also occasion disadvantages for aspiring leaders. Researchers have long argued that cohesive social structures build trust by exerting pressure for consistent norms and reciprocity among individuals within embedded and overlapping relationships (Granovetter 1992). Cohesive networks also increase the likelihood of sanctions against individuals who violate their norms (Coleman 1988) and facilitate communication of reputation effects (Reagans and McEvily 2003). Particularly important in the current context, the potential for information control and political action inherent in brokerage structures does not sit well with cultures obsessed with the open, immediate, transparent flow of information. Consider Raymond's (1998, p. 21) appeal to open innovation communities to "be open to the point of promiscuity.... When writing gateway software of any kind," he urges, "take pains to disturb the data stream as little as possible—and *never* throw away information unless the recipient forces you to" (p. 44, italics added)! Community members understand and fear the potential for abuse and proclaim strong norms against manipulative behavior. Raymond (1998, p. 86) again: "Surreptitiously filing someone's name off a project is, in a cultural context, one of the ultimate crimes." The power of a broker who controls the flow of information to influence the attribution of credit can undermine the fundamental mechanics of open innovation communities inasmuch as attribution is a driving incentive and the basis of status. Most importantly, brokers must consequently allay concerns that they might violate norms against hiding and manipulating information. They must also contend with the inherent and previously recognized challenges of mobilizing across networks (von Krogh et al. 2003). As Coleman (1988) argues and Gould (1991) describes with reference to the Parisian revolts of 1871, mobilization will be easier if leaders can call on dense, cohesive networks, particularly if they are themselves embedded in such networks. Brokers face the added difficulty of balancing conflicting demands of simultaneous membership in multiple groups that might have differing role expectations (Podolny and Baron 1997).

Because brokerage exerts both positive and negative influences on leadership, the overall effect of the role remains an empirical question. We can still test our model of who becomes a leader, however, by developing other potentially observable implications. In particular, if another variable moderates the advantages or disadvantages of brokerage, we should observe an interaction effect. For example, if followers' trust can be gained through assurances that a potential leader will not abuse a brokerage position, we should observe a contingent and positive correlation between brokerage and assumption of a leadership position. Burt (2001b, p. 2) recognizes the problem in contexts beyond open innovation communities: "The social capital of brokerage depends on trust—since the value created by brokers by definition involves new, and so incompletely understood, combinations of previously disconnected ideas—but trust is often argued to require network closure, precisely the condition that brokers rise above."

Lerner and Tirole (2001, p. 222) attest to the fundamental importance of trust within open innovation communities. "The key to a successful leadership," they observe, "is the programmers' trust in the leadership: that is, they must believe that the leader's objectives are sufficiently congruent with theirs and not polluted by ego-driven commercial or political biases." Brennan reports that the Debian community developed a web of trust specifically to repel commercial Trojan horses: "The more deep and tightly interlinked the web of trust is, the more difficult it is to defeat" (Brennan 2003, cited in O'Mahony and Ferraro, forthcoming).[2] Davies (2003, p. 15) cites an official description of the IETF that explicitly claims that trust remains intransitive as a network relationship:

> [IETF processes] are all reliant on personal knowledge of the capabilities of other individuals and an understanding built on experience of what they can be expected to deliver, given that there are almost no sanctions that can be applied beyond not asking them to do a similar task again.... In essence, the IETF is built on a particular kind of one-to-one personal trust relationship. This is a very powerful model but it does not scale well because this *trust is not transitive*. Just because you trust one person, it does not mean that you trust (i.e., know the capabilities of and can rely on) all the people that person trusts in turn. (italics added)

If absence of trust causes potential followers to doubt a leader's motives, opportunities to observe the leader should ameliorate that doubt. We suggest that physical attendance and greater interaction with the community constitute such opportunities. Burt (2001b) proposes that, given sufficient face time and repeated interaction, brokers can overcome potential distrust of the role and turn the position to their advantage. However much electronic interaction predominates in modern open innovation communities, trust is still greatly facilitated by personal contact. Admission to the Debian community's web of trust, for example, involves personal and physical key signings, complete with government-issued identification and a handshake (O'Mahony and Ferraro, forthcoming). A short paper on how open source conventions should be run advises planners to shape the social space and maximize social interactions through physical layout, after-hours meeting places, message boards, and copious refreshments (Raymond 2000). Face time in the community will facilitate members' assessments of whether an aspiring leader might harbor blatantly political motives or be inclined to abuse the brokerage position. For these reasons, we argue that trust developed through physical interaction will increase the likelihood that a broker will advance into leadership. With low physical presence, increasing brokerage decreases the likelihood of becoming a leader. With high physical presence, increasing brokerage increases the likelihood of becoming a leader.

Hypothesis 2 (H2). *The interaction of greater physical presence and social brokerage in working relationships will increase a member's likelihood of becoming an open innovation community leader.*

In addition to the integrating role of social broker, individuals can also work across internal community boundaries and perform the integrating role of boundary spanning. The importance of boundary spanners to technical organizations and the process of innovation were emphasized in early research on the structure of engineering firms (Lorsch 1965, Allen 1977, Tushman 1977). Individuals who occupy these positions tend to hold advanced technical degrees, make the most important technical contributions, earn the respect of their colleagues, communicate with peers in other organizations, and, most important for the current study, advance into management and technical leadership positions. Boundary spanners stimulate the innovation process because formal organizational boundaries correlate with technical boundaries (Henderson and Clark 1990). These boundaries can become barriers to the flow of information due to the evolution of local dialects (Dougherty 1992) and the difficulty of transferring complex information across social boundaries (Sorenson et al. 2006). Technologies can be refined more productively within technologically focused efforts, but at the risk of becoming incremental and obsolete. Boundary spanners reduce this risk of obsolescence by gathering, interpreting, and disseminating nonredundant information across boundaries (Allen 1977, Tushman 1977).

The boundaries within open innovation communities, like those in private firms, usually correspond to the interfaces between technological subsystems (Henderson and Clark 1990). Each boundary demarcates a distinct technological area or module. The boundaries are defined directly by leaders' architectural decisions and then implemented by followers' choices of where to volunteer their efforts. Torvalds, for example, delegates responsibility for different modules to a handful of trusted lieutenants (Hamm 2005), each of whom reviews submitted code, accepts the best, and works with Torvalds and the other lieutenants to resolve issues that cross technical boundaries. Bradner (1998, p. 5; 2003b) describes how individuals who aspire to leadership in the IETF must define their technological boundaries relative to other community efforts as part of their BOF proposals: "Is there a good understanding of any existing work that is relevant to the topics that the proposed working group is to pursue ... and, if so, is adequate liaison in place?" Although technical and social boundaries correlate, brokerage and boundary spanning remain distinct roles. Nothing prevents a boundary spanner from being the only engineer, in which case she would be a social broker. Alternately, she could be one of the many engineers who collaborate across a boundary, in which case the boundary spanner is not a social broker at all.

Open innovation communities are well aware of these issues and value individuals that can span technical boundaries (Bradner 2004). Boundary spanners should be more creative and able to call on more-diverse resources, as explicitly recognized in the expectation that leaders promote "cross-pollinating between groups" (Davies 2003, p. 13). In addition to providing the integrating advantages of social brokerage, boundary spanners are in an even stronger position to control the potential for community forking, which generally occurs when alternative technical solutions attract dedicated coalitions that refuse to compromise or resolve incompatibilities. Davies and Hoffman (2004, p. 6) recount recent discussions of IETF problems that emphasize the difficulty of technological interdependence:

> The IETF has effective mechanisms for dealing with well-defined problems of limited scope. These problems are well handled in IETF Working Groups, where experts in a given technology can convene and solve the problems specific to one technology area. However, we are much less effective at resolving complex problems that affect more than one IETF area.

Boundary spanners, being more aware of other efforts across the larger community, can better negotiate the boundaries of their own groups' efforts. Individuals who span boundaries before attempting to lead are better prepared, for a variety of reasons: simple technical awareness of interdependent technologies and their inventors; expanded personal creativity; and greater likelihood of having observed or directly participated in conflict resolution. Potential followers, particularly the most competent, will be aware of technical interdependencies and value individuals who have worked across their communities' internal boundaries. Individuals with experience working across internal community boundaries are, for these reasons, more likely to become leaders.

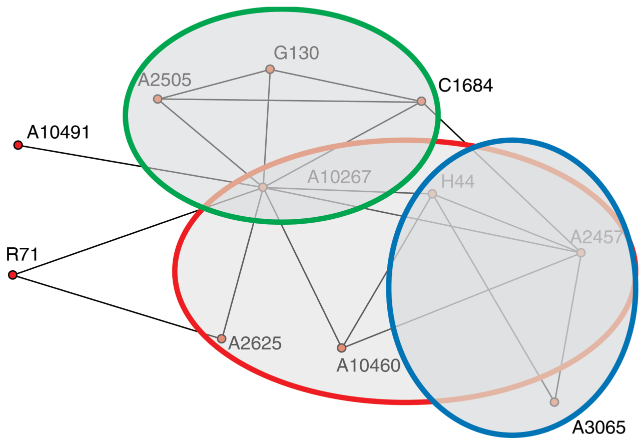
Hypothesis 3 (H3). *Internal community boundary spanning will increase a member's likelihood of becoming an open innovation community leader.*

Despite their correlation, social brokerage and boundary spanning remain distinct social positions. For example, a boundary spanner might be the only collaborative linkage between groups, and thus also be a collaborative broker; alternately, she could be but one of a number of linkages across a socially cohesive boundaries. These distinctions provide an additional opportunity to test our theory in parallel with Hypothesis 2. Our development of Hypothesis 2 held that if a social broker must overcome concerns about trust before becoming a leader, actions and positions that increase trust will exhibit a positive interaction with the social brokerage position. We now propose that the simultaneous pursuit of social brokerage and boundary spanning will exhibit a negative interaction, because spanning a boundary will diminish potential followers' trust.

Consider first an individual who spans multiple boundaries and simultaneously brokers collaborative relationships. This individual's collaborators will be less familiar with one another, not having worked together previously, and, if they contribute in different technical areas, will be less aware of and comfortable with the technologies, jargon, objectives, and reputations of collaborators in the nonlocal areas. In such a situation, concerns about trust will be magnified if an individual invests serious resources in another group's efforts, as argued by Podolny and Baron (1997, p. 676): "Each constituency grows increasingly suspicious that its needs are receiving less attention from the boundary spanner than someone else's needs." Suspicion will be greater, to the extent that a constituency is locally cohesive and insulated from other constituencies.

Now consider the opposite extreme: an individual who spans multiple boundaries and simultaneously works within cohesive collaborative relationships. This individual's collaborators will be quite familiar with one another, owing to prior working relationships. Furthermore, these cohesive relationships by definition would span technological boundaries within the community.

Figure 1 Illustration of IETF Ego Publication Network Around Nodes A10267 and H44, December 1999



Notes. The green node, H44, transitions to a first working group chair leadership appointment in this period. The red nodes represent individuals who do not transition but are at risk of working group chair appointment. The blue nodes represent existing working group chairs and are not in the first appointment risk set (they are included in network position calculations). The ellipses illustrate working group memberships. Both H44 and A10267 span working group boundaries, but H44 is in a relatively cohesive ego network. A10267, by contrast, is in a relatively strong brokering position with multiple structural holes among alters. Consistent with our hazard models, A10267 does not advance to leadership in subsequent periods.

An individual will thus be observed by mutual acquaintances on both sides and be trusted by both sides to resolve technological and organizational boundary conflicts. Such an individual is more likely to satisfy conflicting goals and expectations among collaborating organizations. Cohesive boundaries also signify active technical areas of innovative search and attempts to resolve technological interdependencies. Taken together, these arguments imply that leaders will emerge from among those who make public contributions at the boundaries of socially cohesive and strongly interacting technical organizations within a community.

Hypothesis 4 (H4). *The interaction of greater social brokerage and internal community boundary spanning will decrease a member's likelihood of becoming an open innovation community leader.*

Figure 1 illustrates the hypotheses with two ego networks taken from December of 1999.

Evidence

We draw data primarily from the published proceedings of the IETF's triannual meetings from its inception in 1986 until 2002, and from the organization's RFC (request for comments) publication series over the same time period.[3] Each proceeding represents a snapshot of the IETF at a given time, complete with a list of registered conference attendees (including their e-mail affiliations) and archived charters for each of the active working groups indicating current chairs and area directors. It can thus be determined from the proceedings when individuals first attended an IETF conference and if and when they began to lead a working group. RFC publications indicate, within a designated time frame, which engineers are publishing technical documents and whether—and with which other engineers and established working groups—they are collaborating. We supplemented these archival data with extensive interviews of community leaders (Bradner 2003a, b, 2004; Kaufman 2003) at their places of employment and at the 56th IETF meeting in San Francisco, March 16–21, 2003.

Exactly 15,465 individuals attend the IETF meetings over the course of the study and are observed at four-month intervals. Because the first three years' of conference data were truncated due to leading independent variable creation, the analysis data span 1989–2002. Subjects enter the risk set on first attending an IETF meeting or in the first trimester of 1989 if community involvement occurred before 1989. Subjects exit on the first working group chair appointment or are treated as right censored.[4] For variables relating to the overall state of the community, such as the number of incumbent working group chairs at any point in time, we utilize the full database. Because coauthorship patterns form the basis of the longitudinal social networks we construct, and social brokerage can only occur if an individual writes two or more publications, we restrict analysis of leadership appointment to the 610 individuals who publish two or more documents prior to working group chair promotion.[5] The logic behind this restriction is that observable variation on brokerage or boundary spanning can only occur when an actor has two or more opportunities on which to interact with other individuals (by definition, all coauthors on a single draft are cohesively

Table 1 Descriptive Statistics for Network Risk Set ($N = 5{,}066$)

Variable	Mean	Std. dev.	Min	Max
Incumbent WG chairs	153.77	51.38	16.00	230.00
University affiliation	0.13	0.33	0.00	1.00
Government affiliation	0.04	0.19	0.00	1.00
Affiliated WG chairs	19.27	35.96	0.00	192.00
Executive tie	0.57	0.77	0.00	5.00
U.S. patents	0.37	1.38	0.00	17.00
Degree	6.97	6.19	0.00	39.00
WG contributor	0.82	0.38	0.00	1.00
(*ln*) *IETF publications*	1.34	0.33	1.10	3.09
Presence (*uncorrected*)	4.53	2.87	0.00	9.00
Miles	9.75	0.53	0.00	11.13
Social brokerage[a]	0.00	0.27	−0.66	0.46
Spans 1 boundary[b]	0.23	0.42	0.00	1.00
Spans 2+ boundaries[b]	0.05	0.22	0.00	1.00

Notes. WG = working group.
[a]Mean centered to facilitate interpretation of interaction effects (Friedrich 1982).
[b]The omitted reference category is Zero WGs.

Table 2 Bivariate Correlation Matrix for Network Risk Set ($N = 5{,}066$ Observations)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1 *Incumbent WG chairs*	1.00													
2 *University affiliation*	−0.24	1.00												
3 *Government affiliation*	−0.10	−0.08	1.00											
4 *Affiliated WG chairs*	0.27	−0.12	−0.09	1.00										
5 *Executive tie*	0.17	−0.05	0.01	0.16	1.00									
6 *U.S. patents*	0.05	−0.07	−0.03	0.06	0.07	1.00								
7 *Degree*	0.41	−0.11	−0.02	0.25	0.47	0.08	1.00							
8 *WG contributor*	0.26	−0.08	0.02	0.15	0.09	−0.03	0.31	1.00						
9 (*ln*) *IETF publications*	−0.02	0.04	−0.01	0.06	0.23	0.03	0.27	0.06	1.00					
10 *Presence* (*uncorrected*)	0.16	−0.04	0.06	0.18	0.05	0.02	0.19	0.24	0.21	1.00				
11 *Miles*	0.43	−0.05	−0.08	0.17	0.09	0.01	0.21	0.18	0.07	0.34	1.00			
12 *Social broker (-constraint)*	0.27	−0.01	−0.06	0.15	0.23	0.08	0.56	0.19	0.16	0.14	0.17	1.00		
13 *Spans* 1 *boundary*	0.13	0.03	−0.01	0.04	0.06	0.02	0.18	0.26	0.14	0.18	0.11	0.21	1.00	
14 *Spans 2+ boundaries*	0.10	0.02	0.03	0.06	0.17	0.00	0.32	0.11	0.39	0.16	0.10	0.20	−0.13	1.00

Note. WG = working group.

tied). We discuss solutions to the biased-sample problem in the variables and results sections.

Variables

Tables 1 and 2 describe the 610 subjects who published prior to working group chair appointment or censoring, observed over 5,066 trimesters. The dependent variable *Leader* is a categorical indicator that a subject appears as a working group leader at the end of the current observation period. (*ln*) *IETF publications*, which counts total IETF technical publications authored or coauthored in the prior three years, is our main measure of individual human capital. These documents also reveal working group affiliation. *Spans* 1 *boundary* indicates that an engineer published within two working groups; *Spans* 2+ *boundaries* indicates that a subject published in three or more working groups. Although use of a count variable for the number of working group memberships instead of breakout into mutually exclusive categories returned similar estimates, we prefer the indicators due to the qualitative differences between no membership, membership in one group, and boundary spanning.

We created social network measures based on coauthorship by constructing 51 social networks, one for each trimester with a three-year window on prior linkages, with valued symmetrical ties between all IETF authors. We calculated the following network position variables in UCINET version 6 (Borgatti et al. 2001). *Social Brokerage* measures the opposite of individual constraint, where constraint measures the extent to which ego has ties with few alters and those alters have ties between them (Burt 1992). We enter the negative of constraint (Equation (1)) as our measure of brokering. In Equation (1), p_{ij} is the proportion of i's relations directly invested in contact j. $\sum_q p_{iq}p_{qj}$ is the portion of i's direct connections invested in contacts q who are in turn invested in contact j. The sum in the parentheses is the proportion of i's direct and indirect connections that are invested in contact j. The sum of squared proportions over all contacts j is a measure of individual i's network constraint.

$$\text{Constraint}_{\text{actor } i} = \sum_j c_{ij}, \quad c_{ij} = \left(p_{ij} + \sum_q p_{iq}p_{qj}\right)^2, \quad q \neq i, j. \quad (1)$$

Presence measures the extent of an individual's physical involvement with the IETF community over the prior three years. A physical presence measure that simply counted a subject's conference attendance poses a serious endogeneity problem, in that we cannot directly observe an individual's aspirations for leadership from archival records. We also cannot observe an employer's strategy to coopt the community by influencing the leadership processes, and these unobserved aspirations and strategies would very likely influence conference attendance. Endogeneity violates the standard assumption of no correlation between independent variables and the error term (Greene 1993, p. 285). Unobserved ambition also specifically poses a problem for testing the hypothesis that the effect of network position interacts with presence. To address this problem we employ an instrumental variable approach (Greene 1993, p. 603). We first developed a proxy variable that correlates with attendance but not with the probability of becoming a leader. *Miles* accounts for total distance, in miles, between an engineer's home and all IETF meetings during the prior three years.[6] The instrument implicitly argues that attendance will be correlated with both the cost of attendance and an individual's desire to lead. Given that the location of each conference is probably set without consideration of which members desire to lead, however (and indeed, the conference is moved around intentionally, both within North America and abroad, to facilitate attendance by a wider variety of members; see Bradner

2003b), it should be exogenous to the outcome variable of leadership. To create the instrument, we assigned each subject to a latitude and longitude based on the phone number reported in the conference registration list, then calculated (using the STATA sphdist ado subroutine) spherical distance between the subject and each meeting's latitude and longitude (Sorenson and Stuart 2001). We then regressed three-year conference attendance on *Miles* and all other independent variables, and recovered predicted values for *Presence*. For identification purposes, *Miles* was not included when estimating appointment to leadership. In the first-stage estimation, the predicted variable is purified of the endogenous element that is correlated with the disturbance term. The R^2 for the regression of these variables on attendance (depending on the final specification) ranged from 22% to 24%. Given the high amount of explained variance, the instrument should not be vulnerable to finite sample and weak instrument bias (Bound et al. 1995).

The variable *Nonselection hazard* addresses the nonrandom sample problem. Because coauthors work, by definition, within a cohesive relationship, and because brokerage requires at least two different coauthors, we calculate each of the boundary-spanning and social network measures only for cases in which a subject has two or more publications. Restricting analysis to such a subset produces a nonrandom sample and possibly biases results by neglecting the factors that led to inclusion in the sample in the first place. A standard solution to selection bias is to first estimate a selection parameter on the full data, and then include this parameter in models that use the restricted sample (Heckman 1976). Although not as well developed in the hazard model context, preliminary work (Boehmke et al. 2006) indicates that sample selection can also bias hazard model estimates. To address this problem, following Heckman we estimate the probability that any of our 15,465 at-risk engineers publishes two or more IETF documents. All independent and control variables except those related to boundary spanning and social network position are included.

Finally, the models include a series of control variables. *Entry cohort* is a set of 15 dummy variables based on the year of entry into the IETF community. *Incumbent WG chairs* is a count of incumbents (those not at risk of a first leadership position) holding chair positions and is intended to gauge the extent to which the community currently fulfills its leadership needs. *University affiliation* and *Government affiliation* are dummy variables for employer type. *Affiliated WG chairs* counts the number of extant WG chairs from the same employer, indicating employer commitment to IETF participation or attempts to influence the direction of community evolution (Wade 1995). *Executive tie* measures the number of collaborative working relationships with a current AD within the past three years to control for reliance on personal ties with individuals who have control over the appointment process. *U.S. patents* obtained by a subject during the prior three years identifies individuals with an external reputation and propensity to publish.[7] *Degree* measures, on the basis of all drafts published in the previous three years, the number of individuals with whom an engineer has published a draft. Finally, the indicator variable *WG contributor* records subjects that have participated in an existing working group within the prior three years. Such individuals are more likely to lead because they have already demonstrated affinity for social interaction and gained greater exposure to organizational processes, leadership models, and opportunities for informal leadership.

Models

We estimate hazard models of the duration T from first attendance at an IETF conference until appointment as a working group leader. Rate models are more appropriate than choice models because members do not compete for a limited number of positions. Instead, they create positions based on their understanding of technical issues and opportunities that confront the Internet community. Rate models also accommodate censored observations. Equation (2) defines the instantaneous hazard of appointment as a leader. T represents the time between a member's first attendance at an IETF meeting and (possible) appointment as a working group leader; $r(t)$ is the instantaneous hazard of making the transition from individual member to appointment as a working group leader. The data change at trimester frequency, basically every four months, to correspond to the frequency of the IETF conferences.

$$r(t) = \lim_{t' \to t} \frac{\Pr(t \leq T < t' \mid T \geq t)}{t' \to t}. \tag{2}$$

We estimated semiparametric Cox models (Equation (3)) to avoid making parametric assumptions about the form of duration dependence in the underlying hazard rate (Cox 1972). Given the relatively few events, we used the Breslow method for handling ties. The model's hazard rate is the product of an unspecified baseline rate, $h(t)$, and an exponential term that includes covariates X. The Cox model assumes, however, that the proportional hazards remain constant over time. We tested this assumption both graphically and with a Schoenfeld (1982) test. Neither inspection nor the statistical test indicated a significant relationship between time and model residuals (p value < 0.23). Piecewise exponential models returned very similar substantive results. We estimated all models with STATA version 8.

$$r(t) = h(t)e^{(\beta X)}. \tag{3}$$

Table 3 Cox Hazard Models of Appointment as Working Group Chair Among IETF Participants, 1989–2002

	(1)	(2)	(3)	(4)	(5)	(6)	(7) w/out IV	(8)	(9)	(10) Imputed network values
Entry cohorts	(15,)	(15,)	(15,)	(15,)	(15,)	(15,)	(15, +)	(15,)	(15, +)	(15, +)
Incumbent WG chairs	−0.039	−0.039	−0.037	−0.041	−0.042	−0.040	−0.043	−0.041	−0.039	−0.019
	(0.010)**	(0.009)**	(0.009)**	(0.009)**	(0.009)**	(0.009)**	(0.009)**	(0.009)**	(0.009)**	(0.003)**
University affiliation	−0.116	−0.158	−0.127	−0.241	−0.240	−0.161	−0.099	−0.160	−0.145	−0.017
	(0.279)	(0.275)	(0.277)	(0.276)	(0.280)	(0.283)	(0.282)	(0.283)	(0.284)	(0.133)
Government affiliation	−0.293	−0.265	−0.309	−0.325	−0.327	−0.415	−0.683	−0.434	−0.303	−0.319
	(0.606)	(0.542)	(0.559)	(0.541)	(0.542)	(0.560)	(0.531)	(0.561)	(0.567)	(0.241)
Affiliated WG chairs	−0.002	−0.002	−0.002	−0.001	−0.001	−0.002	−0.004	−0.002	−0.005	0.003
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.002)
Executive tie	−0.326	−0.333	−0.315	−0.315	−0.315	−0.305	−0.228	−0.303	−1.464	−0.443
	(0.166)*	(0.152)*	(0.156)*	(0.152)*	(0.151)*	(0.155)*	(0.145)	(0.156)	(0.298)**	(0.155)**
U.S. patents	−0.047	−0.050	−0.049	−0.035	−0.030	−0.024	−0.011	−0.020	−0.041	−0.012
	(0.090)	(0.089)	(0.089)	(0.089)	(0.087)	(0.087)	(0.088)	(0.087)	(0.088)	(0.041)
Degree	0.003	−0.021	−0.045	−0.022	−0.023	−0.034	−0.020	−0.035	−0.036	0.040
	(0.023)	(0.025)	(0.026)	(0.025)	(0.027)	(0.028)	(0.027)	(0.028)	(0.028)	(0.022)
WG contributor	0.918	1.120	1.279	0.773	0.891	0.997	0.422	0.919	0.804	0.628
	(0.555)	(0.371)**	(0.391)**	(0.362)*	(0.376)*	(0.386)**	(0.329)	(0.402)*	(0.394)*	(0.314)*
(*ln*) *IETF publications*	1.338	1.399	1.301	0.973	0.977	0.862	0.581	0.865	0.957	1.051
	(0.562)*	(0.345)**	(0.404)**	(0.352)**	(0.346)**	(0.386)*	(0.295)*	(0.386)*	(0.383)*	(0.113)**
Presence	−0.117	−0.158	−0.172	−0.156	−0.147	−0.129	0.220	−0.065	−0.228	−0.057
	(0.320)	(0.148)	(0.188)	(0.163)	(0.153)	(0.192)	(0.045)**	(0.213)	(0.214)	(0.147)
Social brokerage		1.110	1.655		1.260	2.235	1.075	2.171	1.974	1.846
		(0.404)**	(0.443)**		(0.443)**	(0.524)**	(0.446)*	(0.526)**	(0.530)**	(0.455)**
Social broker ∗ Presence			0.794			1.023	0.090	1.042	0.913	0.560
			(0.259)**			(0.295)**	(0.130)	(0.301)**	(0.309)**	(0.260)*
Spans 1 boundary				0.734	0.690	0.725	0.472	0.790	0.838	0.855
				(0.260)**	(0.258)**	(0.274)**	(0.233)*	(0.282)**	(0.282)**	(0.327)**
Spans 1 boundary × Social broker					−1.970	−3.662	−1.893	−3.515	−3.242	−3.120
					(0.915)*	(1.035)**	(0.939)*	(1.046)**	(1.034)**	(0.907)**
Spans 1 boundary × Presence								−0.141	−0.171	−0.147
								(0.191)	(0.191)	(0.202)
Spans 2+ boundaries				1.534	1.701	1.883	1.341	1.867	1.848	2.061
				(0.408)**	(0.485)**	(0.505)**	(0.448)**	(0.731)*	(0.751)*	(0.605)**
Spans 2+ boundaries × Social broker					−2.503	−5.543	−2.519	−5.504	−5.135	−3.670
					(1.793)	(2.041)**	(1.822)	(2.040)**	(2.025)*	(1.868)*
Spans 2+ boundaries × Presence								−0.053	−0.011	−0.522
								(0.328)	(0.333)	(0.370)
Nonselection hazard									−1.263	
									(0.262)**	
Log likelihood	−579.48	−575.50	−570.71	−571.86	−567.03	−560.83	−553.00	−560.56	−547.15	−3,653.29
Observations	5,056	5,056	5,056	5,056	5,056	5,056	5,056	5,056	5,056	120,193
Individuals	610	610	610	610	610	610	610	610	610	15,465
Events	118	118	118	118	118	118	118	118	118	472

Notes. Standard errors in parentheses. WG = working group.
*Significant at 5%; **significant at 1%; + jointly significant at 5%.

Figure 2 Publication Spline for Full-Risk Set (15,466 Subjects Over 243,405 Observation Periods)

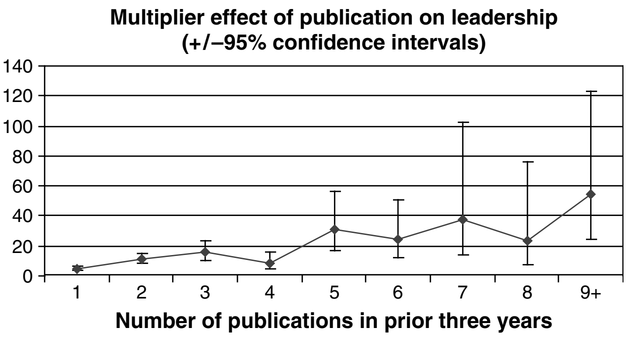
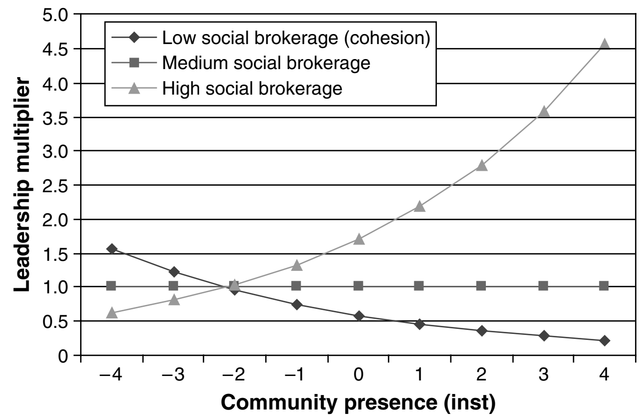


Figure 3 Interaction Between Social Brokerage and Community Presence



Note. (Mean + / − one standard deviation of brokerage from Model 9.

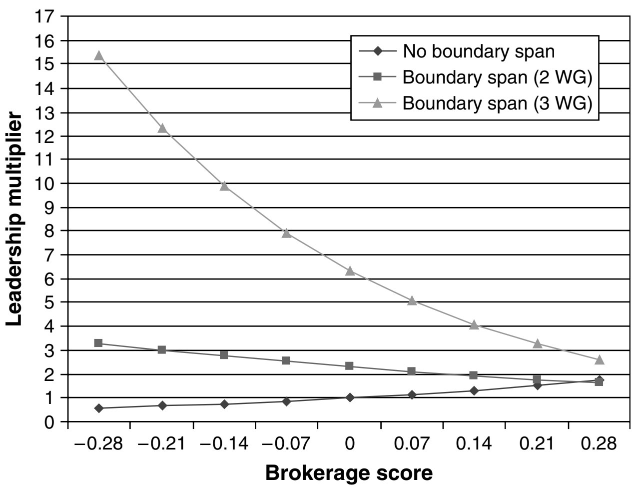
Results

Models 1 through 5 in Table 3 enter explanatory variables individually and then with interactions. We maintain *IETF Publications* and controls in all specifications both because publications has such a powerful effect and because we want to guard against the possibility that differences in social network position are simply the result of the number of times (i.e., number of publications) we observe a subject. Models 6 through 10 present full specifications. Model 7 illustrates results without use of an instrumental variable for presence, while Model 8 illustrates the lack of effect for a boundary-spanning and presence interaction. Model 9 illustrates that the effects remain robust in the presence of a nonselection hazard, and Model 10 establishes that the results hold when estimations are run using imputed network values for the full sample. Before analyzing Table 3, we present an unconditional analysis of the effects of technical contribution, as measured by publications for the entire risk set of 15,466 subjects observed over 243,405 time periods. Figure 2 dramatically illustrates the importance of technical contribution on leadership and supports a logged specification in the hazard models. The effect sizes are quite large; the first publication increases the likelihood of becoming a leader by 143%.[8] The figure indicates that for three publications the point estimate increases 1,590% relative to having none. Hence, the results support Hypothesis 1, which argues for the importance of technical contribution for leadership.

The models also support the hypothesized relationships for both brokerage and boundary spanning. We argued that the net effect of brokerage remains unpredictable due to positive and negative influences of the position on the perceptions of potential followers. We proposed in Hypothesis 2, however, that physical attendance at meetings would ameliorate negative perceptions, such that attendance and interaction should demonstrate a positive interaction. All the models support this interaction effect. Model 9 (from which we draw all interpretations of effect magnitude) indicates that an individual who simultaneously increases attendance and social brokerage position by one standard deviation each will increase the likelihood of becoming a leader by 33%. The one standard deviation increase in brokerage alone increases the likelihood of leadership by 70%, indicating that it has a net positive influence on the chance of leadership. (Instrumented attendance has no impact.) Figure 3, which plots the intersection effect for low, medium, and high values of social brokerage, indicates that most of the positive effect of the brokerage position accrues to individuals who attend many conferences. Conversely, a cohesive ego network structure is preferred for those who do not attend conferences and lack a strong physical presence. Hence, the models and Figure 3 support Hypothesis 2, that the positive effects of social brokerage will be enhanced by increased trust, as developed by physical presence in the community.

Although social brokerage and its interaction with attendance demonstrate strong positive effects, the benefits of boundary spanning are greater. Contributing to two working groups correlates with a 131% increase in leadership (over and above the 123% benefit of contributing within any working group). Contributing to three or more working groups correlates with a 535% increase. Simultaneous brokerage and membership in two groups correlates with a negative interaction effect of 59%; in three or more groups, a negative effect of 75% (although the overall effect remains positive, due to the stronger first-order effects). Figure 4 illustrates the negative consequences of simultaneous brokerage and boundary spanning. With increased brokerage, benefits increase for individuals who work within a group, but decrease sharply for those who straddle boundaries. Hence, the models support Hypothesis 3, that boundary spanning increases the likelihood of leadership, and Hypothesis 4, that simultaneous brokerage decreases this positive effect. Further supporting our theory, Model 9 does not demonstrate an even marginally significant interaction between boundary spanning and the instrumented attendance variable. Although prediction and

Figure 4 Interaction Between Technical Boundary Spanning and Social Brokerage (Model 9)



Note. WG = working group.

observation of a null result do not test theory (hence, are not included in our hypotheses), the lack of significance provides further evidence of our contention that fellow community members perceive the two roles differently.

To directly compare the different structural strategies and their influence on leadership, imagine two individuals: first, a broker whose measure is one standard deviation above the variable mean and contributes to only one working group; and second, a boundary spanner whose measure is at the brokerage mean and contributes to two groups. Both strategies work approximately equally well if they attend an average number of conferences. The broker is then 193% more likely to become a leader (70% for brokerage and 123% for contributing within a working group) compared with a 254% increase for the boundary spanner (131% for spanning one boundary and 123% for contributing within a working group). If both individuals increase their attendance at conferences by one standard deviation, the advantage remains with the boundary spanner, because the broker gains only 33% from the interaction of brokerage and attendance. The advantage shifts decisively in favor of the boundary spanner, however, if she spans additional boundaries; assuming all else stays equal, she is 658% more likely to become a leader (535% for spanning two or more boundaries and 123% for contributing within a working group). While it appears to be moderately effective to bind a single working group with itself, it appears to be even more effective to connect three groups within the larger community. Moderately ambitious individuals should therefore look for local integration opportunities; very ambitious individuals should seek out the intersections of communitywide activity. Furthermore, for individuals who have already authored many drafts, it would be easier to span an additional boundary, given the incremental impact of a single additional collaboration on a deeply embedded structural position.

With respect to control variables, collaboration with an AD subtracts from the likelihood of leadership, a result consistent with community-espoused norms. With the exception of the incumbent working group population, none of the controls demonstrate significance in the networked risk set. Surprisingly, degree never demonstrates a significant effect in any model; working with additional coauthors does not correlate with leadership. We attribute this to a dilution effect of the credit that occurs when an individual coauthors with too many others. Discussion with IETF members lent validity to this interpretation. Finally, the selection term is negative and significant, indicating that individuals who are most likely not to be selected are also less likely to become leaders.

Finally, Model 10 in Table 3 takes an alternative tack at addressing the issue of the small subsample used for primary analysis. Recall that very few IETF participants, just 610 out of 15,465, have two or more publications prior to leadership appointment or censoring. While it is therefore true that many engineers advance to leadership without observed publications, and hence without an observed network position, further investigation reveals that 75% of all working group chairs eventually publish, and 100% are regular conference attendees. There is also anecdotal evidence that most of the working group chairs that never publish have, in fact, attempted to do so. Collaborative networks are thus a pervasive feature of this organization, but the data are limited by a nonrandom propensity to observe collaborative ties. While Model 9 addresses this problem with the nonselection hazard term, Model 10 instead utilizes imputed network values for all subjects. Imputed values for *Degree, Boundary Spanning*, and *Social Brokerage* are derived from regressions using cumulative conference attendance, time at risk, type of employer, location outside the United States, cumulative distance from recent conferences, publishing by coemployees, working group chair selection among coemployees, patenting, and year dummies. The results in Model 10 compare very favorable to all other full specifications.

Discussion

While the models demonstrate the importance of both human and social capital in the attainment of leadership positions, it is important to demarcate their theoretical and empirical limitations. We generalize the theoretical implications for brokerage with caution, because the social context differs from that of private firms (Burt 1992, 1997, 2001a; Podolny and Baron 1997). This research considers how individuals who participate in voluntary communities emerge as leaders, as opposed to how individuals in for-profit firms are promoted to management. Aspiring managers leverage brokerage positions by controlling information, resources, and, ultimately, the perceptions of those who do the promoting. Such efforts are facilitated by focusing on maximizing uncertainty among a few key decision makers. In a voluntary community, however, leadership requires mobilizing efforts across a large (and often unknown) population of peers, friends, and colleagues. Members who gain reputations as controlling individuals who actively manage perceptions will probably face informal sanctions (von Hippel and von Krogh 2003). As the results indicate, aspiring leaders in voluntary communities need to offset these potentially negative perceptions with additional assurances that they intend to benefit and bind the community as a whole.

Caution should also be exercised before generalizing the results to all open innovation communities, particularly those that lack physical interaction. Our results remain conditional on attendance (although our data indicated that every IETF leader attended at least one meeting). The issues of trust and forking remain salient for all open innovation communities and warrant additional research and elaboration. Concern is also merited with respect to the monotonically increasing environmental munificence during the period of observation—the IETF (and the industries that supported it) expanded during all but the last few years of the study. Continuously expanding resources and leadership opportunities facilitate the dispersion of benefits, enable coalition building, and probably ease distrust. The IETF's recent self-examination (Wasserman 2003) admits as much, even as it attempts to grapple with increasingly divisive processes and communal fissures. Elaborating this relationship between munificence and positional efficacy remains an opportunity for future research within communities and firms. Community growth and membership change also highlights an opportunity to study the demographic dynamics of voluntary communities. For example, when do current members or leaders attract demographically extreme recruits? When this occurs, the possibility of forking or disruptive selection (McPherson et al. 1992) should greatly increase.

Empirical limitations also deserve attention. The dependent variable of appointment does not measure organization of a BOF meeting. Ideally, we would also observe which community members attempted to organize a meeting, and then, conditional on that successful organization, which members advance to leadership. Archived BOF meeting documentation remains inconsistent, however, and more complete for meetings that become working groups. Rather than introduce a nonquantifiable success bias into our models, we developed the instrumental variable based on attendance, which controls for members' aspirations and attempts to lead. Without use of the instrument, the influence of technical contribution and structural positions become much weaker, and mere attendance appears to be a very effective route to leadership. Inferences drawn from this model would be wrong, however, because it does not control for the self-selection of aspiring leaders to attend. This incomplete specification would therefore downplay the importance of technical contribution and miss the nuanced influence of brokerage and its interaction with physical presence. The results should still be regarded with caution, despite the strength of the instrument, given that the second stage estimates a nonlinear Cox model, and the use of instruments with nonlinear models remains controversial (see Bowden and Turkington 1981, Box-Steffensmeier and Jones 2004; cf. Hausman et al. 1995). Limitations notwithstanding, the research illustrates a method for analyzing social networks with archival data. Of greater importance to future research in social networks, the use of an endogeneity instrument enables control for strategic behavior in the development of relationships. To date, network analysis has mainly ignored the problem of separating the causal effects of position from the intentions of the individual who may have consciously and strategically created the position. Although our analyses used archival data, the more common use of questionnaire data in network analysis could also benefit from the explicit design and gathering of instrumental variables.

In addition to the methodological contributions, the results contribute to the larger network literature. Given that the quantitative research on the benefits of brokerage has not previously distinguished brokerage from boundary spanning, it remains unclear which position confers the bulk of previously demonstrated benefits. At least in the IETF, leaders are more likely to emerge from cohesive boundaries than from isolated brokerage positions. The results present the only analysis (to the best of our knowledge) that separates the influence of human and social capital on promotion. They highlight the previously unexamined differences between brokerage and boundary-spanning positions and demonstrate that the role conflict previously associated with both positions (Podolny and Baron 1997) appears to be associated only with brokerage.

These findings should be elaborated in open and commercial contexts with dynamic analyses, because boundaries change as often as the relationships that straddle them. The current research studied boundaries that emerge from leaders' architectural and individuals' voluntary choices, remain relatively informal, and dissipate with completion of the group's work. Firm boundaries are more formal, longer lived, and probably support the transformation and hardening of technological boundaries into social boundaries. In private firms, brokerage and boundary spanning probably correlate more strongly, such that individuals who span a cohesive boundary should become increasingly rare over time. If this were true, then the combination of boundary spanning and cohesion might be an even more effective promotion strategy in private firms than in open innovation communities. Such a strategy might be conscientiously directed by a senior manager, for example, by assigning a capable engineer to a critical boundary-spanning project and discouraging brokerage of her colleagues. The capable engineer would then be an ideal candidate for promotion, in either of the groups or in the larger organization, due to her visible and trusted contributions in managing the organization's technological interdependencies.

Conclusion

The IETF archives afford a rare and detailed look at the social, technical, and political dynamics within an open innovation community. We used the entire history of 16 years of meetings and work collaborations to trace the emergence of leaders within the organization. After demonstrating the importance of human capital, as measured by technical contributions, the models also demonstrated the similarly large importance of social capital, as measured by brokerage and boundary spanning of collaborative relationships. That both brokering and boundary-spanning roles greatly increase the likelihood of leadership points to the importance of social positions that can unite open innovation communities. We argued that trust does not come easily to community members who fear cooptation by commercial interests or forking over technical disagreements. Because brokers by definition contrive less cohesive and less trusting contexts, the probability that they will assume leadership roles remains highly contingent on building trust with community members. We argue that aspiring leaders can build trust through physical attendance and, consistent with this argument, find a positive interaction with physical attendance. Also consistent with our emphasis on trust in open innovation communities, brokerage and boundary spanning demonstrated a negative interaction, indicating that brokers who span boundaries remain at a disadvantage. While brokerage alone demonstrates positive influence on becoming a leader, boundary spanning demonstrates a much stronger effect. Finally, we did not observe a contingent relationship between boundary spanning and attendance. Our results emphasize the importance of intermediary and integrating roles—for brokers within technological boundaries, and for boundary spanners across cohesive technological boundaries.

Open innovation communities represent a new and powerful social context in which to generate knowledge and advance technology (von Hippel and von Krogh 2003). Even if they represent merely a recombinant hybrid between the institutions of science and norms of communities of technological practice, their results to date and potential to shape the rate and direction of technological change remain impressive. Increasing technological complexity and interdependence put these communities at perpetual risk of forking and balkanization. Without strong leaders they inevitably splinter and fail. Leaders, in binding their communities together, foster the integration needed to forestall these eventualities. The contribution of this research is to elucidate how leaders emerge from the interactions of ability, position, and strategy.

Acknowledgments

The authors are very appreciative of the help of Scott Bradner and many other members of the Internet Engineering Task Force in completing this study. They would also like to thank two anonymous reviewers, Jóhanna Birnir, Ronald Burt, Brent Goldfarb, Michael Hannan, Eric von Hippel, Jerry Kim, Karim Lakhani, Siobhan O'Mahony, Renée Richardson, Tim Simcoe, John Simon, Mike Tushman, Steve Wheelwright, and, in particular, Woody Powell and Olav Sorenson for their feedback, and the Harvard Business School Division of Research for support. The work has also benefited from seminars given at Harvard and Stanford Universities and the Universities of Chicago, Maryland, and California, Berkeley.

Endnotes

[1] "One of the principal differences between the IETF and many other standards organizations is that the IETF is very much a bottom-up organization. It is quite rare for the executive leaders within the IETF, the IESG (Internet Engineering Steering Group) or the IAB (Internet Architecture Board) members, to create a working group on their own to work on some problem that is felt to be an important one. Almost all working groups are formed when a small group of interested individuals get together on their own and then propose a working group to an Area Director" (Bradner 1999, p. 49). It is much the same with other open innovation communities. "Today, an open source software development project is typically initiated by an individual or a small group with an idea for something interesting they themselves want for an intellectual or personal or business reason" (von Hippel and von Krogh 2003, p. 211). This process for Linux was described thus: " 'The lieutenants get picked—but not by me,' explains Torvalds. 'Somebody who gets things done, and shows good taste—people just start sending them suggestions and patches' " (Hamm 2005, p. 66). See also Rosenkopf et al. (2001).

[2] Open innovation communities use the term *Trojan horse* to refer to the threat of corporate manipulation and malicious sabotage, such as the deliberate planting of insidious bugs (DiBona et al. 1999).

[3] In 1987, 1989, and 1990 the IETF held four meetings each year.

[4] We do not estimate repeated events models because prior leadership experience would strongly influence the probability of an additional appointment. For example, a prior leader, particularly if successful, would probably be more trusted and benefit less from the publication of standards drafts or advantageous network positions.

[5] We investigated other potential measures of relationships, but found each wanting: e-mail interactions and working group attendance lists were inconsistently archived, particularly for earlier interactions; e-mail contributions would be skewed in favor of communicative extroverts; interviewees reported that individuals often had no knowledge or relationship with affiliated (that is, same-employer) colleagues. In contrast, standards

collaborations and the working group publication affiliation measure strong, accurately measured, and very important social relationships in this community.

[6]Zero values for miles occur when a member attends her first conference in her home town.

[7]In order to explore the demographics of IETF members, we linked the IETF data back into the patent database and performed some basic regressions on variables from Fleming et al., forthcoming. Relative to the entire population of inventors of U.S. patents, IETF members with patents had significantly more patents, collaborators, and number of assignees (that is, their patents were owned by a greater number of organizations—this correlates most strongly with personnel movement and boundary spanning). IETF members demonstrated no significant difference in their brokerage of other patented inventors, future prior art citations to their patents (a widely used indicator of quality), references to the nonpatent (generally scientific) literature, and proclivity to repeatedly collaborate with the same colleagues. Hence, IETF members who are inventors and who patent appear to publish more, and work with more people and span more organizations, than do non-IETF members who are inventors and who patent.

[8]As calculated by $100(e^{(0.887*1)} - 1) = 143$.

References

Abbate, J. 1999. *Inventing the Internet*. MIT Press, Cambridge, MA.

Allen, T. 1977. *Managing the Flow of Technology*. MIT Press, Cambridge, MA.

Boehmke, F. J., D. S. Morey, M. Shannon. 2006. Selection bias and continuous-time duration models: Consequences and a proposed solution. *Amer. J. Political Sci.* **50**(January) 192–207.

Borgatti, S., M. Everett, L. Freeman. 2001. *UCINET 6.0 Version 1.00.* Analytic Technologies, Natick, MA.

Bound, J., D. Jaeger, R. Baker. 1995. Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *J. Amer. Statist. Assoc.* **90**(430) 443–450.

Bowden, R., D. Turkington. 1981. A comparative study of instrumental variables estimators for nonlinear simultaneous models. *J. Amer. Statist. Assoc.* **76**(376) 988–995.

Box-Steffensmeier, J. M., Bradford S. Jones. 2004. *Event History Modeling: A Guide for Social Scientists.* Cambridge University Press, New York.

Bradner, S. 1996. The Internet standards process—Rev. 3. RFC 2026.

Bradner, S. 1998. IETF working group guidelines and procedures. RFC 2418.

Bradner, S. 1999. The Internet engineering task force. C. DiBona, S. Ockman, M. Stone, eds. *Open Sources: Voices from the Open Source Revolution*. O'Reilly, Sebastapol, CA, 47–52.

Bradner, S. 2003a. IETF structure and Internet standards process. Presentation *57th IETF Conv.*, Vienna, Austria.

Bradner, S. 2003b. Personal communication. Cambridge, MA (March 24).

Bradner, S. 2004. Personal communication. Cambridge, MA (August 28).

Bradner, S. 2006. Personal communication. Cambridge, MA (April 19).

Brennen, A. 2003. Gnu P G Keysigning Party HOWTO. http://cryptnet.net/fdp/crypto/keysigning-party/en/keysigning_party.html.

Broothaerts, W., H. Mitchell, B. Weir, S. Kalnes, L. Smith, W. Yang, J. Mayer, C. Roa-Rodriguez, R. Jefferson. 2005. Gene transfer to plants by diverse species of bacteria. *Nature* **433** 629–633.

Burt, R. 1992. *Structural Holes*. Harvard University Press, Cambridge, MA.

Burt, R. 1997. The contingent value of social capital. *Admin. Sci. Quart.* **42** 339–365.

Burt, R. S. 2000. The network structure of social capital. *Res. Organ. Behav.* **22** 345–423.

Burt, R. 2001a. Structural holes versus network closure as social capital. N. Lin, K. Cook, R. Burt, eds. *Social Capital: Theory and Research*. Aldine de Gruyter, New York.

Burt, R. 2001b. Bandwidth and echo: Trust, information, and gossip in social networks. A. Casella, J. Rauch, eds. *Networks and Markets: Contributions from Economics and Sociology*. Russell Sage Foundation, New York.

Coleman, J. S. 1988. Social capital in the creation of human capital. *Amer. J. Sociol.* **94** S95–S120.

Cox, D. 1972. Regression models and life-tables (with discussion). *J. Roy. Statist. Soc. Ser. B* **34** 187–220.

Dalle, J., P. David. 2003. The allocation of software development resources in "open source" production mode. Working paper, Stanford University, Stanford, CA.

Davies, E. 2003. IETF problem statement. http://www.ietf.org/Internet-drafts/draft-ietf-problem-issue-statement-02.txt.

Davies, E., J. Hoffman. 2004. IETF problem resolution process. http://rfc3844.x42.com/.

DiBona, C., S. Ockman, M. Stone. 1999. Introduction. C. DiBona, S. Ockman, M. Stone, eds. *Open Sources: Voices from the Open Source Revolution*. O'Reilly, Sebastapol, CA, 1–19.

Dougherty, D. 1992. Interpretive barriers to successful product innovation in large firms. *Organ. Sci.* **3**(2) 179–202.

Economist. 2003. Open source's local heroes. *Econom. Tech. Quart.* **369**(8353) 5.

Fleming, L., C. King, A. Juda. Small worlds and regional advantage. Under revision. *Organ. Sci.* Forthcoming.

Friedrich, R. 1982. In defense of multiplicative terms in multiple regression equations. *Amer. J. Political Sci.* **26**(4) 797–833.

Goetz, T. 2003. Open source everywhere. *Wired* **11**(11).

Gould, R. 1991. Multiple networks and mobilization in the Paris Commune, 1871. *Amer. Sociol. Rev.* **56**(6) 716–729.

Granovetter, M. 1992. Problems of explanation in economic sociology. N. Nohria, R. Eccles, eds. *Networks and Organization*. Harvard Business School Press, Boston, MA.

Greene, W. 1993. *Econometric Analysis*. Prentice Hall, Englewood Cliffs, NJ.

Hamm, S. 2005. Linux, Inc. *Business Week* **3918**(Jan. 31) 60.

Harris, S. 2001. The tao of IETF—A novice's guide to the Internet engineering task force. RFC 3160.

Hausman, J., W. Newey, J. Powell. 1995. Nonlinear errors in variables: Estimation of some Engel curves. *J. Econom.* **65** 205–233.

Heckman, James J. 1976. Sample selection bias as a specification error. *Econometrica* **47** 153–161.

Henderson, R., K. Clark. 1990. Architectural innovation: The reconfiguration of existing product technologies and failure of established firms. *Admin. Sci. Quart.* **35** 9–30.

Hoffman, P., S. Bradner. 2002. Defining the IETF. RFC 3233.

Kaufman, C. 2003. Personal communication. Westford, MA, (July 28).

Kogut, B., A. Metiu. 2001. Open-source software development and distributed innovation. *Oxford Rev. Econom. Policy* **17**(2) 248–264.

Lakhani, K. R. 2004. The determinants of developer mobilization in open source communities. Working paper, Massachusetts Institute of Technology, Cambridge, MA.

Lakhani, K. R., E. von Hippel. 2003. How open source software works: Free user-to-user assistance. *Res. Policies* **32**(6) 923–943.

Lee, G., R. Cole. 2003. From a firm-based to a community-based model of knowledge creation. *Organ. Sci.* **14**(6) 633–649.

Lerner, J., J. Tirole. 2001. Some simple economics of open source. *J. Indust. Econom.* **2** 197–234.

Lorsch, P. 1965. *Product Innovation and Organization*. McMillan, New York.

McPherson, J., P. Popielarz, S. Drobnic. 1992. Social networks and organizational dynamics. *Amer. Sociol. Rev.* **57** 153–170.

Merton, R. 1973. The normative structure of science. N. Storer, ed. *The Sociology of Science: Theoretical and Empirical Investigations*. University of Chicago Press, Chicago, IL.

Mockus, A., R. Fielding, J. Herbsleb. 2002. Two case studies of open source software development: Apache and Mozilla. *ACM Trans. Software Engrg. Methods* **11**(3) 309–346.

Mowery, D., T. Simcoe. 2002. Is the Internet a U.S. invention? An economic and technological history of computer networking. *Res. Policies* **31** 1369–1387.

Neustadt, R. E. 1990. *Presidential Power and the Modern Presidents: The Politics of Leadership from Roosevelt to Reagan*. Free Press, New York.

O'Mahony, S., F. Ferraro. Managing the boundary of an "open" project. J. Padgett, W. Powell, eds. *Market Emergence and Transformation*. MIT Press. Forthcoming.

Padgett, J., C. Ansell. 1993. Robust action and the rise of the Medici, 1400–1434. *Amer. J. Sociol.* **98** 1259–1319.

Perrone, V., A. Zaheer, B. McEvily. 2003. Free to be trusted? Organizational constraints on trust in boundary spanners. *Organ. Sci.* **14**(4) 422–439.

Podolny, J. M., J. N. Baron. 1997. Resources and relationships: Social networks and mobility in the workplace. *Amer. Sociol. Rev.* **62**(5) 673–693.

Raymond, E. 1998. *The Cathedral and the Bazaar*. O'Reilly and Associates, Inc., Sebastapol, CA.

Raymond, E. 2000. Conventions at light speed: What hackers can learn from SF fandom. http://www.catb.org/~esr/writings/sfshows/.

Reagans, R., B. McEvily. 2003. Network structure and knowledge transfer: The effects of cohesion and range. *Admin. Sci. Quart.* **48** 240–267.

Rivlin, G. 2003. Leader of the free world: How Linus Torvalds became benevolent dictator of Planet Linux, the biggest collaborative project in history. *Wired* **11**(11).

Rosenkopf, L., A. Metiu, V. George. 2001. From the bottom up? Technical committee activity and alliance formation. *Admin. Sci. Quart.* **46** 748–772.

Schoenfeld, D. 1982. Partial residuals for the proportional hazards regression model. *Biometrika* **69** 239–241.

Secure Shell secsh. 2003. http://www.ietf.org/html.charters/secsh-charter.html.

Sorenson, O., T. E. Stuart. 2001. Syndication networks and the spatial distribution of venture capital investments. *Amer. J. Sociol.* **106** 1546–1588.

Sorenson, O., J. Rivkin, L. Fleming. 2006. Complexity, networks, and knowledge flow. *Res. Policy* **35** 994–1017.

Tushman, M. 1977. Special boundary roles in the innovation process. *Admin. Sci. Quart.* **22** 587–605.

von Hippel, E., G. von Krogh. 2003. Open source software and the "private-collective" model: Issues for organization science. *Organ. Sci.* **14**(2) 209–223.

von Krogh, G., S. Spaeth, K. R. Lakhani. 2003. Community, joining, and specialization in open source software innovation: A case study. *Res. Policies* **32**(7) 1217–1241.

Wade, J. 1995. Dynamics of organizational communities and technological bandwagons: An empirical investigation of community evolution in the microprocessor market. *Strategic Management J.* **16** 111–133.

Wasserman, M. 2003. IETF problem resolution processes. http://www.ietf.org/Internet-drafts/draft-ietf-problem-process-01.txt.