Penguins, Camels, and Other Birds of a Feather: Brokerage, Boundary Spanning, and Leadership in Open Innovation Communities

Lee Fleming
Morgan Hall T95
Harvard Business School
Boston, Mass. 02163
lfleming@hbs.edu
(617) 495-6613

David M. Waguespack
University at Buffalo
The State University of New York
Department of Political Science
514 Park Hall, Buffalo, NY 14260
(716) 633-7594
dw23@buffalo.edu

April 8, 2005

We are very appreciative of the help of Scott Bradner and many other members of the IETF in completing this study. We would also like to thank Jóhanna Birnir, Ronald Burt, Michael Hannan, Eric von Hippel, Jerry Kim, Karim Lakhani, Siobhan O'Mahony, Renee' 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 at Berkeley.

Abstract: 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 voluntary communities in order to avoid the ever present danger of forking and balkanization. This is enabled by two correlated but distinct social positions: brokerage, and boundary spanning between technological modules. An inherent lack of trust associated with brokerage positions can be overcome through physical interaction or contributions within technological boundaries. Successful leaders are thus the product of strong technical contribution and a structural position that can bind the community together.

Open innovation communities typically lack financial or corporate backing, forgo personal ownership rights to their members' work, rely on unpaid 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 as well as 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, Fielding, and Herbsleb, 2002), and popular: 18 million people use open source operating systems (Rivlin, 2003). Even as open source operating systems challenge the world's most powerful software firm, 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 third world communities 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, non-proprietary, open innovation community.

Though 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, 2003). Despite their bazaar-like, egalitarian, argumentative, unplanned, chaotic appearance, open innovation communities rely heavily upon strong leadership to function effectively and resist splintering, forking, and balkanization (DiBona, Ockman, and Stone, 1999, Lerner and Tirole, 2001, Kogut and Metiu 2001, Lee and Cole 2003, von Hippel and von Krogh 2003, Lakhani 2004). Reputation counts; though 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 was illustrated by Burt (1992), who demonstrated that "brokers," individuals who connect otherwise disconnected actors, can exploit “structural holes” to advance more quickly in their careers. An individual who works with other individuals 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) illustrated a widespread correlation between ability, ties across multiple organizations, and leadership. They described 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.

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 characterized brokers as calculating and politically savvy operators, while Allen and Tushman characterized boundary spanners as well respected guardians who redirected 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 co-optation by commercial interests. The most effective strategies and behaviors of aspiring leaders will therefore be different for brokerage versus 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 be compelled to pay attention to any other member. Ascendancy in such relationships relies purely, to borrow a phrase from politics, upon “the power to persuade" (Neustadt, 1990). In addition to providing field settings in which to research a novel social phenomenon, open innovation communities often archive their interactions electronically. These public records afford social scientists an unprecedented opportunity to construct databases of human and social capital, social and political processes, and a host of outcomes, including the emergence of leadership.

To explain the emergence of leadership within open innovation communities, we induct theory from interviews, archival research, and 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 demonstrate the importance of the first technical contribution for future leaders. Marginal returns to additional contributions decrease, but nevertheless remain positive. Individuals who broker work collaborations are more likely to assume leadership, but the effect is strongly contingent upon physical interaction with 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 simultaneously span technological boundaries within the community. Boundary spanners, in contrast, do not suffer from a lack of trust. Successful leaders are thus the product of strong technical contribution coupled with physical interaction, contingent strategies, and a structural position that can bind the community together.

Open Innovation Communities and the Internet Engineering Task Force

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, Ockman, and Stone, 1999). Consistent with the 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 Internet Engineering Task Force (IETF).[1] 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 freely. While we induct theory from many open innovation communities, we test our hypotheses with a specific dataset 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 freely available (Bradner, 1999). The IETF is also

---

[1] The first part of our paper’s title refers to the animals frequently associated with these projects. The penguin, the logo for Linux, was chosen by founder Linus Torvalds for somewhat idiosyncratic reasons. Perl is associated with camels because the cover of a popular Perl manual features one. "Birds of a feather" is what the IETF terms nascent projects. Apache, though not explicitly referenced in the title, has as its logo an eagle feather. Space considerations preclude detailed histories of the various communities, for which we refer interested readers to books edited by DiBona, Ockman, and Stone; Abbate; Mowery and Simcoe; Feller, Fitzgerald, Hissam, & Lakhani 2005; see also von Hippel and von Krogh 2003.

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 maturing open technical communities Bradner (1999: 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.

In 1986 an amalgam of ad hoc technical committees advising the Defense Advanced Research Projects Administration (DARPA) on the original ARPANET Internet backbone coalesced into the IETF. This occurred just as the Internet was expanding to include the emerging National Science Foundation's NSFNET and as emerging commercial Internet service providers began to replace both networks. The IETF 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. Bradner (1999) asserts that: "all of the basic technology of the Internet was developed or has been refined in the IETF.” Although much of the communication and work of the task force is conducted via electronic media, members meet three times annually, 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), though 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 (figure 1). 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).

New working groups emerge mainly from grass roots interest in a topical area and are typically preceded by a “Birds of a Feather” (BOF) meeting convened during a conference. [2] Members solicit participation in BOFs through 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 (usually the BOF organizer). Appointment can thus be construed as confirmation of successful, if informal, open community leadership. Even though formal appointment is by an AD, the process recognizes only successful bottom up efforts.[3] Very few members advance to leadership; of 955 individuals who first attended an IETF meeting in 1995, 34 have become WG chairs.

---

[2] "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).

[3] This process, commonly observed in open innovation communities, was described by the founder of Linux 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).

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

Human capital explanations of success emphasize the capabilities of individuals (Coleman, 1988); people with greater skills, intelligence, and creativity are more likely to innovate or get promoted. Such explanations resonate strongly within engineering and open innovation communities. Many engineers reject social or “softer” explanations of success and tend to focus on purely functional contributions – they believe that the best solutions and deserved recognition should rely solely upon individual expertise, creativity, and hard work. Open innovation communities, like most engineering cultures, value technical contributions over all else, eschewing positional power and even democracy in favor of proven technology (Wasserman, 2003). Von Krough, Spaeth, and Lakhani (2003) document how applicants must contribute lines of code to be accepted into the community, those who merely post suggestions, critiques, and pleas generally being ignored.

Despite a strong aversion to non-technical sources of prestige and authority, community members readily recognize the status derived from technical contribution. Successful contributors, according to Raymond (1998), gain, “good reputation among one’s peers, attention and cooperation from others…and higher status in the…exchange economy.” Rivlin (2003) illustrates how Torvalds realizes 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 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). Bradner (2003b) maintains that "if someone is seen as strong and as having done good work in the past (and can argue their position well) they gain quite a bit of status and future proposals from them tend to get a better ride." 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' 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 will be sought out and respected and guide the future technical and social evolution of the community.

H1: Technical contribution will increase a member's likelihood of becoming an open innovation community leader.

While an initial technical contribution signals competence, additional contributions provide declining marginal returns. An engineer who aspires to leadership must demonstrate communication, motivational, and organizational skills. Technical challenge being a large part of what motivates prospective contributors, leaders must ensure that enough sufficiently interesting work is left for their followers. Lerner and Tirole (2001: 220) explain it thus.

> The initial leader must…assemble a critical mass of code to which the programming community can react. Enough work must be done to show that the project is doable and has merit. At the same time, to attract additional programmers, it may be important that the leader does not perform too much of the job on his own and leaves challenging

> programming problems to others.… While many make fewer programming contributions, having moved on to broader project management tasks, the individuals we talked to believed that the initial experience was important in establishing credibility to manage the project.

Bradner (2003a) concurs, maintaining that “people who are too involved in the [technical] documents” are less likely to become leaders. “So does the leader/coordinator for a bazaar-style effort really have to have exceptional design talent,” inquires Raymond (1998), “or can he get by on leveraging the design talent of others? I think it is not critical that the coordinator be able to originate designs of exceptional brilliance, but it is absolutely critical that the coordinator be able to *recognize good design ideas from others*.” These arguments imply that contributions past the initial one provide less and less benefit to aspiring leaders of open innovation communities.

H1a: The positive relationship between technical contribution and leadership will demonstrate a threshold effect for the first technical contribution.

H1b: Contributions beyond the first technical contribution will increase a member’s likelihood of becoming an open innovation community leader at a decreasing marginal rate.

Technical contribution remains the primary prerequisite for aspiring open innovation community leaders. Given the ability to contribute, however, an aspiring leader must still choose what and with whom (if anyone) to contribute. Engineers who collaborate more widely are more likely to lead, for a variety of reasons. The technical reputations of collaborative engineers will become more widely known, because their reputation will spread through the opinions of co-authors. Isolates’ reputations will only spread when their engineering is used or read by others. Isolates will also be less able to build consensus, motivate contribution, and affect convergence on difficult technical issues.

Controlling for their amount of technical contribution, engineers who have worked with a wider number of colleagues will be better able to mobilize successful initiatives because they can recruit a greater number of interested colleagues. Aspiring leaders need colleagues, friends, and “birds of a feather” who know and respect the leader’s abilities.

H2: A greater number of working relationships will increase a member’s likelihood of becoming an open innovation community leader.

While working with many co-authors should increase the likelihood of leadership, the most important role – and challenge - for an open innovation community leader is to integrate and bind the community together. Because such communities remain voluntary, members can always leave and mobilize a new effort, a process known as “forking.” (DiBona, Ockman, and Stone 1999, Kogut and Metiu 2001, Lerner and Tirole 2001) Of Torvalds Rivlin (2003) 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 through integrating and intermediary roles such as brokerage and boundary spanning (Perrone, Zaheer, and McEvily, 2003). Leaders who work within cohesive networks or technological boundaries are less likely to be aware of, and less able to influence, balkanization processes. Occupants of brokerage and boundary spanning positions therefore have great influence over the evolution of open innovation communities.

Conflicting arguments have been made, however, for the influence of brokerage upon leadership. Burt (1992, 1997, 2000, 2001a), who defines a broker as the only social connection or bridge among a group of actors, argues that brokering enhances both career mobility and promotion. In the current empirical context, brokering occurs when co-authors do not collaborate with one another 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, 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 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, to wit: "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: 18; Bradner, 2004). For these reasons open innovation communities remain within the boundary conditions of Burt's arguments.

Illustrating Burt's arguments in the context of open innovation communities, brokerage provides definite advantages to aspiring leaders. Because brokers can draw upon different social networks, they bring greater and less redundant resources to bear upon a problem. They will know the skill sets and technical experiences of individuals in a wider diversity of social arenas and be able to recruit them for appropriate future projects. Furthermore, because of their position as the sole intermediary of an information bottleneck, brokers are themselves more likely to come up with innovative ideas (Hargadon and Sutton 1997; Burt 2003). This will enhance their technical reputation and indirectly, their likelihood of becoming a leader.

But brokerage can also occasion disadvantages for aspiring leaders. Researchers have long argued that cohesive social structures build trust by exerting pressure for consistent norms among individuals with 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) 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!" Community members understand and fear the potential for abuse and proclaim strong norms against manipulative behavior. Raymond (1998) 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. 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 Krough, Spaeth, and Lakhani, 2003). As Coleman (1988) has argued and Gould (1991) described with reference to the Parisian revolts of 1871, mobilization will be easier if leaders can call upon 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 leadership, 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 subsequent leadership. Burt (2001b) recognized 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) 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.” This is precisely how Torvalds has retained leadership in Linux according to one contributor: “Linus has power, but he doesn’t have it by fiat.… He has power because people trust him” (Hamm, 2005). 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, 2003).[4] Davies (2003: 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

---

[4] 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, Ockman, and Stone, 1999).

> can be applied beyond not asking them to do a similar task again. The relationship works best when it is two-way—the person being asked to perform a task needs to be able to rely on the behavior of the person doing the asking. 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.

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 based on 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, 2003). 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). The IETF holds three physical conventions a year at various locations around the globe:

> The meetings, held three times a year, are week-long dweebfests whose primary goal is to reinvigorate the Working Groups to get their tasks done, and whose secondary goal is to promote a fair amount of mixing between the Working Groups and the areas…there is lots of work, as well as a fair amount of time for socializing (Harris 2001 pg. 10).

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 achieve leadership.

H3: The interaction of physical attendance and greater 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 was emphasized in early research on the structure of engineering organizations (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 not-invented-here syndrome (Allen, 1977). Technologies can be refined more productively within technologically focused efforts, but at the risk of becoming incremental and obsolete (Fleming, 2001). Boundary spanners reduce this risk of obsolescence by gathering, interpreting, and disseminating non-redundant information across boundaries (Allen, 1977; Tushman, 1977). They also integrate efforts across mutually interdependent organizations (Lorsch, 1965).

Because the foregoing research predates the rise of open innovation communities the internal boundaries of such communities remain unclear (O' Mahony and Ferraro, 2003). These boundaries do exist, however, and as in private firms, usually correspond to the interfaces between technological sub-systems (Henderson and Clark, 1990). Each boundary demarcates a distinct technological module. Internal community boundaries are defined directly by leaders' architectural considerations and indirectly 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, 2004) describes how individuals who aspire to leadership in the IETF must define their technological boundaries relative to other community efforts as part of their Birds of a Feather 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?" Once these boundaries have been suggested, contributors choose where to work and in doing so, create more or less intense areas of social interaction. Technological design boundaries thus become social boundaries as engineers focus on the immediate technical challenge. Nothing, however, prevents a boundary spanner from being the only engineer to collaborate across modules and thereby be a social broker, or from being one of many engineers to collaborate across modules and not be a social broker.

Open innovation communities are well aware of these issues and value individuals that can span technical boundaries (Davies 2003, pg. 13; Bradner 2004), for many of the same reasons that they value social brokers. That boundary spanners should be more creative and able to call upon more diverse resources is explicitly recognized in the expectation that leaders promote "cross-pollinating between groups" (Davies, 2003: 13)." In addition to providing the 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 both attract dedicated coalitions that refuse to compromise or resolve incompatibilities. Hamm (2005) describes how "Linux backers fret that different versions of the software will become incompatible with one another" and Davies and Hoffman (2004: 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 within the 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; 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.

H4: Internal community boundary spanning will increase a member's likelihood of becoming an open innovation community leader.

Despite their probable 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 cohesive boundary. These distinctions provide an additional opportunity to test our theory in parallel with hypothesis three. Our development of hypothesis three 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 the technologies, objectives, and reputations of collaborators in the non-local areas. In such a situation concerns about trust will be magnified if an individual invests serious resources in another group's efforts, to wit, Podolny and Baron (1997): "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 must span technological boundaries within the community. 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.

H5: The interaction of brokerage and internal community boundary spanning will decrease a member's likelihood of becoming an open innovation community leader.

Evidence

We draw data primarily from the published proceedings of the IETF's tri-annual meetings from its inception in 1986 to 2002, and from the organization's RFC publication series over the same time period.[5,6] Each proceedings represents a “snapshot” of the IETF at a given time, complete with a list of registered conference attendees (including their email 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 2003 and 2004, Kaufman 2003) at their places of employment and at the 56th IETF meeting in San Francisco, March 16-21, 2003.

We developed from these sources a longitudinal dataset of observations on 15,465 individual IETF participants at four month intervals. This dataset spans 1989 to 2002, the first three years’ of conference data having been truncated due to leading independent variable creation. Subjects enter the risk set upon first attending an IETF meeting or in the first trimester of 1989 if community involvement occurred before 1989. Subjects exit

---

[5] In 1987, 1989, and 1990 the IETF held four meetings.

[6] The acronym RFC means "Request for Comments." In current parlance an RFC refers to a published document, not one under review. RFCs either detail a technical standard (standards-track RFC) or provide other relevant information (non-standards-track RFC). The Internet Engineering Steering Group (IESG), comprising the IETF chair and area directors, has final authority to approve or reject proposed publications.

on the first working group chair appointment or are treated as right censored. 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. Future research, however, aims to clarify the relationships between position, trust, prior leadership success, and repeated or higher level leadership (for example, appointment as an Area Director).

A subset of data consists of those individuals who publish *two or more* RFC documents. For these 609 individuals we can compute meaningful social network measures based on collaboration patterns. The analysis requires two or more publications because individuals with a single publication are by definition in a cohesive and non-brokering relationship with all collaborators on that one publication.[7] We discuss the problems and remedies associated with this kind of sample selection below. We investigated other potential measures of relationships, but found each wanting: email interactions and working group attendance lists were inconsistently archived, particularly for earlier interactions; email 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

---

Standards-track RFCs typically must also be approved by an established relevant IETF working group. Non-standards-track RFCs are not subject to working group scrutiny prior to publication.

[7] Although common employer is another possible way to infer social relations in this empirical context, firms often send dozens of employees from different divisions to IETF meetings. Interview subjects frequently asserted that they would only know a handful of their affiliated colleagues and that this did not influence or represent their other social relations.

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

Variables

Table 1 presents descriptive statistics for independent variables in the full risk set of 15,465 subjects observed during 243,370 trimesters, Table 2 bi-variate correlations for the same data. Tables 3 and 4 present statistics and correlations for networked subjects, 609 subjects observed over 5,032 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. There are 473 events in the full risk set and 117 events in the network risk set.

Community presence (instrumented) measures the extent of an individual's physical involvement with the IETF community over the prior three years. Conference attendance data and participants’ phone numbers enabled us to develop an instrument to control for endogeneity (for example, we cannot directly observe an individual’s aspirations for leadership, or her employer’s wish to co-opt the community by influencing the leadership processes). We begin by counting the number of IETF conferences an individual attends over the prior three years. Aspirations probably correlate with the probability of attaining a leadership position, and given that the IETF remains a volunteer organization, with the effort expended to attend conferences. That aspirations remain unobserved raises the concern that a variable that simply counts attendances at conferences is endogenously determined. In ordinary least squares estimation endogeneity violates the assumption that an independent variable is uncorrelated with the error term (Greene, 1993: 285). One solution to this problem is to find a proxy that is correlated with the offending independent variable but not the disturbance term. Put another way, the proxy needs to correlate with attendance but not with the probability of becoming a leader.

To develop such a proxy we employ the instrumental variable approach (Greene, 1993: 603). We first regressed three years conference attendance on a series of variables deemed likely to influence the probability of attending a conference but unlikely to influence selection to a leadership position. Generally speaking, the variables used to compute *Community presence (instrumented)* relate to the costs of, and propensity for, participation in the IETF. *Miles* accounts for total distance, in miles, between an engineer's home and all IETF meetings during the prior three years. [8] To create this measure 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 (for a similar calculation see Sorenson and Stuart, 2001). *Miles* enters the first stage equation as a polynomial term. *Non-US* is a categorical indicator that the subject lives outside the United States. *University affiliation* and *Government affiliation* indicate that a subject works for an organization of the respective type and not the reference category of a firm. The assumption is that non-commercial organizations might have limited resources to support member attendance. *Affiliated conference attendees* is the number of conference appearances by a subject's colleagues; we include this as a measure of the resources a given employer devotes to IETF participation.

---

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

After regressing observed three year conference attendance on *Travel miles, Non-US, University, Government,* and *Affiliated conference attendees*, we recovered predicted values for *Community presence (instrumented)*. In OLS 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 upon attendance was 16.3%. Given the high amount of explained variance, the instrument should not be vulnerable to finite sample and weak instrument bias (Bound, Jaeger, and Baker 1995). Stock, Wright, and Yogo (2002) also define a strong instrument as having less than 10% bias, relative to a model without an instrument. The 5% F critical value for this criterion and degrees of freedom is 11.51, well below the first stage F statistic of 90.63.

(ln) IETF Publications, which counts total IETF technical publications authored or co-authored in the prior three years, is our main measure of individual human capital. These documents also reveal working group affiliation. *Two WGs* indicates that an engineer spanned two technical sub-domains; *Three+ WGs* indicates that a subject spanned three or more. Use of a count variable for the number of working group memberships instead of breakout into mutually exclusive categories returned similar estimates. We retain the exclusive categories to enable qualitative differentiation between working group membership (a control variable) and boundary spanning.

We also use publication data to construct longitudinal social networks. We created social network measures based on co-authorship 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, Everett, and Freeman, 2001). *Degree* measures, on the basis of all drafts published in the previous three years, the number of unique individuals with whom an engineer has published a draft. *Social Brokerage* measures the opposite of individual constraint, where constraint measures the extent to which ego has ties with many alters and those alters have ties between them (Burt 1992). We enter the negative of constraint (Eq. 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.

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

We calculate each of the social network measures only for cases in which a subject has two or more publications. The logic behind this restriction is that variation on brokerage or boundary spanning is possible only when an actor has two or more opportunities (i.e., two or more publications) on which to interact with other individuals. Restricting analysis to such a subset produces a non-random 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, 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, Morey, and

Shannon, 2004) indicates that sample selection can also bias hazard model estimates. To address this problem, following Heckman we estimate the probability that any of our at-risk engineers publishes two or more IETF documents on a number of variables likely to influence network entry. These model estimates include *(ln) Community Size* at t, which gauges changes in the baseline level of authorship relative to the size of the internet community. *U.S. Patents* obtained by a subject during the prior three years captures individuals with demonstrated technical expertise and a propensity to publish. Patent counts were arrived at by matching IETF engineers to the population of U.S. patent inventors (as developed by Fleming, King, and Juda, 2004). [9] *Affiliated RFC Authors* and *Affiliated WG chairs* indicate employer resources sufficient to support a high level of IETF participation. We include categorical indicators for *Non-US* and *University* and *Government* affiliation. The selection instrument, the *Inverse Mills Ratio*, indicates the probability of non-selection based on a first stage equation that estimates the probability of authoring two or more IETF publications. A positive coefficient indicates that selected individuals are more likely than the population at large to become leaders.

Finally, the models include a series of control variables. *Summer Meeting, Winter Meeting*, and *Non-North American Meeting* are dummy variables that denote the season

---

[9] 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, King, and Juda (2004). 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). IETF members demonstrated no significant difference in their brokerage of other inventors, boundary spanning, future prior art citations to their patents (a widely used indicator of quality), references to the non-patent (generally scientific) literature, and proclivity to repeatedly collaborate with the same colleagues. Hence, IETF members who patent appear to publish more, and work with more people and organizations, than typical inventors, though we hesitate to claim strong inferences due to the differences in context, measures, and comparison groups.

and location of the current IETF conference. We include these variables to control for community processes possibly being related to the timing of working group chair appointments. Appointments to the IESG typically being made at the spring meeting, it is conceivable, for example, that "lame-duck" area directors approve fewer groups during that trimester. *IETF Executive* is an indicator that a subject is a past or present member of one the IETF's executive committees. It is possible, though rare, for individuals to jump straight to a higher level position and still be at risk of their first working group appointment. *Affiliated Executive* indicates that a subject's co-worker currently holds a position on one of the high level IETF oversight committees. We do not assume that these individuals know one another inasmuch as some firms send hundreds of engineers to the IETF. Rather, we construe that high level leadership to be a proxy for firms that commit significant resources to the community. *(ln) Community size* is the natural log of total meeting registrants over the prior three years meant to capture changes in the supply of engineers and size of the community. *Incumbent WG Chairs* is a count of incumbents (those not at risk of a first leadership position) holding chair positions and intended to gauge the extent to which the community currently fulfills its leadership needs. *Executive Tie* measures the number of collaborative working relationships with a current Area Director within the last three years to control for reliance on personal ties with individuals who have control over the appointment process. Finally, the indicator variable *WG Contributor* records when a subject has 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 are not competing 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.

$$r(t) = \lim_{t' \to t} \frac{\Pr(t \le T < t' | T \ge t)}{t' \to t} \qquad (2)$$

We estimated semi-parametric 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).

Piece-wise exponential models returned very similar substantive results. We estimated all models with STATA version 8.

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

Results

Table 5 reports hazard models for appointment to Working Group leadership for the entire dataset, table 6 for the networked risk set. Models 1 to 5 explore the functional relationship between drafts and appointment for the full risk set. Model 1 provides a baseline with controls only. Model 2 introduces a categorical indicator of any publication, model 3 the ln of drafts, and model 4 both the indicator and ln of drafts. Model 5 adds a second order term for the number of drafts (we introduce the square of the original variable to avoid collinearity between the ln and second order terms). Model 5 does not demonstrate significant improvement over previous models (according to a chi squared test of the difference of maximum likelihoods with one degree of freedom). Model 4 is nested and much better than the earlier models, hence, publication provides an initial benefit for any publication, followed by declining marginal returns for additional contributions. The lack of significance of the second order term indicates that the influence of additional drafts never turns negative, although the ln indicates decreasing importance. Figure 2, based on an unreported model that uses indicator categories for total publications (without any other measures of publication), demonstrates the dramatic nature of the publication effect and is consistent with a threshold and logged specification. The effect sizes are quite large in the models and figure. According to model 4, the first publication increases the likelihood of becoming a leader by 143%.[10] Figure 2 indicates that for three publications the point estimate increases 1,590% relative to having none. These results support the arguments of hypothesis one: a single publication greatly increases an individual's probability of leadership, and that additional publications increase that probability at a declining marginal rate.

Models 6 through 14 in table 6 present results for the networked risk set. Model 6 presents a baseline estimation and model 7 steps in the publication variable. We maintain the publication variable and controls in all subsequent 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 8 through 12 enter additional explanatory variables individually and then with interactions. Model 13 tests a full specification and model 14 includes the networking risk set selection term.

Despite the importance of structural position to leadership, individuals still gain advantage through additional publications. A one standard deviation increase in publications increases the probability of leadership by 22%; having authored two drafts, authoring a third increases the probability by 10%. Surprisingly, degree never demonstrates a significant effect in any model; working with additional co-authors does not correlate with leadership. We attribute this to a dilution effect of the credit that occurs when an individual co-authors with too many others. Discussion with IETF members lent

---

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

validity to this interpretation. Hence, the results support hypothesis one but not hypothesis two.

The models 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 three, 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 14 (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 49%. 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 without a strong community presence. Hence, the models and figure three support hypothesis three.

Although social brokerage and its interaction with attendance demonstrate strong positive effects, the benefits of boundary spanning are potentially greater. Contributing to two working groups correlates with a 63% increase in leadership (over and above the 98%

benefit of contributing within any working group). Contributing to three or more working groups correlates with a 287% increase. Simultaneous brokerage and membership in two groups correlates with a negative interaction effect of 45%, in three or more groups, 64%. 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 three or more. Hence, the models support hypothesis 4, that boundary spanning increases the likelihood of leadership, and hypothesis 5, that simultaneous brokerage decreases this positive effect. Further supporting our theory, unreported models did not demonstrate even marginally significant interaction between boundary spanning and the instrumented attendance variable. Although prediction and observation of a null result do not test theory (hence, are not included in our hypotheses), the lack of significance is nevertheless consistent with our contention that fellow community members perceive the two roles differently.

To directly compare the different structural strategies and their influence on leadership we assume two individuals; first, a broker who scores one standard deviation above the brokerage mean and contributes to only one working group, and second, a boundary spanner who scores at the brokerage mean and contributes to two groups. Both strategies work similarly well if they attend an average number of conferences. The broker is then 70% more likely to become a leader compared with a 63% increase for the boundary spanner. If both individuals increase their attendance at conferences by one standard deviation, the advantage shifts to the broker, since she gains 49% from the interaction of brokerage and attendance. The advantage shifts back to the boundary spanner, however, if she spans additional boundaries; assuming all else stays equal, she is 224% more likely to become a leader. [11] While it appears moderately effective to bind a single working group with itself, it appears even more effective to bind three groups within the larger community. Moderately ambitious individuals should therefore look for local integration opportunities; very ambitious individuals should seek out the cohesive nexuses of community wide 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, considering first the full risk set, increasing the size of the community (*(ln) Community Size*) increases the potential leadership opportunities, and crowding in the leadership ranks (*Incumbent WG Chair Population*) diminishes those opportunities. New appointments occur less often during the summer meetings (*Summer Meeting*). Shared affiliation with a higher level leader (*Affiliated Executive*) increases the probability of leadership, a result we interpret to indicate that the engineer's firm strongly supports participation in the community (rather than cronyism). Supporting this interpretation, in the networked risk set, collaboration with an Area Director 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. Contributing within a working group has a substantial positive effect of 98%. Finally, the selection term, though positive, does not reach conventional levels of significance.

---

[11] 287% advantage for three or more groups minus the 63% advantage for two groups.

Discussion

While the results support most of the predictions, it is important to demarcate their limitations, both theoretical and empirical. We generalize the theoretical implications 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 for-profit firms promote to middle 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 sanction (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. Overall, however, the results illustrate the power of brokerage and boundary spanning concepts in a previously unexamined social context.

Caution should also be exercised before generalizing the results to all open innovation communities, particularly those that lack physical interaction. Our results are conditional upon attendance, though 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. Caution should also be exercised due 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, in which case the possibility of forking or "disruptive selection" (McPherson, Popielarz, and Drobnic 1992) should greatly increase.

While our network analysis provided a novel methodology for the analysis of publication based relationships, the research still suffers from a variety of empirical limitations. While the Heckman selection term reached significance in partial models (not shown), the lack of significance in the final model indicates weakness in the predictors of multiple publications. Nonetheless, the lack of substantive change adds to our confidence. Overall, the approach illustrates a method that can be replicated whenever network variables remain unobserved or undefined for a subset of subjects, for example, for an isolated subject.

In addition to the limitations of the selection instrument, the dependent variable of appointment measures leadership after a Birds of a Feather meeting, and not the organization of the meeting itself. Ideally, we would observe which community members attempted to organize a meeting, and then conditional on that successful organization, which members advance to leadership. Archived Birds of a Feather meeting documentation remains inconsistent, however, and more complete for meetings that become working groups. Rather than introduce unquantifiable 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, technical contribution becomes insignificant, roles and their interactions become less significant, and mere attendance appears as the most 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 miss the importance of technical contribution and 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 (an R^2 of 16.3%), given that the second stage estimates a non-linear Cox model, and the use of instruments with non-linear models remains controversial (see Bowden and Turkington 1981, Box-Steffensmeier and Jones 2004, vs. Hausman, Newey, and Powell 1995).

Despite the limitations, the results contribute to the larger network literature. 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 more strongly associated with brokerage. This finding should be elaborated in different contexts with dynamic analysis, 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. Given that the quantitative research on the benefits of brokerage upon innovation does not distinguish brokerage from boundary spanning, it also remains unclear which position fosters creativity more effectively. These positions must be studied dynamically, since successful innovation probably leads to a rapid accretion of ties and increased cohesion across social or technological boundaries, once a breakthrough has occurred.

Our results can also inform research into the careers and management of scientific and technical professionals. First, as figure 2 dramatically illustrates, the work demonstrates the importance of human capital. To advance in these scientific and technical contexts an individual cannot rely upon political power, affiliation, or relationships with powerful individuals (indeed, prior collaboration with a powerful individual demonstrates a significantly negative influence upon leadership). Given the demonstration of technical ability, however, an individual's social capital gains immense value. Social capital remains contingent, however, and should be cultivated selectively. For example, an oft heard concern of junior academics is the high risk of gaining tenure as a boundary spanner. An economic sociologist, for example, is probably better advised to pursue tenure in one or the other field rather than try to demonstrate legitimacy and contribution in both fields simultaneously. Similar interaction effects —e.g., physical participation within departments and attendance at conferences or, more important, contribution at a cohesive boundary—might obtain in scientific communities. Such cohesive boundaries likely give birth to new communities and active research efforts and provide fertile ground for high visibility contributions as happened, for example, with biology, chemistry, crystallography, and the discovery of DNA. Finally, non-academic engineers and scientists are increasingly mobile whether as long-term or temporary professionals (Bradach, 1997; Rosenkopf, Metiu, and George 2001; Gura, 2004). Such mobility requires extra-organizational reputation, contacts, and access to opportunities. Open innovation communities enable professionals to establish a reputation outside of a single organization, to cultivate contacts in other firms, and to learn about employment opportunities. Economists, who were at first perplexed as to why an individual would contribute to a volunteer effort without pecuniary compensation, have since argued that skill enhancement and quality signaling constitute the primary motivations. This implies that community contributors will earn higher salaries and be more likely to change employers. Our work, in emphasizing the importance of social networks, provides the additional explanation that individuals contribute to develop contacts that enhance career mobility. Future research hopes to differentiate between the economic and social influences on open innovation community contributors' employment mobility.

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 sixteen years of meetings and communications to trace the emergence of leaders within the organization. The models demonstrated a threshold effect for the first technical contribution, with decreasing marginal returns for additional contributions. 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 co-optation by commercial interests or forking over technical disagreements. Because brokers by definition contrive less cohesive and trusting contexts, the probability that they will assume leadership roles is 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 that the positive effect of brokerage demonstrates contingency 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. 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 pursue knowledge and advance technology (von Hippel and 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 delineate and describe the factors that produce those leaders.

References

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

Ahuja, G.
2000 “Collaboration networks, structural holes, and innovation: A longitudinal study.” Administrative Science Quarterly, 45: 426-455.

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

Boehmke, Frederick J., Daniel S. Morey, and Megan Shannon
2004 "Selection bias and continuous-time duration models: Consequences and a proposed solution." Working paper, University of Iowa.

Bound, J., D. Jaeger, and R. Baker
1995 “Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak.” Journal of the American Statistical Association, 90, 430: 443-450.

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

Bowden, R., and D. Turkington
1981 “A comparative study of instrumental variables estimators for nonlinear simultaneous models.” Journal of the American Statistical Association, 76, 376: 988-995.

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

Bradach, J.
1997 “Flexibility: The new social contract between individuals and firms?” Working paper 97-088, Harvard Business School.

Bradner, S.
1996 The Internet Standards Process—Revision 3. RFC 2026.

Bradner, S.
1998 IETF Working Group Guidelines and Procedures. RFC 2418.

Bradner, S.
1999 “The Internet Engineering Task Force.” In C. DiBona, S. Ockman, and M. Stone, (eds.), Open Sources: Voices from the Open Source Revolution:47-52. Sebastapol, CA: O’Reilly.

Bradner, S.
2003a "IETF structure and Internet standards process." Presentation to the 57th IETF Convention, Vienna, Austria. 2003b Personal communications, January 29, June 26, Cambridge, MA.

Bradner, S.
2003b personal communication, Cambridge, Ma. March 24.

Bradner, S.
2004 personal communication, Cambridge, Ma. August 28.

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

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

Burt, R.
1997 "The contingent value of social capital." Administration Science Quarterly, 42: 339-365.

Burt, R. S.
2000 "The network structure of social capital." Research in Organizational Behavior, 22: 345-423.

Burt, R.
2001a "Structural holes versus network closure as social capital." In N. Lin, K. Cook, and R. Burt (eds.), Social Capital: Theory and Research. Aldine de Gruyter. 2001b "Bandwidth and echo: Trust, information, and gossip in social networks." In A. Casella and J. Rauch (eds.), Networks and Markets: Contributions from Economics and Sociology. Russell Sage Foundation.

Burt, R.
2003 "Social origins of good ideas." Working paper, University of Chicago.

Coleman, J. S.
1988 "Social capital in the creation of human capital." American Journal of Sociology, 94: S95-S120.

Cox, D.
1972 "Regression models and life-tables (with discussion)." Journal of Royal Statistical Society, Series B 34:187-220.

Coya, S.

"IETF structures and Internet Standards Process," http://ietf.org/newcomer/frame.htm.

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

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

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

DiBona, C., S. Ockman, and M. Stone
1999 "Introduction." In C. DiBona, S. Ockman, and M. Stone (eds.), Open Sources: Voices from the Open Source Revolution:1-19. Sebastapol, CA: O'Reilly.

Dougherty, D.
1992 "**Interpretive barriers to successful product innovation in large firms**." Organization Science, 3, 2: 179-202.

Economist
2003 "Open source's local heros." The Economist Technology Quarterly, (Dec. 6).

Feller, J., Fitzgerald, B., Hissam, S., & Lakhani, K. R. (Eds.).
2005. Perspectives on Free and Open Source Software. Cambridge: MIT Press.

Fleming, L.
2001 "Recombinant Uncertainty in Technological Search." Management Science, 47, 1: 117-132.

Fleming, L., C. King, and A. Juda
2004 "Small worlds and regional advantage" Working paper, Harvard Business School.

Friedrich, R.
1982 "In defense of multiplicative terms in multiple regression equations." American Journal of Political Science, 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."
American Sociological Review, 56, 6: 716-729.

Granovetter, M.

1992 “Problems of explanation in economic sociology.” In N. Nohria and R. Eccles (eds.), Networks and Organization.Boston, Ma: Harvard Business School Press.

Greene, W.
1993 Econometric Analysis, Englewood Cliffs, N.J.: Prentice Hall.

Gura, T.
2004 “Joining a trend, scientists increasingly say ‘call my agent’.” Science, 303 (16 January): 303-304.

Hamm, S.
2005 “Linux, Inc.” Business Week, Issue 3918 (Jan. 31): 60.

Hansen, M.
1999 “The search-transfer problem: The role of weak ties in sharing knowledge across organization subunits.” Administrative Science Quarterly, 44, 1: 82-111.

Hargadon, A., and R. Sutton
1997 “Technology brokering and innovation in a product design firm.” Administration Science Quarterly, 42: 716-749.

Harris, S.
2001 The Tao of IETF—A Novice's Guide to the Internet Engineering Task Force. RFC 3160.

Hausman, J., W. Newey, and J. Powell
1995 “Nonlinear errors in variables: Estimation of some Engel curves,” Journal of Econometrics, 65: 205-233.

Heckman, James J.
1976 "Sample selection bias as a specification error." Econometrica, 47: 153-161.

Henderson, R. and K. Clark
1990 “Architectural Innovation: The Reconfiguration of Existing Product Technologies and Failure of Established Firms,” *Administration Science Quarterly* 35: 9-30.

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

Kaufman, C.
2003 personal communication, Westford, Ma. July 28, 2003,

Kogut, B., and A. Metiu
2001 “Open-source software development and distributed innovation.” Oxford Review of Economic Policy, 17, 2: 248-264.

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

Lakhani, K. R., and E. von Hippel
2003 "How open source software works: Free user to user assistance." Research Policy 32, 6: 923-943.

Lee, G., and R. Cole
2003 “From a firm-based to a community-based model of knowledge creation.” Organization Science, 14, 6: 633-649.

Lerner, J., and J. Tirole
2001 “Some simple economics of open source.” The Journal of Industrial Economics, L, 2: 197-234.

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

McPherson, J. and P. Popielarz, and S. Drobnic
1992 “Social Networks and Organizational Dynamics.” American Sociological Review 57, 153-170.

Merton, R.
1973 “The normative structure of science.” In N. Storer (ed.), The Sociology of Science: Theoretical and Empirical Investigations. Chicago: University of Chicago Press.

Mockus, A., R Fielding, and J. Herbsleb
2002 “Two case studies of open source software development: Apache and Mozilla.” ACM Transactions on Software Engineering and Methodology, 11, 3: 309-346.

Mowery, D., and T. Simcoe
2002 “Is the Internet a U.S. invention? An economic and technological history of computer networking.” Research Policy, 31: 1369-1387.

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

O'Mahony, S., and F. Ferraro
2003 “Managing the boundary of an ‘open’ project.” Working paper, Harvard Business School.

Padgett, J., and C. Ansell

1993 “Robust action and the rise of the Medici, 1400-1434.” American Journal of Sociology, 98: 1259-1319.

Perrone, V., A. Zaheer, and B. McEvily
2003 “ Free to be trusted? Organizational constraints on trust in boundary spanners.” Organization Science, 14, 4: 422-439.

Podolny, J. M., and J. N. Baron
1997 "Resources and relationships: Social networks and mobility in the workplace." American Sociological Review, 62, 5: 673-693.

Podolny, J. M., and T. E. Stuart
1995 “A role-based ecology of technological change.” American Journal of Sociology, 100, 5: 1224-1260.

Raymond, E.
1998 “The cathedral and the bazaar,” version 1998/11/22. Available at www.tuxedo.org.

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

Reagans, R., and B. McEvily
2003 “Network structure and knowledge transfer: The effects of cohesion and range.” Administrative Science Quarterly, 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, and V. George
2001 "From the bottom up? Technical committee activity and alliance formation." Administrative Science Quarterly, 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., and T. E. Stuart
2001 “Syndication networks and the spatial distribution of venture capital investments." American Journal of Sociology, 106: 1546-1588.

Stock, J., J. Wright, and M. Yogo

2002 “A survey of weak instruments and weak identification in generalized method of moments.” Journal of Business and Economic Statistics, 20, 4: 518-529.

Stuart, T., H. Hoang, and R. Hybels
1999 “Interorganizational endorsements and the performance of entrepreneurial ventures.” Administration Science Quarterly, 44: 315-349.

Tushman, M.
1977 “Special Boundary Roles in the Innovation Process.” Administration Science Quarterly, 22: 587-605.

Von Hippel, E., and G. von Krogh
2003 “Open source software and the ‘private-collective’ model: Issues for organization science.” Organization Science, 14, 2: 209-223.

Von Krogh, G., Spaeth, S., & Lakhani, K. R.
2003. Community, Joining, and Specialization in Open Source Software Innovation: A Case Study. Research Policy, 32(7): 1217-1241.

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

Table 1: Descriptive statistics for full risk set (N=243405).

	Mean	Std. Dev.	Min	Max
Summer Meeting	0.33	0.47	0.00	1.00
Winter Meeting	0.35	0.48	0.00	1.00
Non-North American Meeting	0.25	0.43	0.00	1.00
IETF Executive	0.00	0.04	0.00	1.00
Affiliated Executive	0.15	0.36	0.00	1.00
Patents	0.13	0.82	0.00	33.00
(ln) Community Size	8.49	0.62	5.74	8.99
Incumbent WG Chairs	162.71	47.44	16.00	230.00
Community Presence (Inst.)[a]	0.00	0.40	-1.49	2.92
Miles	16569.67	9776.30	0.00	60072.18
Non US	0.24	0.43	0.00	1.00
University affiliation	0.12	0.33	0.00	1.00
Government affiliation	0.05	0.22	0.00	1.00
Affiliated Conf. Attendees	61.23	142.07	0.00	1286.00
IETF Publication Dummy	0.061	0.239	0.00	1.00
(ln) IETF Publications	0.059	0.237	0.00	3.04

[a] mean centered to facilitate interpretation of interaction effects (Friedrich 1982)

Table 2: Bi-variate correlation matrix for full risk set (N=243405).

		1	2	3	4	5	6	7	8
1	Summer Meeting	1.00							
2	Winter Meeting	-0.52	1.00						
3	Non-N. American Meeting	0.60	-0.42	1.00					
4	(ln) Community Size	0.00	0.06	0.18	1.00				
5	Incumbent WG Chairs	-0.01	0.08	0.17	0.94	1.00			
6	Affiliated Executive	-0.01	-0.01	-0.02	-0.06	-0.06	1.00		
7	IETF Executive	0.00	0.00	-0.01	-0.07	-0.06	0.11	1.00	
8	Patents	0.00	0.00	0.00	0.01	0.00	0.07	0.02	1.00
9	Community Presence (Inst.)	0.01	-0.01	0.06	0.16	0.15	0.24	0.01	0.04
10	IETF Publication Dummy	0.00	-0.01	-0.01	-0.06	-0.06	0.09	0.11	0.06
11	(ln) IETF Publications	0.00	-0.01	-0.01	-0.06	-0.05	0.09	0.15	0.06

		9	10	11
9	Community Presence (Inst.)	1.00		
10	IETF Publication Dummy	0.07	1.00	
11	(ln) IETF Publications	0.07	0.92	1.00

Table 3: Descriptive statistics for network risk set (N=5066).

Variable	Mean	Std. Dev.	Min	Max
Summer Meeting	0.33	0.47	0.00	1.00
Winter Meeting	0.34	0.47	0.00	1.00
Non-North American Meeting	0.23	0.42	0.00	1.00
IETF Executive	0.04	0.20	0.00	1.00
Affiliated Executive	0.31	0.46	0.00	1.00
(ln) Community Size	8.35	0.75	5.74	8.99
Incumbent WG Chairs	153.20	51.51	16.00	230.00
Patents	0.37	1.38	0.00	17.00
Executive Tie	0.57	0.77	0.00	5.00
Community Presence (Inst.)[a]	0.00	1.17	-4.84	4.04
Miles	18872.36	7829.88	0.00	59767.03
Non US	0.15	0.36	0.00	1.00
University affiliation	0.13	0.33	0.00	1.00
Government affiliation	0.04	0.19	0.00	1.00
Affiliated Conf. Attendees	126.53	221.19	0.00	1286.00
(ln) IETF Publications	1.35	0.34	1.10	3.04
Degree	7.01	6.19	0.00	39.00
WG Contributor	0.82	0.38	0.00	1.00
Two WGs[b]	0.23	0.42	0.00	1.00
Three+ WGs[b]	0.05	0.23	0.00	1.00
Social Brokerage[a]	0.00	0.27	-0.66	0.46
Inverse Mills Ratio	2.33	0.23	1.31	2.74
Affilated RFC Authors	19.57	36.63	0.00	203.00
Affilated WG chairs	19.25	35.92	0.00	192.00

[a] mean centered to facilitate interpretation of interaction effects (Friedrich 1982).

[b] the omitted reference category is "Zero WGs".

Table 4: Bi-variate correlation matrix for network risk set (N=5066).

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	Summer Meeting	1.00																
2	Winter Meeting	-0.50	1.00															
3	Non-N. American Meeting	0.55	-0.39	1.00														
4	(ln) Community Size	0.02	0.06	0.21	1.00													
5	Incumbent WG Chairs	0.01	0.07	0.20	0.95	1.00												
6	Affiliated Executive	-0.02	-0.01	-0.02	-0.03	-0.02	1.00											
7	IETF Executive	-0.01	-0.01	-0.04	-0.20	-0.18	0.31	1.00										
8	Patents	0.00	0.00	0.01	0.07	0.05	0.04	-0.05	1.00									
9	Community Presence (Inst.)	0.02	0.00	0.11	0.38	0.41	0.14	0.06	0.04	1.00								
10	(ln) IETF Publications	0.00	0.00	-0.01	-0.04	-0.03	0.07	0.18	0.02	0.10	1.00							
11	Executive Tie	0.01	0.02	0.03	0.16	0.17	0.03	-0.13	0.07	0.12	0.21	1.00						
12	Degree	0.01	0.03	0.08	0.38	0.41	0.09	-0.02	0.08	0.25	0.28	0.47	1.00					
13	WG Contributor	0.01	0.02	0.05	0.26	0.25	0.01	-0.05	-0.03	0.22	0.06	0.09	0.31	1.00				
14	Two WGs	0.01	0.01	0.03	0.12	0.13	0.00	-0.02	0.02	0.11	0.12	0.06	0.18	0.26	1.00			
15	Three+ WGs	0.00	0.01	0.02	0.09	0.09	0.03	0.04	-0.01	0.10	0.42	0.16	0.32	0.11	-0.13	1.00		
16	Social Brokerage	0.00	0.01	0.07	0.24	0.27	0.07	0.02	0.08	0.18	0.17	0.23	0.56	0.20	0.21	0.21	1.00	
17	Inverse Mills Ratio	0.01	0.03	0.06	0.28	0.24	-0.45	-0.05	-0.39	-0.10	-0.09	-0.07	-0.03	0.01	0.02	0.02	-0.05	1.00

Table 5: Cox hazard models of appointment as Working Group Chair among IETF participants, 1989-2002.

	1	2	3	4	5
Summer Meeting	-0.427	-0.419	-0.392	-0.399	-0.397
	(0.137)**	(0.136)**	(0.137)**	(0.137)**	(0.137)**
Winter Meeting	-0.029	-0.009	0.017	0.011	0.012
	(0.115)	(0.115)	(0.116)	(0.116)	(0.116)
Non-N.A. Meeting	0.043	0.035	0.029	0.030	0.028
	(0.159)	(0.158)	(0.158)	(0.158)	(0.158)
IETF Executive	1.127	0.249	-0.463	-0.251	-0.216
	(0.273)**	(0.275)	(0.284)	(0.286)	(0.286)
Affiliated Executive	0.642	0.482	0.530	0.492	0.486
	(0.109)**	(0.109)**	(0.109)**	(0.109)**	(0.109)**
Patents	0.067	0.031	0.041	0.032	0.031
	(0.034)+	(0.040)	(0.039)	(0.040)	(0.040)
(ln) Community Size	0.104	0.320	0.389	0.399	0.417
	(0.165)	(0.166)+	(0.166)*	(0.166)*	(0.166)*
Incumbent WG Chairs	-0.016	-0.017	-0.018	-0.018	-0.019
	(0.003)**	(0.003)**	(0.003)**	(0.003)**	(0.003)**
Community Pres. (Inst.)	0.046	-0.057	-0.099	-0.098	-0.094
	(0.161)	(0.160)	(0.158)	(0.158)	(0.159)
IETF Pub. Dummy		2.025		0.887	0.663
		(0.099)**		(0.188)**	(0.240)**
(ln) IETF Publications			1.678	1.137	1.399
			(0.070)**	(0.143)**	(0.223)**
IETF Pubs. squared					-0.006
					(0.004)
Log Likelihood	-4132.69	-3964.11	-3948.15	-3937.45	-3936.11
Observations	243405	243405	243405	243405	243405
Individuals	15466	15466	15466	15466	15466
Events	473	473	473	473	473

Standard errors in parentheses

+ significant at 10%; * significant at 5%; ** significant at 1%

Table 6: Cox hazard models of appointment as Working Group Chair among IETF participants, 1989-2002.

	6	7	8	9
Summer Meeting	-0.133	-0.106	-0.106	-0.089
	(0.275)	(0.275)	(0.275)	(0.276)
Winter Meeting	0.344	0.353	0.352	0.346
	(0.243)	(0.245)	(0.245)	(0.245)
Non-North American Meeting	0.307	0.295	0.295	0.273
	(0.298)	(0.297)	(0.298)	(0.297)
IETF Executive	0.162	-0.071	-0.074	-0.115
	(0.382)	(0.386)	(0.387)	(0.385)
Affiliated Executive	0.128	0.156	0.155	0.125
	(0.217)	(0.218)	(0.218)	(0.219)
Patents	-0.042	-0.046	-0.047	-0.052
	(0.089)	(0.089)	(0.089)	(0.088)
(ln) Community Size	0.447	0.525	0.527	0.568
	(0.358)	(0.360)	(0.361)	(0.361)
Incumbent WG Chair Population	-0.017	-0.017	-0.018	-0.019
	(0.006)**	(0.006)**	(0.006)**	(0.006)**
Executive Tie	-0.106	-0.208	-0.214	-0.244
	(0.135)	(0.136)	(0.145)	(0.138)+
WG Contributor	0.701	0.625	0.618	0.677
	(0.285)*	(0.287)*	(0.292)*	(0.297)*
Community Presence (Instrumented)	-0.037	-0.041	-0.042	-0.053
	(0.126)	(0.125)	(0.125)	(0.126)
(ln) IETF Publications		0.878	0.867	0.772
		(0.228)**	(0.245)**	(0.230)**
Degree			0.003	
			(0.022)	
Social Broker (-constraint)				0.871
				(0.376)*
Log Likelihood	-593.21	-586.82	-586.81	-584.06
Observations	5066	5066	5066	5066
Individuals	610	610	610	610
Events	117	117	117	117

Standard errors in parentheses
+ significant at 10%; * significant at 5%; ** significant at 1%

Table 6 (cont.): Cox hazard models of appointment as Working Group Chair among IETF participants, 1989-2002.

	10	11	12	13	14
Summer Meeting	-0.091	-0.115	-0.116	-0.107	-0.120
	(0.276)	(0.276)	(0.274)	(0.275)	(0.275)
Winter Meeting	0.342	0.329	0.343	0.343	0.332
	(0.244)	(0.245)	(0.246)	(0.245)	(0.245)
Non-North American Meeting	0.281	0.290	0.295	0.300	0.311
	(0.299)	(0.299)	(0.298)	(0.298)	(0.299)
IETF Executive	-0.231	0.014	-0.034	-0.159	-0.268
	(0.386)	(0.388)	(0.389)	(0.390)	(0.401)
Affiliated Executive	0.103	0.150	0.145	0.136	0.264
	(0.219)	(0.218)	(0.219)	(0.220)	(0.245)
Patents	-0.056	-0.038	-0.033	-0.032	0.015
	(0.088)	(0.088)	(0.086)	(0.085)	(0.094)
(ln) Community Size	0.562	0.472	0.542	0.530	0.330
	(0.356)	(0.362)	(0.365)	(0.363)	(0.403)
Incumbent WG Chair Population	-0.020	-0.018	-0.018	-0.018	-0.017
	(0.006)**	(0.006)**	(0.006)**	(0.006)**	(0.006)**
Executive Tie	-0.301	-0.227	-0.223	-0.221	-0.204
	(0.140)*	(0.136)+	(0.137)	(0.145)	(0.146)
WG Contributor	0.724	0.440	0.514	0.670	0.685
	(0.299)*	(0.302)	(0.316)	(0.327)*	(0.327)*
Community Presence (Instrumented)	-0.029	-0.048	-0.041	-0.022	-0.011
	(0.124)	(0.126)	(0.127)	(0.125)	(0.125)
(ln) IETF Publications	0.703	0.478	0.501	0.579	0.592
	(0.228)**	(0.272)+	(0.275)+	(0.278)*	(0.278)*
Degree				-0.030	-0.027
				(0.027)	(0.027)
Social Broker (-constraint)	1.398		1.142	1.975	1.957
	(0.419)**		(0.439)**	(0.493)**	(0.491)**
Social Broker X Comm. Presence	1.160			1.275	1.265
	(0.288)**			(0.289)**	(0.288)**
Two WGs		0.539	0.511	0.506	0.489
		(0.227)*	(0.225)*	(0.225)*	(0.226)*
Two WG X Social Broker			-2.017	-2.261	-2.222
			(0.878)*	(0.895)*	(0.894)*
Three+ WGs		1.049	1.382	1.384	1.353
		(0.369)**	(0.445)**	(0.443)**	(0.445)**
Three+ WG X Social Broker			-3.292	-3.692	-3.749
			(1.657)*	(1.760)*	(1.760)*
Inverse Mills Ratio					0.694
					(0.633)
Log Likelihood	-576.14	-581.89	-576.82	-567.21	-566.58
Observations	5066	5066	5066	5066	5066
Individuals	610	610	610	610	610
Failures	117	117	117	117	117

Standard errors in parentheses
+ significant at 10%; * significant at 5%; ** significant at 1%

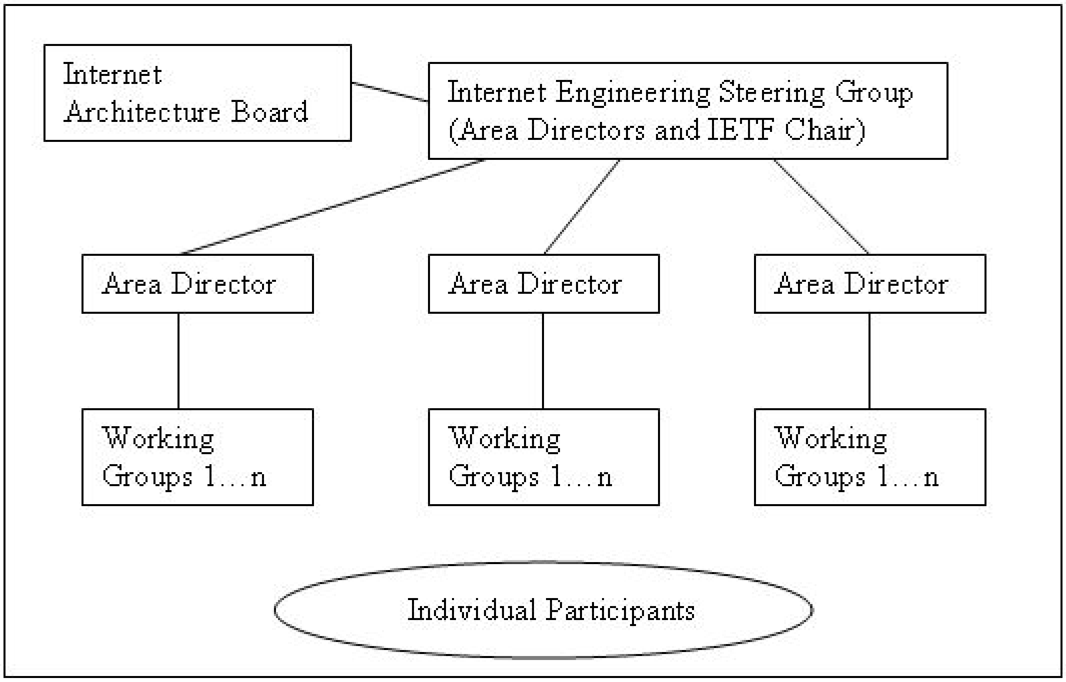

Figure 1:Diagram of IETF structure. Adapted from Steve Coya's IETF newcomer's orientation presentation entitled "IETF Structures and Internet Standards Process," http://ietf.org/newcomer/frame.htm.

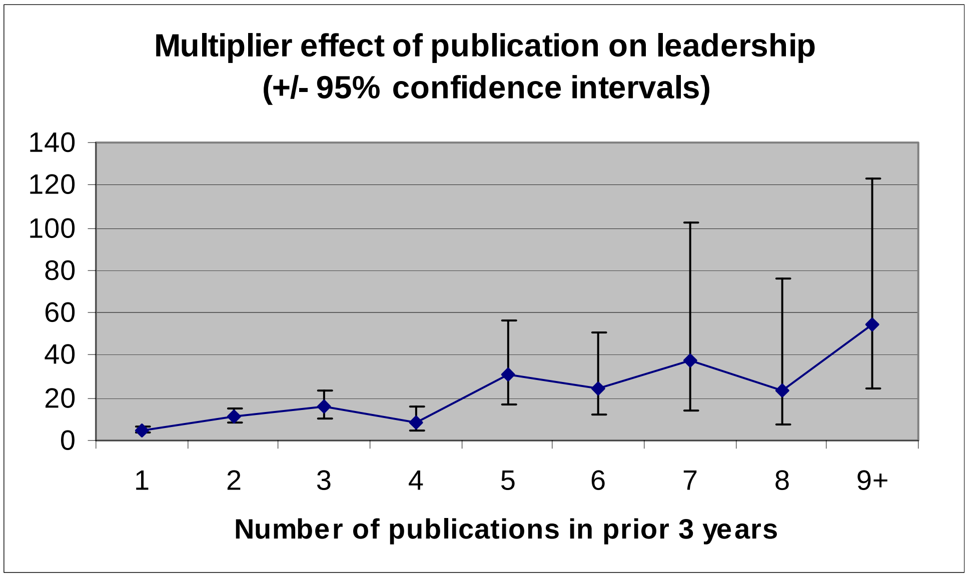

Figure 2: Publication spline for full risk set.

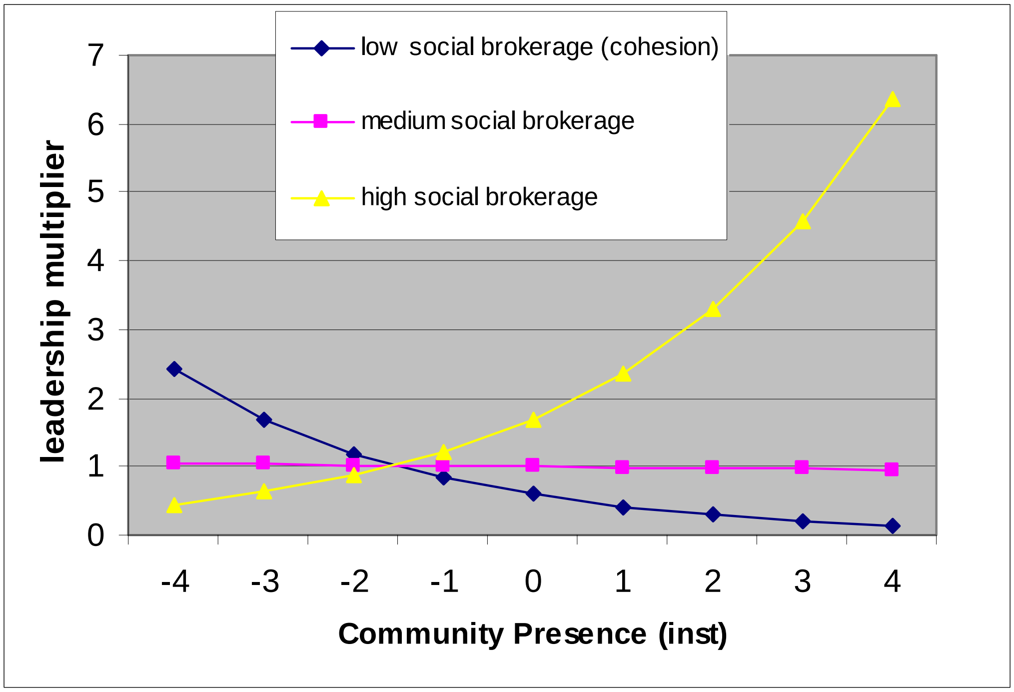

Figure 3: Interaction between social brokerage and community presence (mean +/- one standard deviation of brokerage from model 14).

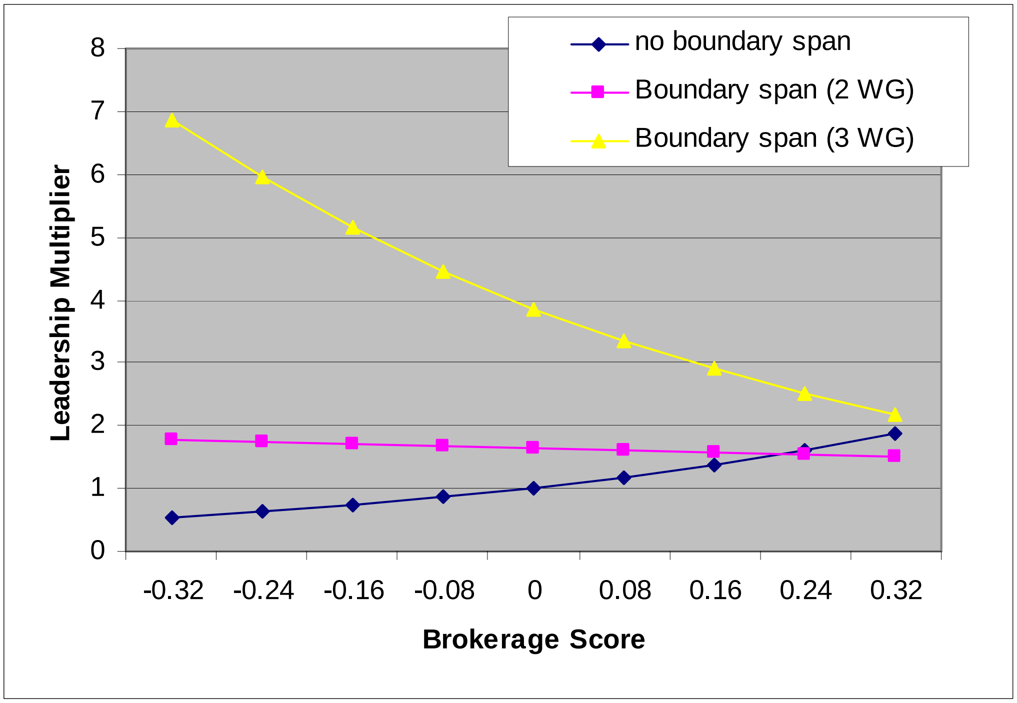

Figure 4: Interaction between technical boundary spanning and social brokerage (model 14).