Boundary spanning in a for-profit research lab: Scientific publication and the interface between commerce and academe

Chris Liu & Toby Stuart

(authorship alphabetical)

Abstract

In innovative industries, private sector firms increasingly are participants in open communities of science or technology. As part of the norms of exchange and engagement in such communities, firms often publicly disclose what would otherwise remain private discoveries. In a quantitative case study of one firm in the biopharmaceutical sector, we explore the consequences of scientific publication—an instance of public disclosure—for a core set of activities within the firm. Specifically, we link publications to human resource practices, finding in researcher-level, fixed effects regressions that bonuses are tied to publications. Second, using a unique electronic mail dataset, we show that researchers within the firm who author publications are much better connected to external (to the firm) members of the open scientific community. This result directly links publishing to current understandings of firms' absorptive capacity. Third, in an unanticipated finding, our analysis raises the possibility that the firm's most prolific publisher begin to migrate to the periphery of the intra-firm social network, which may occur because their strong external relationships induce them to reorient their focus to a community of scientists beyond the firm's boundary.

I. Introduction

A burgeoning literature investigates the porous boundary between the firm and the university. This work has developed along two macroscopic streams. First, a number of studies examine the emergence of the university as an engine of entrepreneurship, singling out its role in spawning startup companies (e.g., Zucker and Darby 1996; Shane and Stuart 2001; DiGregorio and Shane 2003), as well as the significant rise in faculty patenting rates and faculty engagement in other forms of technology transfer (Mowery et al. 2001; Agrawal and Henderson 2002; Owen Smith and Powell 2003; Azoulay, Ding, and Stuart 2007; Colyvas and Powell 2007). A second body of work approaches the interface from the reverse direction; it evaluates the potential gains to for-profit firms for contributing to open Science, especially the role of academic publishing in the development of firms' innovative capacity (Cohen, Nelson and Walsh 2002; Cockburn and Henderson 1998 Liebeskind *et al*. 1996; Stern 2004; Lim 2009; Murray 2004).

This is a large literature, but distilled, much of it concerns the process of boundary spanning. On one side of the divide, entrepreneurial faculty members have ventured into the world of commerce by building relationships and reputations in industry. On the other side, company researchers and dealmakers have navigated the academic landscape, seeking access to the community of scholars. The obstacles and incentives to traverse the university-industry "divide", however, differ on the two sides. For university faculty, the literature has pondered the collision and potential reconciliations of traditional scientific norms and values with the exigencies of commercial science, most notably the need for formal intellectual property rights on research advances. In negotiating their roles in industry, academic scientists have grappled with the normative challenges of appropriating private returns to a supposedly public good—scientific knowledge—and the construction of role identities that can accommodate juxtapositions between open and commercial Science. By contrast, as firms have adopted publication policies that result in private knowledge crossing into the public sphere, the questions they face surround the balance between the time and disclosure costs that are incurred when research staff publish their scientific findings, versus the potential benefits of open publication policies, including access to informal networks of science.

In this paper, we place the spotlight on, or, more precisely, *inside*, a life sciences firm (hereafter, "BTCO"). This particular company owes its existence to entrepreneurial boundary spanners; BTCO's cofounders hailed from both academic and private sector backgrounds. Along with a group of other biotechnology industry pioneers, BTCO heralded the emergence of a new type of company with unusually permeable boundaries and the adoption of core organizational design elements that were modeled after universities. As we will show, BTCO possesses an impressive publication record commensurate with its deep-seated ties to academe and its inheritance of strong scientific values at birth. Today, BTCO is a research-intensive organization employing many Ph.D. scientists; it has successfully recruited prestigious, senior scientists from prominent university appointments; its researchers have published many articles in highly regarded journals; and the internal organization of research at the company parallels a biology department's structure: its core organizational units are laboratories. We therefore regard BTCO as the archetype of a straddler: it is a for-profit company but, in many ways, it behaves like a university. Indeed, the integrative activities of this and similar organizations have been a central force in the erosion of the boundary between for-profit and public Science.

The general concern of our paper is the relationship between publishing, the allocation of rewards within the firm, and the structure of the communication network inside and beyond the borders of the organization. We ask three specific questions. First, at the researcher level, what are the effects of publishing on discretionary compensation in the firm? Or, put in different terms, does the firm pay scientists to contribute to open Science (cf. Stern, 2004)? Second, does publishing influence the architecture of communication networks inside and beyond the boundaries of BTCO? And, in examining these aspects of BTCO, what broader lessons might we draw about the within-firm consequences of for-profit companies' contributions to communities of open science and innovation? To address these questions, we exploit a rich and unique data archive that includes demographic, publication, and compensation information for all researchers in BTCO. In addition, although for a shorter duration of time that will only permit analysis in the cross section, BTCO has provided daily downloads of all electronic mail of the members of the research division. These electronic correspondences enable us to observe the correlates of publishing on the shape of the within-firm network, as well as basic characteristics of firm-to-university communication patterns.

We focus on three core findings. First, we confirm that BTCO rewards successful publishers. Using internal compensation data, regressions including researcher fixed effects show that publication success substantially increases the percentage of target bonuses that researchers' receive in a given year. Next, we utilize email data to examine the networks of publishers relative to non-publishers. Using these data, we report three results: first, prolific publishers are the recipients of a greater proportion of the messages sent by their immediate supervisors. Thus, not only do they receive comparatively more remuneration, publishers attract greater shares of their managers' attention than do non-publishers. Second, researchers who publish have significantly more correspondents in (predominantly, prestigious) universities. Using email data, we are able to provide direct evidence that publishing enables members of BTCO to access the informal networks in the invisible colleges of open Science.

The third finding, however, intimates a trade-off: there is a negative relationship between researchers' publication counts and their centrality *within* the BTCO email network. This result was unanticipated. Given BTCO's heritage and values, its pro-publication strategy, and the apparent norms for allocating discretionary compensation and managerial attention, we reasoned that researchers' standing in the broader academic community would reach inside the company to order its internal status hierarchy. Although these considerations suggest that prolific publishers will be central in the company's communication network, our findings imply one or more offsetting processes: as boundary spanning researchers become increasingly embedded in the academic community, they in fact show signs of migrating toward the periphery of the internal BTCO communication network. Either because of time constraints that require them to eschew internal interactions in which they might otherwise engage, because they value an identity that is more purely associated with academic science, or simply because the act of publishing spawns relationships that draw researchers into the collaborative networks in academe, prolific publishers begin to withdraw from certain internal interactions.

This result, we believe, raises a challenge for maximizing the benefits of boundary spanning: in a world of widely distributed scientific expertise, the individuals within an organization who are most well networked beyond its boundaries are precisely those people who would ideally occupy central positions within the firm. Yet, it is these very same individuals who have the greatest opportunity to shift the locus of their interaction to communities beyond the boundaries of the firm.

II. Publishing and Social Networks in the Private Sector

Two, related insights frame the literature on the publication strategies of private-sector organizations. First, in science- and technology-based industries, the knowledge base that is the foundation for innovation can be very broadly distributed—so much so that Powell, Koput and Smith-Doerr (1996) conceptualize the locus of innovation as residing in highly variegated networks, rather than within the boundaries of single organizations (or, for that matter, even single organizational forms). In contexts such as biopharmaceuticals, software development, medical devices, and microelectronics, innovation is a process of spotting and borrowing: actors must spot discoveries that are pertinent to them and then borrow externally generated insights to seed their own, internal development efforts. The second idea is absorptive capacity (Cohen and Levinthal 1990). In essence, to identify and assimilate externally developed ideas, organizations must first create the capacity to absorb. This is accomplished by investing in basic research to cultivate scientific and engineering understandings, and by encouraging researchers within the firm to connect to ideas that are developed beyond it.

These two ideas, we argue, drive for-profit firms' decisions to publish. Moreover, their reach extends to implications for the adoptions of a set of human resource practices to manage in-house researchers (Cockburn and Henderson 1998), and they imply advantageous forms of communication within and across the firm's boundaries.

II.a The Locus of Innovation. Scholars have long recognized that science and technology are collective, cumulative endeavors. In classical accounts as well as historical and evolutionary perspectives, the development of technology and the progression of science are understood to be processes in which new discoveries are improvements to, or novel combinations of, antecedent ones (e.g., Schumpeter 1942; Basalla 1988; Dosi 1988; Tushman and Rosenkopf 1992). Building on this understanding, those who study the sociology of technology often employ the metaphor of a "seamless web" to describe the many and multiplex relationships among participants in the development of any technical field (Hughes 1987; Pinch and Bijker 1987). A hallmark of this work, and of sociological characterizations of the innovation process in general, is its emphasis on the relational context in which innovation unfolds (Podolny and Stuart 1995)—new discoveries are never regarded as *de novo* creations; even path-breaking inventions emerge from antecedents that fall within the continuity of an interconnected set of ideas.

The canvas painted in this broad-brushed work aligns with findings from analyses of the innovation process in contemporary science- and technology-based industries. For example, in a set of empirical papers, Zucker and Darby illustrate the dependence of the biotechnology industry on the scientific discoveries of star scientists in universities and research institutes (Zucker and Darby 1996). In a case study, Liebsekind et al. (1996) demonstrate two biotechnology firms' use of external social networks to source scientific discoveries through entry into multiple collaborative research projects with academic scientists. These authors argue that a network-based approach is most suited to "prospecting" for external developments in an increasingly vast scientific landscape.

Drawing us back to the macro consequences of actor-level efforts to build connections with other participants in a technical arena, Powell et al. (2005) illustrate the implications of diffuse expertise for the collaborative structure in the overall organizational field in biotechnology. They depict the evolution of the network among firms, universities, research institutes, and financiers, and the changing rules of attachment that appear to drive the aggregate structure of the field-wide network over time. In a growing set of technical and scientific fields, then, a central task for innovators is to devise a strategy for developing points of contact with the individuals and organizations who collectively architect a field of ideas.

II.b Absorptive Capacity. In what is now one of the most familiar ideas in the literature on organizational learning, Cohen and Levinthal (1989, 1990) argue that the background knowledge required for innovative activity is cumulative: new ideas are aptly assimilated only if a solid foundation of understanding has been put in place. First, background knowledge is a prerequisite for opportunity identification. Without detailed knowledge of a particular area of science or technology, actors may not understand the significance of new opportunities in the area and may even lack the ability to formulate feasible questions to explore. Second, even if new opportunities were recognized, a lack of sufficient background understanding effectively excludes the ability to exploit external developments to further internal innovation objectives.

If we accept the premise that the production of knowledge is widely distributed, then absorptive capacity hinges on a means to reach beyond the boundary of the firm to screen monitor, appraise, and perhaps assimilate external development that are deemed relevant. For example, Cockburn and Henderson (1998) propose that ties to universities are an essential element of the R&D process in the pharmaceutical industry; they find that R&D productivity is correlated with having staff scientists who coauthor with university faculty. Lim (2000) pushes the link between these ideas a step further: in a study of the diffusion of copper interconnect technology among semiconductor producers, he frames absorptive capacity specifically in terms of connectedness. In his fusion, external connectedness itself determines absorptive capacity. Of course internal R&D remains paramount, but its function is largely to enhance firms' access to external knowledge sources.

Organizations enact multiple, often concurrent and complementary strategies to achieve external connectivity in domains of distributed innovation. First, they enter myriad, formal collaborative agreements to exchange, license, or co-develop technologies (e.g., Ahuja 2000; Stuart 2000), which assemble into a dense alliance network within communities of innovators (Powell et al. 2005). Second, knowledge traverses organizational boundaries through employee mobility (e.g., Almeida, Dokko and Rosenkopf 2000; Rosenkopf and Almeida 2003) and through organizational members' participation in formal knowledge sharing venues, such as standard setting bodies and industry associations (Rosenkopf, Metiu and George 2001). A third avenue of interchange is through myriad, informal associations. These range from participation in open source communities to the cultivation of informal collaborative relationships between members of a focal firm and other actors in the broader innovation arena.

Our analysis explores the latter phenomena. We consider the multifaceted consequences of a private sector firm's participation in open Science. Viewed narrowly, open publication is just a manifestation of a corporate policy to permit the selective disclosure of the firm's research discoveries and, in instances of co-authorship with researchers from other organizations, an incomplete image of the set of informal scientific collaborations in which the firm is embedded. (The network is incomplete because much informal collaboration yields outputs other than published articles, such as simple idea exchange or material transfers.) However, we believe that the consequences of open publication are considerably broader: echoing the findings of prior work, the firm's policy vis-à-vis publication may affect its ability to recruit researchers, its policies for allocating rewards, its capacity to foster a broad network of informal collaborators, and even the status ordering and social structure within the firm.

III. Hypotheses

For some time, scholars struggled to understand what seemed a puzzling phenomenon—given the costs, why do companies permit employees to publish and present scientific and technical findings in the venues of open Science? The costs of publication are borne in at least three forms. First, substantial expenses are incurred in the consumption of employee time to craft research results into publications and to shepherd articles through peer review. In fact, given the sizeable time costs of writing and revising research papers, BTCO's current management has recently introduced policies to reduce the number of submissions to second- and third-tier academic journals.[1] Second, publication is disclosure. Although it is possible to time the submission of publications so that they do not interfere with filing for intellectual property claims, firms that publish unavoidably disclose a great deal of information about the focus of their research endeavors. Thus, open publication is tantamount to a revelation of strategic intent.

Third, publication contributes to the conversion of firm-specific human capital to its general form, which in turn increases employee mobility and bargaining power. When firms permit researchers to publish, they not only endow specific individuals with the credit for their discoveries, but de facto they divulge this allocation to the public. It then becomes possible for external parties to link a firm's interesting technical developments to the specific individuals who contributed most to their creation. Efforts by competitors to poach talent is an inevitable result.

[1] Interestingly, BTCO management emphasized that they were not discouraging public disclosure of scientific findings (conference submissions and presentations are encouraged), but they actively discourage the submission of these results to low quality journals. In other words, generating peripheral papers is viewed to be off the critical path to the firm's bottom-line.

What, then, are the compensatory benefits that offset these costs, and what are the implications of an open publication policy for how an organization behaves? In our discussions at BTCO (which closely parallel those reported in Cockburn and Henderson 1998), interviewees underscored two points. First, a permissive publication policy is an essential component of any strategy to recruit and retain the highest quality researchers, especially individuals who hold doctoral degrees. Second, our interviewees suggested that it is necessary not just to permit researchers to publish their work in the outlets of open Science, but also to take one step further: employees of the firm should be rewarded based on their standing in the larger scientific community.

This brings us to a larger point, which is that the labor market for top caliber researchers itself contributes to the blurred boundary between academic and commercial science. Because private sector firms must compete with universities and research institutes for talent, firms attempt to create university-like milieu to cater to the preferences and values held by the researchers whom they endeavor to recruit. After years in graduate school and, in many cases, additional training as post-doctoral fellows, candidates for employment will have extensive exposure to the norms of and reward system in open Science. This means that potential recruits for whom firms will compete will value publications as a core element of their professional identity. Moreover, they view publications as the currency of professional achievement, and may prefer employment systems in which internal rewards reflect the professional esteem accorded to publication.

There is an additional benefit of tying compensation levels to publication outcomes. Not only may this be a matter of employee preference, but pegging rewards to publications potentially helps firms to resolve a perennial dilemma: how does one evaluate and reward researchers who work on very long term and highly uncertain projects, the vast majority of which will fail to deliver revenues for the firm (and none will do so in the near term)? Under these circumstances, peer-reviewed publications provide a semi-objective method of evaluating performance to allocate discretionary compensation in a context in which the quality of research is difficult to assess, and effort is challenging to measure. We therefore hypothesize,

H1: Within BTCO, researchers' remuneration will increase in their publication success.

If encouraging publication is necessary to recruit talented scientist, success in this activity is vital to building the firm's external network. In short, publications are the passkeys to the invisible colleges of open Science. When scientists publish important findings, they gain the visibility that leads to invitations to conferences and seminars; they attract the interest of potential collaborators; they become nodes in discussion networks about new developments in a field; and they establish the types of relationships that result in the exchange of tacit knowledge that circulates in the networks of the profession. In the informal rules of exchange in open Science, access to these networks is contingent both on contributing to the corpus of public science, and correlated with the importance of the contributions one makes.[2] In other words, the connections that are developed through publication are central to the accumulation of absorptive capacity in domains of distributed scientific discovery. We hypothesize,

H2: Within BTCO, publishers will occupy more central positions in the informal networks in academe.

Inspired by cautionary tales such as Xerox's historic failure to capitalize on the myriad, groundbreaking discoveries developed at its Palo Alto Research Center (PARC) (cf. Mitzik, 1999), the literature on absorptive capacity underscores that a firm's ability to apply its existing knowledge base toward further innovation depends critically on the patterns of communication and distribution of knowledge within an organization. Thus, there are internal, formal and informal organizational components to absorptive capacity: investing in basic research is necessary, but it may be insufficient for persistent innovation. For example, Sorensen and Stuart (2000) argue that when organizations age, they tend toward an increased rigidity and an ossification of communication patterns among positions and roles within the firm, which hamper innovation. Mowery and Rosenberg (1991) observe that the objective of basic research often is not to produce a good per se; it is to create the understandings that lay the groundwork for subsequently developing the good. Because these foundational understandings are often exploited in other areas of the organization, these authors admonition that when there are silos within basic research or between it and the rest of the firm, research is prone to become "sterile and unproductive."

We know that one objective of open publication strategies is to embed the firm's researchers in scientific networks. This means that publishing scientists are metaphorically on the boundary of the firm. Their focus is partly external, as the value they bring to the organization is enhanced when they build relationships in the broader research community. At the same time, the maximization of this value may well depend on the positions that externally networked scientists occupy in the communication flows *within* the organization. As the aforementioned studies highlight, innovative organizations must cultivate intraorganizational processes and networks to transfer knowledge from their periphery to the core. In practice, this means that the individuals on the boundary of a firm must be well networked inside of it. [See J.P. ASQ 1994]

There are thus functional reasons for why it is advantageous that the individuals with the deepest scientific expertise and a broad range of beyond-the-firm connections to become central within the firm's internal network. But independent of any practical purpose, it is likely that researchers' standing in the broader scientific community will contour their social positions

---

[2] In addition to publications, BTCO also has a policy of actively sharing reagents with the external scientific community. The exchange of reagents and other research materials is another illustration of conformance to community norms that further contributes to the embedding of BTCO researchers in the exchange system of open Science. In addition to this indirect benefit, sharing reagents also enables BTCO to directly observe external research that builds upon their proprietary materials, which is another low-cost mechanism of external monitoring.

inside the firm. Because of the close connection between universities and organizations like BTCO, because many such firms have been imprinted with the scientific values of academic founders, and because these firms have cultures that often resemble university departments, we anticipate an organic correlation between positions in the internal status hierarchy and contributions to open Science.

H3: Within BTCO, publication will be associated with central positions in the firm's internal communication network.

IV. Data and Methods

a. Context. To test these hypotheses, we have collected data at one prominent biotechnology firm, BTCO. The biopharmaceutical industry has served as a fertile testing ground for much of the literature on the relationship between innovation and collaboration across organizations and organizational forms. To name just a few examples, the influential studies of Cockburn and Henderson (1998), Zucker and Darby (1996) and Powell, Koput and Smith-Doerr (1996) have all examined this industry.

BTCO is a first-generation biotechnology firm, founded more than 25 years ago. Since the time of its founding, BTCO has dedicated significant resources to in-house research, and today its research division comprises hundreds of scientists. The mandate of the firm's research group, distinct from its development arm, is to conduct basic and applied research to feed molecules into the drug development pipeline.

In line with the firm's historical origins, the internal organization of BTCO's research division is modeled after a university biology department. Researchers are subdivided into groups that map to scientific specializations, including immunology, neurobiology, molecular biology, and oncology. These groups are then further divided into the core organizational units, which are laboratories led by (and named after) individual scientists.[3] Though we analyze different subsets, the company provided current and some historical data on all members of the research division.

b. **Publications**. BTCO scientists have published extensively—in recent years, the firm's staff has produced well over 100 papers per year—and they have succeeded in placing some of their work in the preeminent outlets in life science publication, including journals such as *Science*, *Nature*, and *Cell.*

To measure the publication outputs of all individuals in the firm's research department, we collected all articles by BTCO authors that were indexed in the ISI *Web of Science*. We hand matched the roster of employed scientists to the list of authors on papers to correct for spelling discrepancies. In addition, we collected Journal Impact Factors, which allow us to weight publication counts by for the prestige of the publishing venue. Finally, we gathered information on whether or not papers were coauthored with non-BTCO individuals. In most of the results

[3] Laboratories are formed (and disbanded) in concert with the employment of the laboratory head. If the laboratory head departs, lab-members are reassigned.

we will report, we considered all contributors to be equivalent, regardless of their position within the author list.

C. Compensation and Rewards Structure. Quantifying individual contributions to the biopharmaceutical development pipeline is a notoriously challenging problem. At BTCO, scientists are eligible for three distinct forms of merit-based compensation, which are allotted to recognize individuals' contributions to the firm. First, all members of the research division receive stock option grants. The size of these grants is a forward-looking measure that reflects the company's desire to retain certain employees. In addition, the firm dispenses end-of-year bonuses that are backward-looking measures of performance that reflect each employee's contribution to the company over the prior year. Over the course of the year, the department's total research bonus pool increases as a number of pre-set milestones are met. At year end, each manager is given a customized target bonus for each of their direct reports, which is determined by the size of the total bonus pool, the individual's salary-band and other responsibilities. After receiving a target bonus, managers adjust the target (up or down) to reflect perceived performance. Importantly, each laboratory is not forced into a normal curve, although BTCO Research as a whole approaches one. A second, distinct bonus pool is distributed to "top contributors", which includes approximately 5% of all scientists who have been identified as star performers in the prior year.

We combined these two numbers to create a "proportion of target bonus-received" for each scientist, which is used to test hypothesis 1 (publication success will influence bonus allocations).[4] For the median individual in the dataset, end-of-year bonus is approximately 20% of their base salary.

d. Network Data. To map the network structure within and beyond the borders of BTCO's research organization, the company provided us with log files that contain a record of all emails that are exchanged via the company's servers. These data were archived each day and then sent to us. We have taken two steps to insure the privacy of company employees. First, before transferring the email logs to us, BTCO's IT staff stripped the subject headings and email content from all files. Second, in constructing the dataset we analyze, after matching publications to individual names but before we merged in compensation data, the company assisted us in replacing all names with hashed identification numbers.

In meetings with senior leadership and rank-and-file members of BTCO, we were repeatedly told that BTCO is an "email place" and that a great deal of the research division's business is conducted over email exchanges on the company's servers. This assertion is consistent with the ebb and flow of email traffic in the data, which very much confirm a priori suppositions about when communications are most likely to occur in the company, and who in the organization is most likely to be most central in the network.

---

[4] Because of the presence of the extra "top contributor" bonus pool, this proportion is centered 1.05 rather than 1.0.

For all cross-sectional analyses, we used the email logs from the month of January 2009. Before aggregating the daily emails into a sociomatrix for the month, we deleted all messages with more than four recipients to cull broadcast mailings (Quintane and Kleinbaum).

We use the email data to construct three measures of individuals' positions. First, using a detailed organizational chart provided by the company, we are able to identify the immediate supervisor of all individuals in the dataset. To analyze the proportion of a supervisor's attention devoted to each BTCO scientist, we matched the email network to the reporting structure of the firm. For each manager, we counted the number of direct reports and computed the expected proportion of each report's electronic in-volume from the manager. We then divided the observed over the expected in-volume to generate a weighted measure of each report's "managerial attention."

The email data are limited to messages that reach BTCO's servers and we have very limited information about individuals outside the firm who communicate with BTCO researchers. However, for all incoming messages, we were able to retain senders' email addresses. To measure the range of each individual's connections to scientists in universities, we count the unique number of email correspondents that contain a *.edu identifier[5]. We assume that these emails are a residue of ties between BTCO scientists and colleagues in academic institutions, and that the greater the degree score for an individual in BTCO, the better networked he or she is likely to be in academic circles.

Lastly, we use the email network ties to construct a dichotomous, symmetrized adjacency matrix for all BTCO research employees. Although emails are directed ties (indeed, we differentiate between sender and receiver in the previous two measures) the vast majority of communicating pairs within BTCO participate in reciprocal interactions. Thus, for the purpose of identifying internal network positions, we treat correspondences as symmetric ties. As a measure of researchers' positions within the BTCO email network, we use this adjacency matrix to calculate betweenness centrality.

V. Results

We begin our discussion of results with a set of descriptive statistics. Table 1 reports the recent history of publishing and patenting at BTCO. These statistics provide interesting insight into the scientific strategy of the firm. First, the company has published papers and filed patents in a ratio of approximately 2:1 favoring papers. Second, almost exactly half of the scientific articles BTCO has published in the past seven years have been coauthored with researchers at universities, and at least 40% of coauthored papers were written with collaborates at one of the five most prestigious universities in the life sciences.

Table 2 lists, in order of frequency, the universities with which BTCO staff have coauthored the greatest number of papers and with which they have exchanged the most electronic messages. There are two things of note in this table. First, the table underscores the fact that BTCO scientists have established relationships with collaborators and colleagues at

[5] Note that this means that we undercount the number of interactions with scientists in industry because limiting the data to *.edu addresses basically eliminates correspondences with individuals at non-US-based universities.

many elite institutions in the academic life sciences. And second, while there is clearly overlap between the rosters of institutions where the firm has collaborators and coauthors, there are differences as well. As expected, the list of communication partners in universities is both broader and different from the roster of coauthors. Thus, although the coauthorship graph does inform the true information exchange network, it substantially misrepresents the network's shape, reach, and density.

To provide a greater sense for the data, a lower bound on the number of emails exchanged between members of BTCO's research staff and individuals at *.edu email addresses in January and February of 2009 was 10,244. This number excludes virtually all non-U.S. universities and many research institutes, so the actual volume of communication is much higher than this. Moreover, when we break down the aggregate number by type of researcher, we find that active publishers who hold Ph.D. are, by a wide margin, the most extensively networked to scientists in industry. For instance, Ph.D. holders who do not publish averaged 26 email exchanges with individuals at *.edu addresses in January 2009, while Ph.D.s who do publish average 88 email exchanges. For the 19 individuals who are the most prolific publishers in the firm, this number doubles again—they engage in an average of 182 email exchanges with *.edu contacts.

Panels A and B of Table 3 present, respectively, descriptive statistics for the scientists in our dataset in both the panel and the cross section. Consistent with Smith-Doerr's (2004) examination of gender issues in the scientific workforce in the biotechnology industry and with National Science Foundation data on the gender composition of recent Ph.D. cohorts in the life sciences, women actually makeup slightly more than 55% of the scientific staff at BTCO. Of note, Panel B shows that in 2008, approximately 70% of the researchers did not author any scientific papers. For those who hold doctoral degrees, the proportion of non-publishers falls to about 60%. We then bin publishers into three categories, low, medium, and high, to allow for a flexible spline specification of the effect of publication on the outcome variables we analyze. The omitted category is zero publications in a year.

Recall that researchers' target bonus payouts are centered on 1.05 to reflect the addition of compensation from the "key contributors" pool to produce an aggregate measure of incentive compensation to reflect the previous year's performance. Based on the ranges in Panel A, 0 to 2.24, it is clear that managers' perceive and recognize a wide range of performance. Figure 1 illustrates the overall distribution of target bonus, which is approximately normal for research overall and relatively widely distributed.

Table 4 presents the first set of regression results, which examine the effect of publication on researchers' annual bonuses. For these regressions, no email data are required and we were able to exploit the full 7-year panel that includes compensation, publication, and reporting structure data. In this and the subsequent table, we report results based on two different cuts of the data. In the columns labeled "Full", the data include all members of the research organization. In the columns labeled "PhD Only", we subset the data to limit them to Ph.D. holders only. If our hypotheses are correct, we expect that the results will be stronger for the subsample of doctoral degree holders. Obviously, these individuals take primary responsibility for the firm's publication, and their rewards and networks should be much more consequentially influenced by publication activities.

The findings strongly support the first hypothesis. Column (1), which is based on the full sample, shows that in a within-person model, high publishers are estimated to receive a 6.5% increment to their target bonus in years in which they fall into this category. Column (2) excludes publications and shows that the firm similarly rewards individuals who are listed as inventors on many patents, and Column (3) confirms that the effect of the contribution of open publications to bonuses is independent of the one of the filing of proprietary IP rights. Columns (4-6) then repeat each of these regressions for the subsample of doctoral degree holders, and the results for this cut of the data are stronger and cleaner. In these regressions, there is evidence of a monotonic increase in the effect of publication on discretionary bonuses across the three levels of publication. Moreover, the effect of falling in the high publication category more than doubles, to approximately 14.5 percent.

Table 5 reports on the relationships between publications and individuals' positions in the communication structure within the firm, as well as between the firm and *.edu contacts. We explore three different dependent variables: "supervisor attention", the span-of-control-weighted proportion of outgoing messages a supervisor directs to the focal individual; indegree in *.edu correspondents, or the distinct number of contacts in academe who email the focal individual; and the individual's betweenness centrality in the BTCO email network. In general, the findings in Table 5 parallel those in Table 4 with respect to the partitioning of the data; the results are generally stronger when we subset on Ph.D.s, whose efforts are more directly tied to the publication outputs of the firm. Finally, because we do not have a multi-year panel of email data, all of the regressions in Table 5 are estimated in a cross section—we correlated 2008 publication records with January 2009 network data.

In a finding that we regard as reinforcing results on the effect of target bonus, Column (4) of Table 5 shows that not only do high publishers garner greater bonuses, they also monopolize a higher share of supervisor attention. Among the Ph.D.s in the firm, the parameter estimate suggests that, after adjusting for salary grade, gender, and tenure, high publishers attract an additional 64% of a supervisor's attention relative to non-publishers.

Columns (2) and (5) report quasi-maximum likelihood Poisson estimates of the count of university indegree in the BTCO email data. In the full dataset (2), all three levels of publishing covariates are positive, but none are statistically significant. In Column (5) in which we restrict the sample to Ph.D.s, however, the effect becomes monotonic and high publishers have a statistically significant, greater level of connectivity with university researchers. The coefficient indicate a 70% higher indegree, even after conditioning on salary grade and lab head status.

The final pair of regressions, Columns (3) and (6), examine the determinants of individuals' betweenness centralities in the firm's email network in January 2009. In both columns, the (cross sectional) results indicate that betweenness centrality correlates positively with firm tenure, but at a decreasing rate. Also in both columns, men have statistically significantly *lower* levels of betweenness centrality than do women. (In unreported analyses of eigenvector and closeness centrality, there is no gender effect.) In the full dataset, Column (3) shows a strong correlation between centrality and educational level; relative to the omitted category (Ph.D.-holder), individuals with masters and bachelor degrees are much less central in the internal email network.

Columns (3) and (6) contain an unexpected finding. Our third hypothesis forecasted a positive association between individuals' publication counts and their centralities within the internal network. Prolific publishers clearly are distinguished in terms of the quantity and quality of their science. In a research organization with deeply held scientific values and a steadfast belief in the power of novel science to drive the drug development pipeline, we expected a positive effect of scientific standing on internal centrality. However, not only do we reject the hypothesis that the largest publishers will be central in the firm's network, but the opposite finding obtains—very active publishing is negatively correlated with individuals' betweenness centrality.

What might account for this unexpected finding? As we reconsider the possibilities, the logic of McPherson, Popielarz and Drobnic (1992) looms large. In their analysis of the dynamics of voluntary group memberships, McPherson *et al*. find that group attachments are the shifting outcome of competing forces: the number of cohesiveness of one's ties within a group, versus the strength of ties to members of different groups. These authors show that turnover in group memberships depends on the balance of these relational forces. Viewed in this light, the internal centrality finding becomes quite reasonable. As a direct function of their contributions to open Science, prolific publishers within the firm appear to greatly strengthen and extend their relationships beyond it. In consequence, they are naturally drawn toward identification with and greater commitments within the external research community, which may come at the expense of internal interaction.

If this supposition is correct, then there is a trade-off associated with pro-publication policies. On one hand, the evidence in the paper unambiguously shows that publication contributes to BTCO's external connectivity, which in turn is a core element of the firm's absorptive capacity. But a byproduct of the external connectivity gained through publication may be to draw those individuals who develop rich connections beyond the firm to the periphery of the internal networks, which raises the specter of the search-transfer problem paradox described in Hansen (1999). In effect, those who become most able to identify promising external developments (search) may, in so doing, compromise the internal networks that facilitate transfer.

Table 1: Descriptive Stats on Yearly Publishing (limited to individuals who appear in this dataset)

Year	# of Patents	# of Papers	Papers w/ Universities	Papers w/ ”Top 5”**	Papers w/ Industry	Papers in Cell/Nature/Science
2001	83	210	108	45	32	5
2002	76	156	69	24	22	4
2003	77	150	74	27	27	6
2004	63	164	82	26	25	12
2005	77	149	78	25	29	5
2006	92	136	60	25	27	9
2007	26*	161	89	36	27	10

*incomplete data collection.
**”Top 5” are collaborations with Harvard University, MIT, Stanford, UCBerkeley, or UCSF.

Table 2: Prevalent Institutions of Coauthors and Correspondents

Rank Order	Coauthored Universities	Count	Email Correspondence in January, 2009	Count
1	UCSF	102	UCSF	753
2	Stanford	79	Stanford	553
3	Harvard	71	Salk Institute	123
4	UCLA	48	UCDavis	117
5	Duke	34	UCBerkeley	98
6	Yale	32	Yale	82
7	UColorado-Denver	31	U. of Iowa	74
8	UWashington	30	Harvard	64
9	UPenn	24	U. of Chicago	58
10	NIH	22	UCLA	49

Table 3: Descriptive Statistics

Panel A: Pooled Cross-Section Descriptive Statistics (n = scientist-years = 1964)

	Mean	SD	Min	Max
Age	39.210	8.704	22	69
Male	0.462	0.499	0	1
Highest Education-BA	0.368	0.482	0	1
Highest Education-MA	0.233	0.423	0	1
Highest Education-PhD	0.399	0.490	0	1
Firm Tenure	5.964	6.543	0	30
Firm Tenure^2	78.356	148.615	0	900
Lab Head	0.24	0.427	0	1
No Publications	0.768	0.422	0	1
Low Publications	0.136	0.343	0	1
Med Publications	0.057	0.231	0	1
High Publication	0.040	0.195	0	1
Patents	4.561	69.216	0	1670
% of Target Bonus Received	1.058	0.254	0	2.47

Panel B: Descriptive Statistics for Year 2008 (n = scientists = 348).

	Mean	SD	Min	Max
Age	40.730	8.537	24	70
Male	0.445	0.498	0	1
Highest Education-BA	0.345	0.476	0	1
Highest Education-MA	0.224	0.418	0	1
Highest Education-PhD	0.431	0.496	0	1
Firm Tenure	7.147	6.547	1	31
Firm Tenure^2	93.813	173.957	1	961
Lab Head	0.216	0.412	0	1
No Publications	0.698	0.460	0	1
Low Publications	0.195	0.397	0	1
Med Publications	0.057	0.233	0	1
High Publication	0.049	0.216	0	1
% of Target Bonus Received	1.057	0.251	0	2.143

Table 4: Fixed Effects (Panel) Linear Model on Share of Discretionary Bonus

	(1)	(2)	(3)	(4)	(5)	(6)
Dataset	**Full**	**Full**	**Full**	**PhDs Only**	**PhDs Only**	**PhDs Only**
Low Pubcount	0.0205		0.0203	0.0488+		0.0480+
	(0.0140)		(0.0140)	(0.0249)		(0.0248)
Med Pubcount	0.0078		0.0073	0.0674*		0.0659*
	(0.0210)		(0.0210)	(0.0326)		(0.0325)
High Pubcount	0.0651*		0.0667*	0.1424**		0.1463**
	(0.0284)		(0.0283)	(0.0397)		(0.0396)
# of patents		0.0002**	0.0002**		0.0002*	0.0002*
applied-yearly		(0.0001)	(0.0001)		(0.0001)	(0.0001)
tenure	0.0164**	0.0180**	0.0178**	0.0259**	0.0310**	0.0288**
	(0.0046)	(0.0046)	(0.0046)	(0.0083)	(0.0084)	(0.0084)
tenure^2	-0.0003	-0.0004*	-0.0003	0.0001	-0.0003	0.0001
	(0.0002)	(0.0002)	(0.0002)	(0.0004)	(0.0004)	(0.0004)
Constant	0.9303**	0.9400**	0.9288**	0.8758**	0.9073**	0.8650**
	(0.0558)	(0.0554)	(0.0557)	(0.0766)	(0.0764)	(0.0765)
Observations	1963	1963	1963	782	782	782
R-squared	0.03	0.04	0.04	0.11	0.09	0.12
rho	1	1	1	1	1	1
F-test	5	5	5	6	6	6
# of scientists	543	543	543	209	209	209

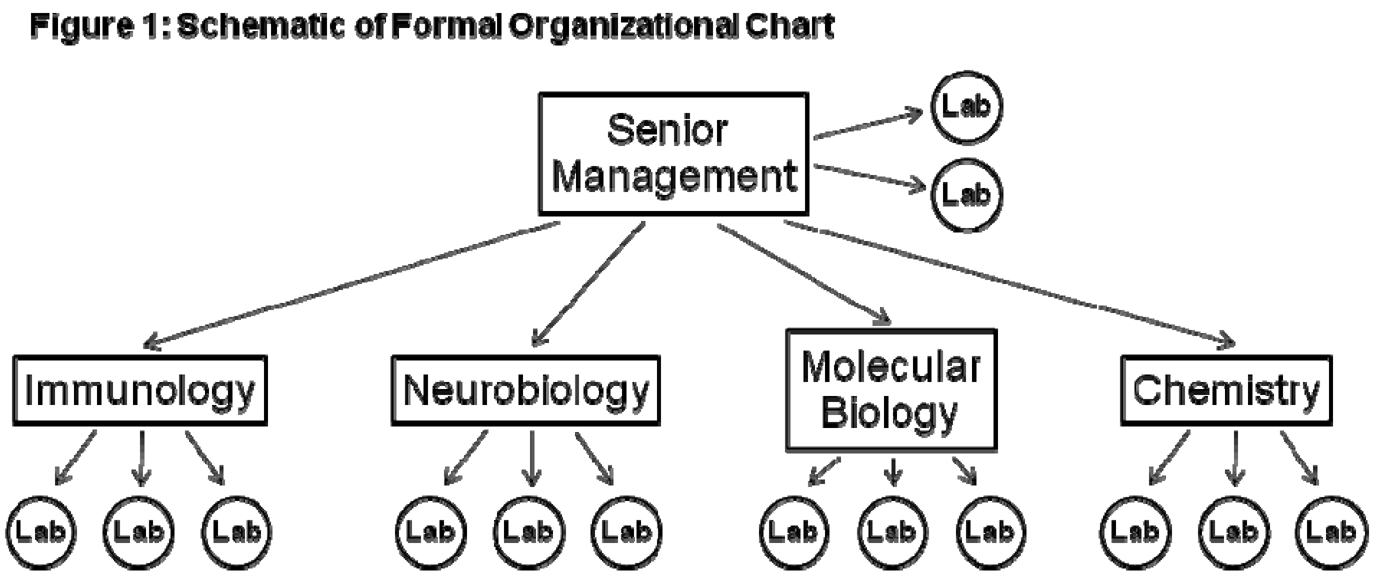
Note: Estimates are displayed as raw coefficients. All models include salary-band dummies and year dummies. These variables are not shown. Robust standard errors in parentheses below; + significant at 10%; * significant at 5%; ** significant at 1%

Table 5: Impact of Scientist Publishing

	(1)	(2)	(3)	(4)	(5)	(6)
Dataset	Full	Full	Full	PhDs Only	PhDs Only	PhDs Only
Model	Linear	Poisson QML	Poisson QML	Linear	Poisson QML	Poisson QML
Dependent Variable	proportion of supervisor attention-log	January '09 University Indegree	January '09 Between-ness Centrality	proportion of supervisor attention-log	January '09 University Indegree	January '09 Between-ness Centrality
Low Pubcount	0.369*	0.021	0.276+	0.508*	0.066	0.340+
	(0.184)	(0.131)	(0.160)	(0.251)	(0.199)	(0.195)
Med Pubcount	0.482*	0.415	0.334	0.555+	0.368	0.388
	(0.228)	(0.296)	(0.214)	(0.318)	(0.348)	(0.286)
High Pubcount	0.283	0.411	-0.634**	0.638*	0.527*	-0.606*
	(0.260)	(0.298)	(0.239)	(0.281)	(0.258)	(0.263)
male	-0.121	0.052	-0.286*	0.128	0.069	-0.400**
	(0.147)	(0.113)	(0.126)	(0.174)	(0.159)	(0.143)
Laboratory Head	-0.499	0.288	0.438	-0.725+	0.637*	0.422
Indicator	(0.300)	(0.213)	(0.288)	(0.377)	(0.287)	(0.302)
company tenure	-0.063	0.136**	0.123**	-0.145+	0.056	0.168**
	(0.050)	(0.037)	(0.035)	(0.083)	(0.057)	(0.048)
tenure-squared	0.001	-0.004**	-0.004**	0.006+	-0.002	-0.006**
	(0.002)	(0.001)	(0.001)	(0.004)	(0.002)	(0.002)
laboratory size	-0.022	0.002	-0.005	-0.064	0.037	-0.031
	(0.019)	(0.016)	(0.018)	(0.046)	(0.024)	(0.022)
Bachelor's-	0.078	-0.004	-0.606**			
highest degree	(0.231)	(0.163)	(0.216)			
Master's-highest	0.049	-0.147	-0.505**			
degree	(0.230)	(0.137)	(0.145)			
Constant	-2.614**	0.010	-1.079*	-2.168**	-0.013	-0.946*
	(0.833)	(0.447)	(0.466)	(0.591)	(0.562)	(0.475)
Observations	378	392	389	179	192	191
R-squared	0.17			0.24		
Log-pseudolikelihood	-649	-981	-130	-286	-562	-82
# of lab clusters	82	82	82	71	76	76

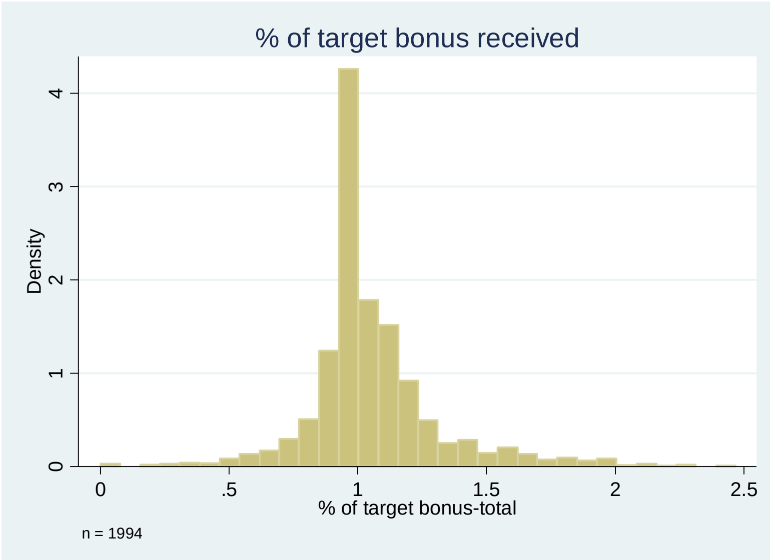
Note: Estimates are displayed as raw coefficients. Errors are clustered at the laboratory level. All models include salary-band dummies and division dummies. These variables are not shown. Robust standard errors in parentheses below; + significant at 10%; * significant at 5%; ** significant at 1%.

Figure 1: Organizational Chart of BTCO



Note: A schematic of the BTCO organizational chart. Fundamental units are laboratories, organized into 4 major, disciplinary divisions.

Figure 2: Share of Discretionary Bonus



Note: Each manager is provided (by Human Resources) with a customized target bonus for each of their direct reports. This target is then adjusted to reflect performance. We present received/target bonus to reflect a weighted measure of performance in each year.