MANAGEMENT SCIENCE
Vol. 56, No. 9, September 2010, pp. 1439–1461
ISSN 0025-1909 | EISSN 1526-5501 | 10 | 5609 | 1439

informs®
DOI 10.1287/mnsc.1100.1195

The Impact of Information Technology on Academic Scientists' Productivity and Collaboration Patterns

Waverly W. Ding
Haas School of Business, University of California, Berkeley, Berkeley, California 94720,
wding@haas.berkeley.edu

Sharon G. Levin
Department of Economics, University of Missouri–St. Louis, St. Louis, Missouri 63121,
slevin@umsl.edu

Paula E. Stephan
Andrew Young School of Policy Studies and NBER, Georgia State University, Atlanta, Georgia 30302,
pstephan@gsu.edu

Anne E. Winkler
Department of Economics, University of Missouri–St. Louis, St. Louis, Missouri 63121,
awinkler@umsl.edu

This study investigates the impact of information technology (IT) on productivity and collaboration patterns in academe. Our data combine information on the diffusion of two noteworthy innovations in IT—BITNET and the Domain Name System (DNS)—with career-history data on research-active life scientists. We analyzed a random sample of 3,114 research-active life scientists from 314 U.S. institutions over a 25-year period and find that the availability of BITNET on a scientist's campus has a positive effect on his or her productivity and collaborative network. Our findings also support the hypothesis of a differential effect of IT across subgroups of the scientific labor force. Women scientists and those working at nonelite institutions benefit more from the availability of IT in terms of overall research output and an increase in the number of new coauthors they work with than do men or individuals at elite institutions. These results suggest that IT is an equalizing force, providing a greater boost to productivity and more collaboration opportunities for scientists who are more marginally positioned in academe.

Key words: diffusion; innovation; technology; life sciences; professional labor markets; gender
History: Received August 13, 2009; accepted March 8, 2010, by Lee Fleming, entrepreneurship and innovation. Published online in *Articles in Advance* July 2, 2010.

1. Introduction

The Internet and other advancements in information technology (IT) have changed the workplace (e.g., Brynjolfsson 1993, Kelly 1994, Dewan and Kraemer 2000). The impact of IT arguably plays a particularly important role in the production of knowledge, given that scientific inquiry is highly dependent on equipment, materials, and the knowledge and skills of others, access to which can be greatly enhanced by IT.

Despite the conviction that IT has brought about major changes in scientific research and contributed to increased productivity in science, few studies have looked at how the adoption of IT directly affects productivity and collaboration patterns of individual scientists. Instead, investigation of the relationship between IT and productivity has largely been inferred using time-period measures of the availability of IT rather than actually establishing the availability or use of IT. (A notable exception in this regard is Agrawal and Goldfarb 2008.) Moreover, some of these studies rely on data aggregated at the journal or institutional level. The few studies that have investigated the impact of IT on productivity and collaboration patterns of individual scientists with access to IT have generally relied on self-reported data on IT usage. Although this approach has much to recommend it, a weakness is that it is almost impossible to date accurately when the technology initially became available to the scientists surveyed in the studies. Another weakness is that these studies generally rely on cross-sectional data, thereby making the proper identification of causal effects a challenge.

This paper addresses many of these deficiencies. Namely, this study combines longitudinal data on individual scientists with explicit measures of whether IT was available at the scientist's institution in order to examine whether the availability of IT led to systematic differences in the scientist's research outcomes. The longitudinal data come from a random sample of over 3,000 research-active life scientists working at universities and colleges between the years 1969 and 1993. These data are linked to explicit information regarding the date that the institution adopted one of two early innovations in IT: BITNET and a domain name (e.g., http://www.gsu.edu). The data also contain detailed information on institutional and individual characteristics related to productivity.

Our empirical work is based on a model of scientific productivity that postulates that scientific output depends upon inputs such as effort, materials, equipment, and knowledge. We argue that IT affects access to inputs and reduces communication costs. Three specific hypotheses follow from the theoretical framework: (1) IT has a productivity-enhancing effect; (2) IT has a collaboration-enhancing effect; and (3) IT has a "democratizing" effect. The latter is tested by examining the extent to which IT increases the productivity of individuals at the margin of the profession, defined here to be women and individuals working at nonelite institutions. Our hypotheses are tested using our longitudinal sample. Productivity is measured using counts of publications, and collaboration is measured using the gain in new coauthors. The empirical work closely matches the theoretical model set forth. Controlling for institutional and individual characteristics related to productivity, we find evidence for productivity and collaboration enhancing effects of IT. Our results also support the democratizing effects of IT: women benefit from IT relative to men; individuals at nonelite institutions benefit relative to those at elite institutions.

Our results inform policy and practice in several respects. First, although the lag may be long, the evidence is convincing that scientific research contributes to economic growth (Adams 1990). Thus, innovations that contribute to increased research output such as IT arguably contribute to growth. Second, there is a long history of studying the "outer circle" of scientists, and policies such as affirmative action have been created to try to lower barriers between the outer and inner circle. IT, to the extent it has a democratizing effect, can be a powerful tool to accomplish this. Third, the IT revolution continues. A starting point for understanding how new innovations will affect productivity is to study how early advances in IT contributed to productivity and collaboration.

The plan of this paper is as follows: Section 2 presents the theoretical framework and our hypotheses. Section 3 reviews the extant literature concerning the impact of IT on publishing outcomes and collaboration. Democratizing effects of IT are also discussed. Section 4 provides information on BITNET and the Domain Name System (DNS). Section 5 describes the data, estimation strategy, and variables. Section 6 presents the empirical findings. Section 7 provides the conclusions and discussion.

2. Scientific Productivity, Collaboration, and IT

We use a simple model of scientific productivity (Stephan and Levin 1992) to formalize how IT affects productivity and collaboration. We assume that the production of scientific knowledge requires effort, materials, equipment, skills, and knowledge. The relative proportions are discipline dependent:

$$P = f(\text{effort, materials, equipment, skills, knowledge}), \tag{1}$$

where P is some measure of output such as an article, measured at the individual level. Effort, materials, equipment, skills (which includes ability), and available knowledge represent the inputs. We assume the marginal product of each argument to be positive; we also assume that the scientist does not supply all of the arguments directly but can draw on others to increase resource intensity. For example, equipment can be augmented by using that in another lab, materials can be borrowed or exchanged, and graduate students can provide expertise that they learned working in another scientist's lab. From (1) it follows that other things being equal, an increase in any argument increases P.

A challenge for studies of scientific productivity is that most of the arguments in the production function are difficult to measure directly. As a result, studies of scientific productivity usually rely on proxies to measure inputs. The skill and knowledge base of the scientist, for example, is often inferred by characteristics of the institution where the scientist trained. Access to materials and equipment is generally proxied by a measure of the quality of the institution where the scientist works. Other things being equal, higher quality institutions have more equipment and can provide access to materials more readily. If an individual does not have a needed piece of equipment, it is likely that it can be found in another's lab. The same is true for materials. Johns Hopkins School of Medicine's Transgenic Core Laboratory, for example, can custom-make a mouse genetically altered to suit a scientist's research purpose (Anft 2008). Scientists working at highly rated programs also have access to colleagues who possess greater knowledge and skills than those working at lower-tier institutions. The knowledge and skills of others who may work in the lab can be proxied by measures of quality of the graduate program and the availability of graduate research assistants by the size of the graduate program. Moreover, because many of the resources used in research (including the scientist's time) are funded through external grants, one would expect a positive relationship between funding and productivity.[1]

We argue that IT, by lowering the cost of communication, provides increased access to the knowledge of others. Prior to the advent of IT, the knowledge of others who were not in the same geographic space was available to a researcher only through face-to-face contact or via telephone exchange. Both were relatively expensive.

IT also lowers the cost of accessing codified knowledge. Again, prior to the availability of IT, accessing codified knowledge required physically going to the library or making a request for material to be sent from another library or host. Innovations in IT also improve the way knowledge is stored and retrieved. Journal articles are now available online through services such as JSTOR and PubMed, and data can be transmitted more quickly and at a lower cost.

IT also has the potential of increasing access to equipment and materials. The Internet, for example, allows for remote access to equipment such as synchrotrons and telescopes (Stephan 2010). It also enhances access to materials by allowing individuals to request materials online.

Thus, other things being equal, IT is expected to lead to an increase in P.[2] The effects are expected to be discipline dependent given that the relative importance of the arguments in the production function are discipline specific. Moreover, Walsh and Bayma (1996) and Walsh et al. (2000) find that the way researchers incorporate IT into their work varies by field. The way in which IT enhances research has also changed over time. For example, the degree to which IT is used to access codified knowledge only began to take off after the development of the World Wide Web.

Hypothesis 1 (H1). *Access to IT increases productivity.*

Both the knowledge and material/equipment effect come about through increased interaction with others, leading to a second hypothesis that access to IT increases collaboration. It is not only that IT provides greater access to materials and equipment, but also that IT allows researchers to share ideas with each other at a lower rate and across greater distances, holding other resources constant.

Hypothesis 2 (H2). *IT leads to increased collaboration.*

Access to the knowledge and skills of others has arguably become increasingly important because of the rate of growth in the breadth and depth of scientific inquiry. Jones (2009), for example, argues that scientists are acquiring narrower expertise and that scientific production has become increasingly specialized. At the same time, many of the breakthroughs in scientific research have come about through interdisciplinary research, and scientific discoveries have become more dependent on access to specialized materials and equipment (Stephan 2010). Because of these changes, collaboration has become increasingly important in science as scientists with differing skills,[3] knowledge, materials, and equipment work together on a question of joint interest. This suggests that the effect of IT on collaboration may have become increasingly important over time, a hypothesis which the relatively short span of our data precludes us from testing.

There is also a strong argument to be made that the effects of IT may differ across subsets of the target population. Barley (1986) and Orlikowski's (1992) work suggests that whether a technological innovation is adopted and how it is used depends on the type of individuals using it and the organizational environment in which the innovation is introduced. We further argue that the *relative incremental* effect that access to IT has on productivity and collaboration depends on how marginally positioned a scientist is in the scientific hierarchy.[4] Those at the "top" have access to strong colleagues, excellent graduate students, and state-of-the-art equipment. Those at the "bottom" have considerably less access. Thus, the effect of IT is greater for those at the margin than those at the top.[5]

To investigate the democratizing effect of IT, we focus on two attributes of a scientist: employer's institutional tier (elite versus nonelite) and gender.[6] To

[1] Funding is arguably based on cumulative productivity, not current productivity (Xie and Shauman 1998), mitigating the concern that funding is endogenous in a reduced form estimation of Equation (1).

[2] Our theoretical underpinnings for how IT affects the output of individual scientists is consistent with the theoretical framework developed by DeLone and McLean (2003). In their model of information systems success, they emphasize the downstream impact of technology on job outcomes such as satisfaction and performance. They also discuss how outcomes depend on organizational factors such as system quality and service quality, factors which we are unable to control for in our empirical analysis.

[3] Rosenblat and Mobius (2004) as well as Van Alstyne and Brynjolfsson (2005) argue, on the other hand, that the Internet tends to encourage individuals who share narrow expertise and interests to work together.

[4] Our hypothesis is consistent with the work of Sproull and Kiesler (1992, p. xii) who find that "computer-based communication can reduce the isolation of physically and socially peripheral workers through increasing organizational participation and personal ties."

[5] Our hypothesis is that those at the margin benefit relatively more from the availability of IT than those who are not at the margin. A related hypothesis that builds on the work of Burkhardt and Brass (1990) focuses on who within an organization in terms of power and social network position is the first to adopt when a new technology is introduced.

[6] Our choice of dimensions of stratification to investigate is informed by prior research on scientific careers (for a review, see Long and Fox 1995). Another dimension worthy of investigation is minority status, but the data do not permit such an inquiry.

further elaborate, and building upon Equation (1), scientists who come from top-tier institutions usually have access to more research inputs than those at lower-tier institutions. As a result, the incremental effect that IT has on productivity and on collaboration should be greater for those at lower-tier institutions, leading to the following hypotheses.

Hypothesis 3A (H3A). *Access to IT has a greater positive effect on productivity for scientists employed at lower-tier universities than for those at higher-tier universities.*

Hypothesis 4A (H4A). *Access to IT has a greater positive effect on collaboration for scientists employed at lower-tier universities than for those at higher-tier universities.*

We also expect the effect of IT to differ by gender. In addition to their lack of access to resources due to their positions in the academic hierarchy, women also face obstacles arising from family and child-care obligations (Ginther and Khan 2009). For example, women scientists may travel less often to conferences and seminars, where academic networking takes place, ideas are exchanged, and collaboration opportunities emerge. Another way of expanding one's academic network is through job mobility. However, there is evidence that the mobility of women scientists is more constrained than that of men (Marwell et al. 1979), particularly when they have children (Shauman and Xie 1996). Thus, we expect that women scientists, who traditionally are marginally positioned in academia, have less extensive networks, and face greater mobility constraints, benefit more from IT in terms of increased productivity (H3B) and gains in coauthors (H4B).

Hypothesis 3B (H3B). *Access to IT has a greater positive effect on productivity for female scientists than for male scientists.*

Hypothesis 4B (H4B). *Access to IT has a greater positive effect on collaboration for female scientists than for male scientists.*

3. Literature on the Relationship Between IT, Productivity, and Collaboration

Our hypotheses regarding the impact of IT on the productivity of scientists have received some previous attention. Specifically, empirical research has looked at how advancements in IT enhance scientists' productivity and collaboration regardless of their "location" in the profession and the extent to which IT enhances the productivity and collaboration of some subgroups (e.g., women or those employed by lower-tier institutions) more than others. The review that follows offers a representative sampling of this prior research and places the data and methodology used in this study in the context of this earlier work.

3.1. IT and Research Productivity

Investigations of the relationship between IT and research productivity generally have found support for the view that IT enhances productivity. Hesse et al. (1993) surveyed the subset of oceanographers who used the electronic network SCIENCEnet and found a positive relationship between frequency of use and both publication counts and professional recognition. Subsequent research by Cohen (1996) and Walsh et al. (2000), among others, expanded the number of disciplines surveyed to include philosophy, political science, and sociology, as well as math and a number of natural sciences, and found a relationship between IT usage and productivity. Winkler et al. (2010) found limited evidence of a positive IT–productivity relationship, using information on life scientists from the Survey of Doctorate Recipients (SDR) and institutional-level information on the adoption of various indicators of IT. Evidence of a positive IT–productivity relationship, where IT was measured using concurrent computer usage, was also reported by Kaminer and Braunstein (1998).

3.2. IT and Research Collaboration

A number of studies have identified a significant increase in the number of coauthored papers by individuals at different academic institutions and in different countries, as well as in the number of coauthors per paper. An analysis of approximately 13 million published papers in science and engineering from 1955 to 2000, for example, found an increase in team size in all but one of the 172 subfields studied, and average team size was found to have nearly doubled, going from 1.9 to 3.5 authors per paper (Wuchty et al. 2007).[7] Adams et al. (2005) found similar results for the top-110 research universities in the United States, reporting that the average number of authors per paper in the sciences grew by 53.4%, rising from 2.77 to 4.24 over the period 1981–1999.

Growth in the number of authors on a paper is due not only to a rise in collaboration within a university—and an increase in lab size—but more importantly to an increase in the number of institutions collaborating on a research project. A study of 662 U.S. institutions that had received National Science Foundation (NSF) funding one or more times found that collaboration across these institutions in science and engineering, which was rare in 1975, grew in each and every year between 1975 and 2005, reaching approximately 40% by 2005 (Jones et al. 2008). Collaboration has increased internationally as well. The Levin et al. (2009) study of authorship patterns

[7] Team size even increased in mathematics, generally seen as the domain of individuals working alone and the field least dependent on capital equipment.

across a wide array of four-year colleges and universities in the United States found that the percentage of papers with one or more international authors went from 6.6% in 1991 to 19.2% in 2007.

The coincidence of the increase in collaboration since the 1990s with the diffusion of several innovations in information technologies has not gone unobserved. Some of this research has focused on the field of economics. For example, Hamermesh and Oster (2002) compared publishing activity in three economics journals for the period 1970–1979 with that for the period 1992–1996. They found almost 20% of authors of jointly produced articles to be located at distant locations in the more recent period compared to 5% in the earlier period. Rosenblat and Mobius (2004) looked at coauthorship patterns in economics from 1969 to 1999 based on papers published in eight top economics journals. A novel feature of their study is they also look at the changing nature of the coauthorship—the degree of similarity of the author's research fields. Their analysis found that, at least in the field of economics, as communication costs fall, researchers seek to collaborate with more distant colleagues who share similar interests. Adams et al. (2005) also identified a growing mean distance among coauthors in their analysis of 2.4 million scientific papers for a number of disciplines, going from 77.7 miles in 1981 to 159.4 miles in 1999. They attributed the change to the improvement in electronic networks connecting scientists after 1987.

3.3. Differential Effects of IT

Studies have also looked at the degree to which IT has a "democratizing" effect and may have benefited some subgroups (e.g., women and those at lower-tier institutions) relative to others, thereby helping to level the research "playing field." In these studies, the role of IT is often inferred from period effects. Kim et al. (2009), for example, examined publishing productivity of faculty in economics and finance who were located at an elite institution at some point in time during the period 1970–2001. They found that the advantage to being located at an elite institution fell starting in 1970 and had, in fact, disappeared by the 1990s. Butler et al. (2008) examined collaboration (measured as coauthorship) across universities in the fields of economics and political science using publication data from three top journals in each field. They inferred the time that IT became available based on a review of NBER working papers published during the 1990s. They found that prior to January 1997 an e-mail address was never included; since January 1999 almost all papers have an e-mail address. Using this indicator of IT, they found that coauthorship increased with IT, especially at lower-ranked institutions. They also examined whether IT differentially affected women relative to men, but found no significant difference.

Hesse et al. (1993) used geographic location to proxy for institutional status, given that the more prestigious departments in oceanography tend to be located closer to the coasts and the less prestigious ones more inland. They found that geographically disadvantaged scientists received a higher productivity gain from IT. However, Cohen's (1996) survey of scientists from a broader set of disciplines found no support for the hypothesis of disproportionate benefits for scientists employed at lower-tier institutions.

Agrawal and Goldfarb (2008) examined the impact of BITNET, as measured by date of institutional adoption, on collaboration (coauthorship) at the institutional level. In their study, they used publication data from seven top journals in the field of electrical engineering for the period 1977–1991 and separated out institutional affiliations of authors into three groups: elite, medium, and lower tier. They found that faculty at medium-ranked research universities benefited the most from the adoption of BITNET in the form of increased collaboration with top-tier institutions and increased publishing productivity. Winkler et al. (2010) appended this same institutional-level measure of BITNET (plus several others) to individual-level data on a cross-section of life scientists drawn from the SDR. They found some evidence, albeit modest, that individuals at lower-tier institutions benefited relatively more from IT, but found no significant differences by gender.

Our study links detailed longitudinal data on life scientists to explicit measures of IT that reflect institutional adoption of IT, thereby advancing the extant literature in several ways. First, only with detailed data on the personal characteristics of scientists and the institutions in which they are located (e.g., Ph.D., tier, career stage, gender, research funding) can one adequately investigate the effect of IT on an individual scientist's productivity and collaboration network. Second, in such investigations longitudinal data is strongly preferred to cross-sectional data because it permits identification through differences in timing (Walsh et al. 2000). This is a major deficiency of the studies that estimate productivity as a function of a self-reported measure of IT use. Third, short of knowing the actual date that an individual adopts IT, it is far better to use the date that an institution adopted IT as an independent variable rather than to infer adoption of IT via time-period effects.

4. Early IT Innovations: BITNET and DNS

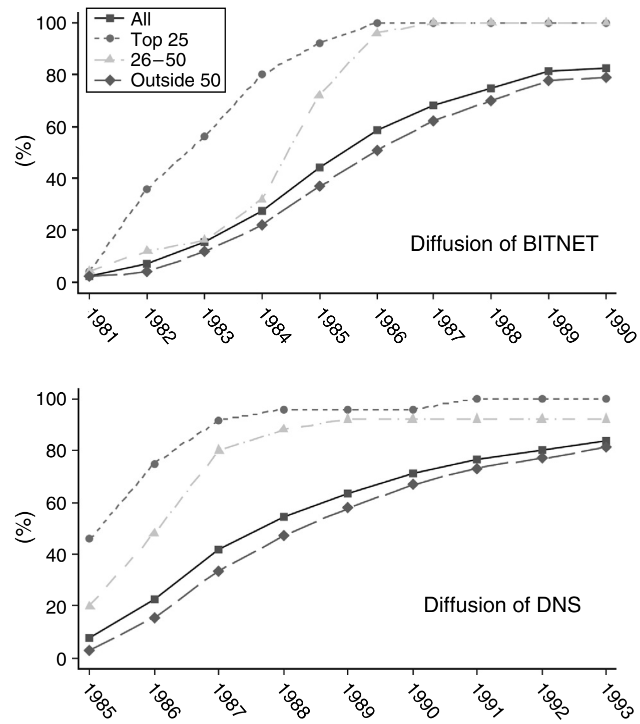
In this study, we analyze the research impact of two early indicators of IT: the availability of BITNET and a domain name (DNS) at a scientist's employing institution.[8] Previous work by Agrawal and Goldfarb (2008) used BITNET data and work by Winkler et al. (2010) examined both indicators.

The IT revolution can be dated to the creation of the Advanced Research Projects Agency Network (ARPANET) by the Department of Defense in 1969. Restricted access to ARPANET, however, led to the development of other networks (NSF 2009). Among these was BITNET, conceptualized by the vice chancellor of University Systems at the City University of New York (CUNY), and first adopted by CUNY and Yale University in May 1981. BITNET provided electronic communications across a range of scientific disciplines and universities. At its peak in 1991–1992, BITNET connected about 1,400 organizations (almost 700 academic institutions) in 49 countries (Corporation for Research & Educational Networking (CREN) 1997). Indeed, BITNET (plus connected networks) has been referred to as "the embryonic Internet" (Humphrys n.d.). By the mid-1990s BITNET was eclipsed by the Internet as we know it today and began to fade away.[9]

Different from the present-day Internet, BITNET linked university scientists using mainframe computers. It was a "store and forward" network built in a tree structure with only a single path from one computer (node) to another (node) (CREN 1997). BITNET enabled researchers to send messages and mail, transfer documents and data, and conduct online discussions (LISTSERVE). Nonetheless, the early "Internet" as characterized by BITNET was a forbidding, uninviting technology. Moreover, much of what is taken for granted today could not be done: researchers could not "attach" files to messages, access library resources, or search for information online. Indeed, it was not until 1991 that the World Wide Web became the first widely used global hypertext system making it possible to simply click on a hot spot in one document to jump to another, and it was not until 1993 that the first graphical Web browser, Mosaic (the forerunner to Netscape), was introduced.

An early and essential development in the Internet's evolution that contributed to its growth was the development of the Domain Name System (DNS) in 1984 (Zakon 2005). This system, which became the industry standard, classified addresses initially according to whether the host computer connecting to the network was an educational (edu), commercial (com), governmental (gov), or international (org) institution; it also provided for a series of country codes. No longer did every host on the Internet need to know the exact name and IP address[10] of every other system on the network, nor did it need to continuously update the file containing this information as the number of hosts on the Internet grew exponentially (DNS History, http://www.livinginternet.com).

Figure 1 **Diffusion of BITNET and DNS**



Notes. Cumulative percentage of institutions adopting BITNET and DNS. Diffusion patterns are graphed for all institutions (black with squares), institutions ranked in the top 25 (grey with circles), institutions ranked between 26 and 50 (light grey with triangles), and institutions not ranked (dark grey with diamonds), by the Gourman Reports (Gourman 1980–1997).

Figure 1 shows the diffusion of BITNET over the period 1981–1990 and the diffusion of DNS over the period 1985–1993 at the 314 academic institutions where the research-active life scientists in our study are located.[11] Both figures exhibit the typical S-curve associated with diffusion of an innovation over time (Rogers 2003)—especially among the nontop institutions; adoption first rises at an increasing rate and then levels off.

[8] Data for these indicators were initially collected for a set of 1,348 four-year colleges, universities, and medical schools in the United States that had been in existence since 1980. See Levin et al. (2010).

[9] By 1992–1993, the number of academic organizations connected to the Internet actually exceeded the number participating in BITNET, and by 1993, the number connected to BITNET began to fall. See BITNET History, http://www.livinginternet.com.

[10] The Internet Protocol (IP) address is a numerical address of four sets of numbers assigned to a specific computer. To communicate using the Internet, one must use an IP address. The DNS system provides an easy-to-remember name that maps into an IP address.

[11] Data on the adoption dates of BITNET beyond 1990 are not available. See Atlas of Cyberspaces (contact Martin Dodge at the University of Manchester for BITNET data, http://www.sed.manchester.ac.uk/geography/staff/dodge_martin.htm).

Diffusion patterns vary considerably by tier. Among the top 25 research institutions, BITNET and DNS diffused rapidly; in the case of DNS, approximately all of the top institutions had adopted the technology within a span of two years. The diffusion of BITNET was a bit slower, but among the top institutions, approximately all had access within five years. Diffusion was somewhat slower among the mid-tier institutions and considerably slower among those institutions outside the top 50. Across all tiers, 82% of institutions had access to BITNET and 83% had access to DNS by the end of 1993.

5. Data, Estimation Strategy, and Variables

5.1. Data

We use individual level data to analyze whether institutional availability of BITNET and DNS led to systematic differences in scientists' research outcomes. We begin our analysis with a random sample of 12,000 life scientists in the United States, drawn from the UMI ProQuest Dissertations database.[12] We restrict our sample to those who earned Ph.D.'s between 1967 and 1990 to isolate the effects of the first two major information technologies, BITNET and DNS, on research productivity and collaboration. We use the Web of Science's Science Citation Index to collect the publications, coauthors, and employment affiliations. Because our interest lies in the research outcomes of academic scientists, we retain only individuals who have publishing experience in academic institutions after receiving the doctorate, creating a data set of scientist-year observations from 1969 to 1993 with annually updated covariates for the individual and the employer institution.[13]

Each scientist begins in the data the year he or she receives a Ph.D. and continues until (i) there is a five-year interval during which that scientist does not publish, (ii) the scientist starts publishing exclusively under a corporate affiliation, or (iii) the year is 1993. We also exclude employment spans when a scientist temporarily affiliates with a nonuniversity research institution (e.g., one of the NIH-affiliated research institutes) because status of IT in such institutions cannot be determined. These restrictions result in a data set of 3,114 scientists with 32,398 scientist-year observations.

5.2. Estimation Strategy

We use a Poisson quasi-maximum likelihood estimator (PQMLE) to examine the effects of BITNET and DNS availability on a scientist's campus on his or her research productivity and collaboration. PQMLE is preferred to ordinary least squares regression because our dependent variables—count of research publications and count of new coauthors—are discrete variables. In addition, because the Poisson model is in the linear exponential family, the coefficient estimates remain consistent as long as the mean of the dependent variable is correctly specified. Thus, the PQMLE does not impose an equi-dispersion condition as in a standard Poisson estimator (Wooldridge 2002). Furthermore, "robust" standard errors are consistent even if the underlying data generating process is not Poisson. In fact, the PQMLE can be used for any nonnegative dependent variable, whether integer or continuous, as long as our conditional mean is correctly specified (Gourieroux et al. 1984, Santos Silva and Tenreyro 2006). Thus, PQMLE is chosen as our main estimator because it imposes less restriction on the distribution of the data and provides more robust results than alternative models.

Let y be our outcome variable in a given year. We want to explain the expected value of y_i given the distribution of a set of covariates. Because our outcome is a count variable, we assume it follows a Poisson distribution. The conditional density function of y_i is

$$P(y_i \mid \cdot) = \frac{\exp(-\lambda_i)\lambda_i^y}{y!}. \tag{2}$$

The functional form for λ_i can be parameterized as below. For each individual scientist i in year t, his or her expected level of research productivity y_{it} *can* be expressed as

$$\begin{aligned} E\{y_{it} \mid x_{it}, z_i, w_{it}, t, u, \varepsilon_{it}\} \\ = \exp\{x'_{it}\beta + z'_i\delta + w'_{it}\gamma + t + u + \varepsilon_{it}\}, \end{aligned} \tag{3}$$

where x_{it} are our IT variables of interest, z_i is a vector of time-invariant variables controlling for individual characteristics and for doctoral-degree-granting

[12] The 12,000 scientists' names are randomly drawn from UMI's ProQuest Dissertations database. The fields and degree years sampled were chosen in proportions that matched the distribution of Ph.D. fields and graduation years for faculty serving on the Scientific Advisory Board (SAB) of biotechnology companies that made initial public offerings between the years 1970 and 2002. The sampling frame was structured in this way because the initial research project was designed to study university faculty members' commercialization activities. Despite the sampling method, our sample is highly representative of the underlying population of academic life scientists. See Appendix A for more information on how our sample compares to the NSF and Scientists and Engineers Statistical Data System's (SESTAT's) definition of life sciences.

[13] We start our estimation window before the onset of BITNET rather than in 1980, when BITNET was first available. We use this window because our goal is to assess how availability of IT affects a scientist's research patterns, not to study the diffusion of BITNET or DNS per se. In our models, we compare a scientist who has access to some form of IT (e.g., BITNET) with one who does not, either because BITNET has not yet been introduced or because his or her employer has not adopted it.

institutions, w_{it} is a vector of time-varying controls for characteristics of scientists, their collaboration networks, and their employer institutions. The model also includes a series of calendar year dummies (t). Year dummies control for other structural changes in academe that may have occurred during the observation period: these might include a rise in the number of publication outlets as well as a general increase in collaboration networks. In addition, we also include the institutional fixed effects (u) to control for unobserved influences introduced by differences across institutions. The random error is ε_{it}.

5.3. Variables

5.3.1. Dependent Variables. Two dependent variables are examined: (a) *research productivity*, defined as the number of research papers published by a scientist in a given year; and (b) *coauthorship gain*, which measures the increase in the number of new coauthors found in a scientist's publications in a given year. The estimated models also include a large number of individual-level and institutional-level covariates, including the availability of the two IT-related technologies, namely BITNET and DNS. Variable descriptions are provided below and in Table 1.

The first dependent variable, *research productivity*, measures the flow of scientific publications and is updated every year.[14] This "count" measure directly tests the hypothesis set forth in §2: with reduced communication costs and greater access to specialized materials and knowledge, the annual volume of a researcher's output increases.[15] This measure has been frequently used in previous studies of scientific productivity (e.g., Hesse et al. 1993, Cohen 1996, Kaminer and Braunstein 1998, Walsh et al. 2000, Azoulay et al. 2009, Winkler et al. 2010). The other dependent variable examined, *coauthorship gain*, measures the number of new coauthors found in papers published by a scientist in a given year. For example, if A and B appear for the first time in year t as a coauthor to John Doe, Doe's coauthorship gain value is 2 in year t. In the case when A and B have appeared in Doe's papers before t, they are not counted as new coauthors, and Doe's coauthorship gain measure for year t is zero. This count measure tracks changes in a scientist's coauthor network and is used to test the second hypothesis set out above. Inclusion of a new coauthor arguably reflects access to new expertise, equipment, or materials brought along by a new collaborator.[16]

5.3.2. IT Measures. IT availability is coded as the availability of IT at the university where the scientist is employed. Information on if and when the scientist's university adopted BITNET and DNS comes from the Atlas of Cyberspaces and ALLWHOIS (http://www.networksolutions.com/whois/index.jsp) websites, respectively. BITNET is coded 1 for the years when BITNET is available at the scientist's university and 0 otherwise. The same coding scheme is used for DNS. Both *BITNET* and *DNS* variables are lagged by one year in the estimated publication and collaboration models.

As discussed in the literature review, the explicit measure of IT employed here is strongly preferred to inferring IT availability from period effects because the latter may capture confounding factors and does not reflect the fact that diffusion occurs over time. Furthermore, information on the timing of institutional access to IT combined with longitudinal individual-level data permit identification of causal effects. Nevertheless, the IT measures used here are not ideal. The ideal measure would consist of information on when (and if) individual scientists adopted IT, given institutional availability of IT. To the best of our knowledge, no such data are available for a longitudinal sample of scientists. Such a measure, if available, would be preferred because scientists may differ in their speed of adoption within a given institution (Walsh and Bayma 1996).[17] Thus, the relationship between IT availability on a campus and a scientist's productivity estimated here may reflect factors other than the scientist's use of IT. Nonetheless, we believe the reach of this concern to be limited, given that the scientists in our data set are active researchers who presumably would adopt IT innovations offering research advantages sooner than would nonactive researchers.

5.3.3. Scientist's Characteristics. The scientist's characteristics captured in this study include gender, subfield, professional experience, number of jobs held, stock of publications, citation count, coauthorship history, research funding, and Ph.D. training.

[14] As a robustness check, we also examine an author-count-deflated publication measure.

[15] A quality-adjusted measure is another possibility (see Kim et al. 2009). However, in this case, the expected hypotheses are not nearly as well developed. For example, though IT may increase production of knowledge by more efficiently matching expertise and equipment across researchers, it may also increase the "balkanization" of the research community, leading to more insular collaborations (Rosenblat and Mobius 2004, Van Alstyne and Brynjolfsson 2005). One unintended consequence may be a reduction in the quality of research. Our preliminary analysis found no significant relationship between our measures of IT and a quality-adjusted dependent variable.

[16] We use an alternative measure as a robustness check, measuring collaboration as the number of coauthor incidences found in a scientist's publications.

[17] There is also the possibility that those without access to IT on their own campus have access to IT through another organization, such as a research institute. We are unable to control for this in our analysis although we control for whether the individual had access to IT in graduate school.

Table 1 Variable Definition and Sources of Information

Variable name	Description	Source
research productivity	Number of research papers published by a scientist in a given year (publication flow)	Web of Science
coauthorship gain	Increase in the number of new coauthors found in a scientist's publications in a given year; for example, if A and B appear for the first time in year t as a coauthor to John Doe, Doe's coauthorship gain value is 2 in year t; however, if A and B have appeared in Doe's papers before year t, they are not counted as a coauthorship gain for year t	Web of Science
female	1 = Yes; 0 = No	Naming convention
Ph.D. subject field	Field in which a scientist is awarded his Ph.D. degree	UMI ProQuest Dissertation
professional experience	Number of years elapsed from the year scientist receives Ph.D. degree	UMI ProQuest Dissertation
number of jobs	Number of employers for which a scientist has worked between Ph.D. grant year and the current year	Web of Science
publication stock	Number of research papers published by a scientist between Ph.D. grant year and the current year	Web of Science
average citation count	Predicted number of citations received per paper for all papers published by a scientist between his Ph.D. grant year and the current year; approximations are used to construct this variable (details in Appendix B)	Web of Science
past five-year coauthoring ties	Number of coauthorship dyads in papers published by a scientist between $t-5$ and t; for a paper written by a scientist with two coauthors, two coauthorship dyads are counted; we then sum up the dyads in all papers by the scientist during the past five years, regardless of whether the scientist has repeated collaboration relations with certain coauthors	Web of Science
total NIH grant	Total amount (in real U.S. dollars) of extramural grant funding awarded to the scientist by NIH before a given year; amount combines direct and indirect costs	Consolidated Grant/Applicant File (CGAF)
Ph.D. university rank	Ranking category of the university where a scientist's Ph.D. was granted (1–25, 26–50, or outside of top 50)	The Gourman Reports
Ph.D. university BITNET	1 = Ph.D.-granting university has adopted BITNET by time of graduation; 0 = otherwise	Atlas of Cyberspaces
Ph.D. university DNS	1 = Ph.D.-granting university has adopted DNS by time of graduation; 0 = otherwise	ALLWHOIS
employer rank	Ranking category of employer university (1–25, 26–50, or outside of top 50)	The Gourman Reports
employer BITNET	1 = Employer university has adopted BITNET; 0 = otherwise	Atlas of Cyberspaces
employer DNS	1 = Employer university has adopted DNS; 0 = otherwise	ALLWHOIS
number of life-science doctorate awards	Number of doctoral degrees awarded in the life sciences by the employer university in a given year	NSF Survey of Earned Doctorates
federal S&E obligations	Amount of federal obligations to support S&E at the employer university in a given year	Survey of Federal Science and Engineering Support to Universities, Colleges, and Nonprofit Institutions (NSF)
year	Calendar year	

Gender of scientists is primarily determined based on first names.[18] When a first name is androgynous, we searched the Web for the scientist's vita, bio-sketch, or pictures, and code gender accordingly. This strategy permits us to confidently identify gender for 98% of the scientists in our data. All remaining scientists with androgynous first names and no gender-related information from the Web are assumed to be male. Our rationale is that most of the gender-ambiguous names belong to foreign-born scientists of East Asian decent. It is reasonable to assume that the vast majority of these are male given the well-documented gender imbalance in science education in these countries. Such a method for determining gender was previously used by Ding et al. (2006).

Apart from gender, we control for other characteristics about the scientist that may affect her productivity. First, because IT usage can differ substantially across fields (Walsh and Bayma 1996), we collected information on the scientific field in which a scientist was trained and used a series of field dummy

[18] The literature on conventions regarding naming suggests that gender is the primary characteristic individuals seek to convey in the selection of given names (Alford 1988, Lieberson and Bell 1992).

variables to tease out productivity differences across fields.

Second, to control for life-cycle effects (Levin and Stephan 1991), we include a scientist's professional experience and its squared terms. Experience is defined as the number of years lapsed since a scientist obtained her Ph.D. degree.

Third, a scientist who has held jobs at different universities might have broader research networks. Job mobility may also be an indicator of productivity, because research active scientists may have more job offers. Hence we control for the number of jobs a scientist has held.

Fourth, the production of scientific knowledge is a cumulative process. Past research performance may have an impact on current productivity. To control for factors related to past research performance, we include two variables—lagged publication stock and lagged average citation per paper to all published papers by the scientist. Explanation of the construction of these variables is found in Table 1. Details on how the citation count variable is computed can be found in Appendix B.

Fifth, collaboration has been found to have a positive relationship with count of research publications (Lee and Bozeman 2005, Fox and Mohapatra 2007). Hence, we control for a scientist's past research collaboration network. Specifically, we take all of a scientist's published papers during the past five years and add up the total number of coauthoring ties reflected in these papers. Note that this variable does not discount repeated coauthoring ties. If a collaborator has coauthored with a focal scientist twice in the last five years, two coauthoring ties will be counted in this measure. In short, this measure captures the intensity of a scientist's collaboration rather than the breadth of her network.[19]

Sixth, to control for the resources a scientist has available for conducting research, we collected data from the Consolidated Grant/Applicant File (CGAF) from the U.S. National Institutes of Health (NIH). This data set records information about grants awarded to extramural researchers funded by the NIH since 1938; it has been used previously to control for input of scientific research (Azoulay et al. 2010). For each scientist, the variable *total NIH grant* measures the total amount of direct and indirect costs (in real U.S. dollars) of funding she has received from NIH by a given year. This variable is also lagged when entered in the models.

Seventh, to control for the scientist's training and thus implicitly to control for the scientist's knowledge base, we include a measure of the ranking of the scientist's Ph.D-granting institution based on the Gourman Reports, which began ranking graduate schools in 1980. For all periods prior to 1980 we assign universities the 1980 Gourman ranking. We group institutions into three categories—top 25, between 26 and 50, and below 50. We also control for whether the scientist received her Ph.D. from an institution that had adopted BITNET or DNS by the time she graduated. Our rationale is that access to IT while a graduate student can influence productivity at the employer university even if the faculty member takes a position at a university that does not have IT.[20] Moreover, controlling for whether the individual had access to IT as a graduate student distinguishes cohorts that early in their careers had access to IT from cohorts who did not.

5.3.4. Institutional Characteristics. In addition to information on the availability of IT at the university where a scientist works, we also include several other institutional-level variables that proxy other inputs in the knowledge-production function presented in §2.

First is the tier of the current university (Gourman Reports). As in the case of Ph.D-granting institutions, universities are grouped into three categories: top 25, between 26 and 50, and below 50. We check for robustness by defining alternative groupings of universities in the empirical work. Tier is a measure of the quality of colleagues, graduate students, and postdocs with whom the scientist interacts on a regular basis.

We also include the number of doctoral degrees awarded by the scientist's university,[21] as a proxy for the availability of graduate students working in the labs. Further, we control for the amount (in real U.S. dollars) of federal obligations for supporting science and engineering at the employer university. Although this variable does not measure the actual science and engineering research expenditures in a given year, it serves as a crude proxy of the available resources at a scientist's employer university.

Descriptive statistics of the variables are provided in Table 2, and the correlation matrix is in Table 3. We see that the scientists in our sample publish on average 1.62 articles a year and gain 2.44 coauthors a year. Approximately one-fifth are women. Almost one-fifth work at mid-tier institutions; 50% work at lower-tier institutions.

[19] In unreported models, we control for an alternative measure of collaborative network, which is the number of unique coauthors in the past five years instead of number of coauthoring ties (incidences). This measure captures more of the reach of a scientist's network. Our results hold with this alternative measure of past collaboration.

[20] For instance, it could have helped the researcher build strong networks prior to leaving graduate school that she could take with her when she transitioned to a faculty position.

[21] We also experimented with the alternative of including a dummy variable indicating whether a scientist's current university grants doctoral degrees in the life sciences. The result does not differ significantly from those reported in the paper.

Table 2 Descriptive Statistics

Variable name	Mean	Standard deviation	Min	Max
Time-varying ($N = 32{,}398$)				
research productivity (publication flow)	1.621	2.254	0	35
coauthorship gain	2.442	4.248	0	69
professional experience	9.903	5.598	2	26
number of jobs	1.363	0.585	1	5
publication stock	14.05	20.39	0	298
average citation count	12.92	15.49	0	341.7
past five-year coauthoring ties	19.23	29.79	0	688
total NIH grant (in million dollars)	0.197	0.579	0	13.6
employer rank 1–25	0.327	0.469	0	1
employer rank 26–50	0.182	0.386	0	1
employer rank outside 50	0.492	0.500	0	1
number of life-science doctorates	49.19	41.69	0	188
federal S&E obligations (in million dollars)	77.37	102.3	0	793.3
BITNET (1 = yes)	0.517	0.500	0	1
DNS (1 = yes)	0.367	0.482	0	1
year	1983.8	6.057	1969	1993
Time-constant ($N = 3{,}114$)				
female	0.206	0.405	0	1
Ph.D. university rank 1–25	0.377	0.485	0	1
Ph.D. university rank 26–50	0.193	0.394	0	1
Ph.D. university rank outside	0.431	0.495	0	1
Ph.D. university 50 BITNET (1 = yes)	0.176	0.381	0	1
Ph.D. university DNS (1 = yes)	0.116	0.321	0	1

6. Results

6.1. Effect of IT on Research Productivity

The PQML regression results are reported in Table 4(a) (for BITNET) and Table 4(b) (for DNS). The baseline model 1 includes a set of control variables: calendar-year dummies, Ph.D. subject-field dummies, university fixed effects, and individual and institutional factors that can affect productivity.

The baseline results are consistent with major findings of previous studies of scientific productivity (e.g., Levin and Stephan 1991, Xie and Shauman 1998; see Fox 1983 and Long and Fox 1995 for reviews). Research productivity is a concave function of a scientist's professional age and peaks about 18 years after she has received her Ph.D.[22] Women scientists have lower productivity than men. Number of jobs held, publication stock, past coauthoring ties and NIH grants all show a positive and significant association with the current year's publication count. None of the variables relating to institution of employment are significant. This is due in part to the inclusion of university fixed effects in this model, as demonstrated in model 4 (discussion to follow). Rank of the institution of Ph.D. training is not significant and neither is whether one had access to IT during one's Ph.D. program.

We hypothesize in H1 that access to IT increases research productivity. We first test this hypothesis for BITNET and report findings in model 2 of Table 4(a). Holding baseline factors constant, we find that the availability of BITNET at a scientist's institution, measured with a one-year lag, was associated with a 7.6% ($=\exp[0.073]$) increase in publication count.

To test H3A regarding differential IT effects across scientists employed at different tiers of universities, we include an interaction term between BITNET and employer ranking. To be more specific, model 3 includes two employer-rank categories (26–50, and outside top 50, with the rank category 1–25 used as the reference group), the *BITNET* variable, and interaction terms between BITNET and the two employer-rank categories. The coefficients for employer-rank categories describe the effects of being employed at a university of a particular rank category when BITNET was not yet available. From model 3 we see that when BITNET is not yet on a campus, the productivity of scientists employed at mid- and lower-tiered universities does not differ significantly from those employed at top-tier universities. The coefficient for the variable *BITNET* describes the effect of BITNET for those employed at top-tier universities (ranked 1–25) relative to the reference category (top-tier universities that do not have BITNET). The results suggest that the availability of BITNET does not lead to a difference in productivity for those at top-tier institutions. The interaction terms reflect the differential effect of BITNET across tiers. We find no evidence of a differential effect for those employed at mid-tier (26–50) universities. We find that the availability of IT leads to a gain of 18% ($=\exp[0.169]$) in publication counts for those employed at the lowest tier.

Thus far, the models presented include university fixed effects, thereby controlling for unobserved differences across institutions. For comparison purposes, we reestimate model 3 without university fixed effects and report results in model 4. As one would expect, university-ranking effects are larger in magnitude when institutional fixed effects are removed. Scientists working at lower-tier universities (outside of top 50) publish significantly less than the base group of top-tier university scientists. In addition, positive and significant interaction effects with BITNET

[22] Professional age (experience) is highly correlated with a scientist's actual age (Stephan and Levin 1992).

Table 3 Correlation Matrix

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)
(1) *publication flow*	1.000																					
(2) *coauthorship gain*	0.536	1.000																				
(3) *experience*	0.168	0.155	1.000																			
(4) *female*	−0.061	−0.030	−0.077	1.000																		
(5) *number of jobs*	0.137	0.125	0.370	−0.045	1.000																	
(6) *publication stock*	0.652	0.542	0.543	−0.089	0.260	1.000																
(7) *avg. citation count*	0.136	0.107	0.364	−0.053	0.205	0.260	1.000															
(8) *past five-year coauthor ties*	0.733	0.712	0.291	−0.054	0.167	0.789	0.197	1.000														
(9) *total NIH grant*	0.215	0.176	0.396	−0.046	0.195	0.371	0.307	0.272	1.000													
(10) *Ph.D. univ. rank 1–25*	0.030	0.017	0.022	0.040	0.039	0.027	0.114	0.023	0.055	1.000												
(11) *Ph.D. univ. rank* 26–50	−0.005	−0.018	0.027	−0.036	−0.013	0.007	0.024	−0.017	0.032	−0.405	1.000											
(12) *Ph.D. univ. rank outside 50*	−0.026	−0.002	−0.044	−0.010	−0.028	−0.032	−0.133	−0.009	−0.081	−0.658	−0.422	1.000										
(13) *Ph.D. univ. BITNET*	−0.054	−0.004	−0.243	0.089	−0.111	−0.128	−0.084	−0.056	−0.081	0.043	−0.019	−0.026	1.000									
(14) *Ph.D. univ. DNS*	−0.040	0.002	−0.197	0.070	−0.097	−0.099	−0.074	−0.044	−0.067	0.065	−0.018	−0.050	0.698	1.000								
(15) *employer rank 1–25*	0.046	0.050	−0.093	0.033	−0.058	−0.004	0.069	0.049	0.012	0.199	−0.066	−0.143	0.069	0.058	1.000							
(16) *employer rank 26–50*	0.034	0.020	0.011	0.000	0.008	0.048	0.023	0.036	0.016	0.016	0.145	−0.135	−0.002	−0.012	−0.328	1.000						
(17) *employer rank outside 50*	−0.070	−0.062	0.079	−0.031	0.048	−0.034	−0.082	−0.074	−0.024	−0.199	−0.050	0.238	−0.064	−0.046	−0.685	−0.464	1.000					
(18) *number of life-science doctorates*	0.051	0.069	0.017	0.036	−0.029	0.057	0.044	0.077	0.013	0.132	−0.023	−0.112	0.090	0.075	0.556	0.030	−0.545	1.000				
(19) *federal S&E obligations*	0.083	0.109	0.001	0.071	−0.025	0.073	0.149	0.127	0.064	0.147	−0.040	−0.112	0.134	0.108	0.426	0.030	−0.423	0.405	1.000			
(20) *BITNET*	0.113	0.164	0.472	0.043	0.204	0.288	0.223	0.219	0.238	0.036	−0.007	−0.029	0.303	0.222	0.058	−0.003	−0.052	0.131	0.183	1.000		
(21) *DNS*	0.104	0.155	0.416	0.046	0.162	0.268	0.183	0.209	0.230	0.031	−0.005	−0.026	0.345	0.279	0.077	0.017	−0.086	0.203	0.206	0.723	1.000	
(22) *year*	0.125	0.178	0.580	0.043	0.275	0.339	0.263	0.252	0.272	0.011	−0.009	−0.004	0.318	0.256	−0.014	0.007	0.008	0.107	0.148	0.812	0.731	1

Table 4(a) Effect of BITNET on Research Productivity

	Model 1	Model 2	Model 3	Model 4	Model 5
Individual characteristics					
female	−0.139	−0.139	−0.138	−0.145	−0.205
	(0.037)**	(0.037)**	(0.037)**	(0.039)**	(0.055)**
experience	0.053	0.053	0.054	0.060	0.052
	(0.012)**	(0.012)**	(0.012)**	(0.013)**	(0.012)**
experience^2	−0.003	−0.003	−0.003	−0.003	−0.003
	(0.001)**	(0.001)**	(0.001)**	(0.001)**	(0.001)**
number of jobs	0.170	0.169	0.168	0.167	0.170
	(0.018)**	(0.018)**	(0.018)**	(0.020)**	(0.018)**
publication stock$_{t-1}$	0.006	0.006	0.005	0.007	0.005
	(0.001)**	(0.001)**	(0.001)**	(0.001)**	(0.001)**
average citation count$_{t-1}$	0.001	0.001	0.001	0.0001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
past five-year coauthoring ties$_{t-1}$	0.008	0.008	0.008	0.007	0.008
	(0.001)**	(0.001)**	(0.001)**	(0.001)**	(0.001)**
total NIH grant (in million\$)$_{t-1}$	0.091	0.091	0.092	0.103	0.091
	(0.030)**	(0.030)**	(0.030)**	(0.035)**	(0.030)**
Ph.D. university characteristics					
Ph.D. university rank 26–50	−0.006	−0.006	−0.005	0.005	−0.006
	(0.036)	(0.036)	(0.036)	(0.040)	(0.036)
Ph.D. university rank outside 50	−0.025	−0.025	−0.023	−0.022	−0.025
	(0.033)	(0.033)	(0.033)	(0.035)	(0.032)
Ph.D. university BITNET	−0.087	−0.086	−0.078	−0.111	−0.091
	(0.075)	(0.075)	(0.075)	(0.077)	(0.076)
Ph.D. university DNS	0.061	0.063	0.068	0.084	0.061
	(0.085)	(0.085)	(0.085)	(0.087)	(0.085)
Employer university characteristics					
employer rank 26–50	0.065	0.062	0.020	−0.062	0.064
	(0.092)	(0.092)	(0.096)	(0.049)	(0.092)
employer rank outside 50	−0.070	−0.075	−0.155	−0.222	−0.072
	(0.122)	(0.123)	(0.123)	(0.048)**	(0.122)
number of life-science doctorates (in 100)$_{t-1}$	−0.008	−0.007	−0.004	−0.019	−0.007
	(0.049)	(0.049)	(0.047)	(0.038)	(0.049)
federal S&E obligations (in billion\$)$_{t-1}$	−0.072	−0.074	0.036	0.083	−0.078
	(0.242)	(0.241)	(0.244)	(0.129)	(0.241)
BITNET$_{t-1}$		0.073	−0.021	−0.080	0.056
		(0.035)*	(0.042)	(0.049)	(0.036)
employer rank 26–50 × *BITNET*$_{t-1}$			0.094	0.193	
			(0.069)	(0.072)**	
employer rank outside 50 × *BITNET*$_{t-1}$			0.169	0.240	
			(0.055)**	(0.060)**	
female × *BITNET*$_{t-1}$					0.116
					(0.057)*
Scientific field dummies	Yes	Yes	Yes	Yes	Yes
Institutional dummies	Yes	Yes	Yes	No	Yes
Calendar year dummies	Yes	Yes	Yes	Yes	Yes
Log pseudo-likelihood	−53,207.9	−53,201.9	−53,170.4	−54,480.4	−53,191.6
df_m	300	300	304	50	305

Notes. Number of observations = 32,398; number of researchers = 3,114; number of institutions = 314. Robust standard errors are in parentheses, clustered around scientists.

†Significant at 10%; *significant at 5%; **significant at 1%.

are found for both mid- and lower-tier universities, though the size of the BITNET effect is larger for the lower-tier than for the mid-tier universities. Most important, the main message that IT boosts productivity more for nonelite-university scientists remains unchanged regardless of whether or not institutional fixed effects are included.

To test H4A that access to IT has a greater positive effect for women, model 5 includes an interaction term between female and BITNET. We find that women with access to BITNET enjoy a 12.3% [$=\exp(0.116)$]) edge over men who have access to BITNET. We find that BITNET does not lead to a significant increase in productivity for men (reflected in

Table 4(b) Effect of DNS on Research Productivity

	Model 1	Model 2	Model 3	Model 4	Model 5
Individual characteristics					
female	−0.139	−0.140	−0.138	−0.146	−0.183
	(0.037)**	(0.037)**	(0.037)**	(0.039)**	(0.047)**
experience	0.053	0.053	0.054	0.060	0.052
	(0.012)**	(0.012)**	(0.012)**	(0.013)**	(0.012)**
experience^2	−0.003	−0.003	−0.003	−0.003	−0.003
	(0.001)**	(0.001)**	(0.001)**	(0.001)**	(0.001)**
number of jobs	0.170	0.169	0.167	0.167	0.170
	(0.018)**	(0.018)**	(0.018)**	(0.019)**	(0.018)**
publication stock$_{t-1}$	0.006	0.006	0.005	0.007	0.005
	(0.001)**	(0.001)**	(0.001)**	(0.001)**	(0.001)**
average citation count$_{t-1}$	0.001	0.001	0.001	0.0001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
past five-year coauthoring ties$_{t-1}$	0.008	0.008	0.008	0.007	0.008
	(0.001)**	(0.001)**	(0.001)**	(0.001)**	(0.001)**
total NIH grant (in million\$)$_{t-1}$	0.091	0.091	0.092	0.103	0.091
	(0.030)**	(0.030)**	(0.030)**	(0.036)**	(0.030)**
Ph.D. university characteristics					
Ph.D. university rank 26–50	−0.006	−0.006	−0.005	0.003	−0.006
	(0.036)	(0.036)	(0.036)	(0.039)	(0.036)
Ph.D. university rank outside 50	−0.025	−0.025	−0.025	−0.025	−0.025
	(0.033)	(0.033)	(0.032)	(0.035)	(0.033)
Ph.D. university BITNET	−0.087	−0.087	−0.075	−0.108	−0.092
	(0.075)	(0.075)	(0.075)	(0.077)	(0.076)
Ph.D. university DNS	0.061	0.064	0.070	0.083	0.060
	(0.085)	(0.085)	(0.085)	(0.087)	(0.085)
Employer university characteristics					
employer rank 26–50	0.065	0.063	0.022	−0.048	0.065
	(0.092)	(0.092)	(0.094)	(0.044)	(0.092)
employer rank outside 50	−0.070	−0.073	−0.141	−0.186	−0.072
	(0.122)	(0.122)	(0.122)	(0.043)**	(0.122)
number of life-science doctorates (in 100)$_{t-1}$	−0.008	−0.007	0.002	−0.014	−0.007
	(0.049)	(0.049)	(0.046)	(0.038)	(0.049)
federal S&E obligations (in billion\$)$_{t-1}$	−0.072	−0.082	0.019	0.075	−0.086
	(0.242)	(0.243)	(0.248)	(0.130)	(0.242)
BITNET$_{t-1}$		0.073	0.061	0.050	0.074
		(0.035)*	(0.033)†	(0.034)	(0.035)*
DNS$_{t-1}$		0.018	−0.093	−0.145	0.002
		(0.032)	(0.057)	(0.063)*	(0.034)
employer rank 26–50 × *DNS*$_{t-1}$			0.122	0.226	
			(0.075)	(0.079)**	
employer rank outside 50 × *DNS*$_{t-1}$			0.203	0.252	
			(0.064)**	(0.067)**	
female × *DNS*$_{t-1}$					0.108
					(0.054)*
Scientific field fixed effect	Yes	Yes	Yes	Yes	Yes
Institutional fixed effect	Yes	Yes	Yes	No	Yes
Year fixed effect	Yes	Yes	Yes	Yes	Yes
Log pseudo-likelihood	−53,207.9	−53,201.5	−53,159.0	−54,474.2	−53,192.7
df_m	300	304	303	51	303

Notes. Number of observations = 32,398; number of researchers = 3,114; number of institutions = 314. Robust standard errors are in parentheses, clustered around scientists.

†Significant at 10%; *significant at 5%; **significant at 1%.

the nonsignificant coefficient for the BITNET variable in the model).

Table 4(b) reports the same set of tests using DNS as the IT variable.[23] The main effect of DNS is positive but not significant (model 2). This may be because DNS was a "successor" technology to BITNET and most of the productivity-enhancing effect was realized through BITNET rather than through DNS. The insignificance of the main effect of DNS does not preclude differential effects of this technology across subgroups such as we have hypothesized. Indeed, this is exactly what we observe in the next three models. Models 3 and 4 test the effects of DNS across different tiers of universities, with and without institutional fixed effects, respectively. We find the same pattern of

[23] The DNS model also controls for availability of BITNET, since DNS was a "successor" technology.

Table 5(a) Effect of BITNET on Collaboration Network

	Model 1	Model 2	Model 3	Model 4	Model 5
Individual characteristics					
female	−0.084	−0.084	−0.082	−0.085	−0.157
	(0.040)*	(0.040)*	(0.040)*	(0.044)†	(0.059)**
experience	0.018	0.019	0.019	0.023	0.018
	(0.012)	(0.012)	(0.012)	(0.013)†	(0.012)
experience^2	−0.002	−0.002	−0.002	−0.002	−0.002
	(0.001)**	(0.001)**	(0.001)**	(0.001)**	(0.001)**
number of jobs	0.194	0.194	0.193	0.193	0.194
	(0.020)**	(0.020)**	(0.020)**	(0.023)**	(0.020)**
publication stock$_{t-1}$	0.003	0.003	0.003	0.004	0.003
	(0.001)**	(0.001)**	(0.001)**	(0.001)**	(0.001)**
average citation count$_{t-1}$	−0.002	−0.002	−0.002	−0.002	−0.002
	(0.001)*	(0.001)*	(0.001)*	(0.001)*	(0.001)*
past five-year coauthoring ties$_{t-1}$	0.009	0.009	0.009	0.008	0.009
	(0.001)**	(0.001)**	(0.001)**	(0.001)**	0.001)**
total NIH grant (in million$)$_{t-1}$	0.065	0.066	0.067	0.080	0.066
	(0.027)*	(0.027)*	(0.027)*	(0.033)*	(0.027)*
Ph.D. university characteristics					
Ph.D. university rank 26–50	−0.006	−0.006	−0.006	−0.001	−0.006
	(0.039)	(0.039)	(0.039)	(0.042)	(0.039)
Ph.D. university rank outside 50	0.029	0.030	0.031	0.029	0.030
	(0.037)	(0.037)	(0.037)	(0.040)	0.037)
Ph.D. university BITNET	−0.102	−0.101	−0.094	−0.115	−0.106
	(0.069)	(0.069)	(0.069)	(0.071)	(0.070)
Ph.D. university DNS	0.066	0.069	0.072	0.083	0.067
	(0.080)	(0.080)	(0.080)	(0.083)	(0.080)
Employer university characteristics					
employer rank 26–50	0.001	−0.003	−0.002	−0.059	−0.000
	(0.108)	(0.108)	(0.111)	(0.056)	(0.108)
employer rank outside 50	−0.051	−0.058	−0.131	−0.236	−0.054
	(0.140)	(0.140)	(0.139)	(0.053)**	(0.140)
number of life-science doctorates (in 100)$_{t-1}$	−0.106	−0.104	−0.107	−0.012	−0.104
	(0.060)†	(0.060)†	(0.058)†	(0.040)	(0.060)†
federal S&E obligations (in billion$)$_{t-1}$	0.069	0.070	0.187	0.288	0.065
	(0.249)	(0.249)	(0.252)	(0.159)†	(0.249)
BITNET$_{t-1}$		0.118	0.048	−0.021	0.100
		(0.042)**	(0.053)	(0.057)	(0.043)*
employer rank 26–50 × *BITNET*$_{t-1}$			0.003	0.124	
			(0.076)	(0.078)	
employer rank outside 50 × *BITNET*$_{t-1}$			0.152	0.246	
			(0.058)**	(0.063)**	
female × *BITNET*$_{t-1}$					0.116
					(0.057)*
Scientific field dummies	Yes	Yes	Yes	Yes	Yes
Institutional dummies	Yes	Yes	Yes	No	Yes
Calendar year dummies	Yes	Yes	Yes	Yes	Yes
Log pseudo-likelihood	−80,936.8	−80,913.9	−80,868.4	−83,172.0	−80,898.1
df_m	316	316	317	50	318

Notes. Number of observations = 32,398; number of researchers = 3,114; number of institutions = 314. Robust standard errors are in parentheses, clustered around scientists.

†Significant at 10%; *significant at 5%; **significant at 1%.

results as with BITNET: nonelite-university scientists benefit more from DNS than do scientists at elite universities. Model 5 tests the effect of DNS for women. As in the case of BITNET, we also find stronger effect of DNS for women than for men.

6.2. Effect of IT on Research Collaboration

Tables 5(a) and 5(b) report estimations concerning the effect of IT on scientists' collaboration networks. To test H2, we examine how the availability of IT changes the number of new coauthors in a scientist's

Table 5(b) Effect of DNS on Collaboration Network

	Model 1	Model 2	Model 3	Model 4	Model 5
Individual characteristics					
female	−0.084	−0.084	−0.083	−0.086	−0.138
	(0.040)*	(0.040)*	(0.040)*	(0.044)†	(0.051)**
experience	0.018	0.019	0.020	0.023	0.018
	(0.012)	(0.012)	(0.012)	(0.013)†	(0.012)
experience^2	−0.002	−0.002	−0.002	−0.002	−0.002
	(0.001)**	(0.0005)**	(0.001)**	(0.001)**	(0.000)**
number of jobs	0.194	0.194	0.192	0.193	0.194
	(0.020)**	(0.020)**	(0.020)**	(0.023)**	(0.020)**
publication stock$_{t-1}$	0.003	0.003	0.003	0.004	0.003
	(0.001)**	(0.001)**	(0.001)**	(0.001)**	(0.001)**
average citation count$_{t-1}$	−0.002	−0.002	−0.002	−0.002	−0.002
	(0.001)*	(0.001)*	(0.001)*	(0.001)*	(0.001)*
past five-year coauthoring ties$_{t-1}$	0.009	0.009	0.009	0.008	0.009
	(0.001)**	(0.001)**	(0.001)**	(0.001)**	(0.001)**
total NIH grant (in million$)$_{t-1}$	0.065	0.066	0.067	0.081	0.066
	(0.027)*	(0.027)*	(0.027)*	(0.033)*	(0.027)*
Ph.D. university characteristics					
Ph.D. university rank 26–50	−0.006	−0.006	−0.006	−0.002	−0.005
	(0.039)	(0.039)	(0.039)	(0.042)	(0.039)
Ph.D. university rank outside 50	0.029	0.030	0.031	0.026	0.030
	(0.037)	(0.037)	(0.037)	(0.040)	(0.037)
Ph.D. university BITNET	−0.102	−0.101	−0.092	−0.112	−0.106
	(0.069)	(0.069)	(0.069)	(0.071)	(0.070)
Ph.D. university DNS	0.066	0.069	0.073	0.082	0.065
	(0.080)	(0.080)	(0.080)	(0.083)	(0.080)
Employer university characteristics					
employer rank 26–50	0.001	−0.003	−0.018	−0.052	−0.001
	(0.108)	(0.108)	(0.109)	(0.049)	(0.108)
employer rank outside 50	−0.051	−0.058	−0.124	−0.194	−0.055
	(0.140)	(0.140)	(0.140)	(0.048)**	(0.140)
number of life-science doctorates (in 100)$_{t-1}$	−0.106	−0.104	−0.099	−0.005	−0.103
	(0.060)†	(0.060)†	(0.058)†	(0.040)	(0.060)†
federal S&E obligations (in billion$)$_{t-1}$	0.069	0.069	0.175	0.282	0.064
	(0.249)	(0.252)	(0.255)	(0.160)†	(0.251)
BITNET$_{t-1}$		0.118	0.110	0.103	0.120
		(0.042)**	(0.041)**	(0.041)*	(0.042)**
DNS$_{t-1}$		0.002	−0.083	−0.163	−0.017
		(0.038)	(0.058)	(0.065)*	(0.039)
employer rank 26–50 × *DNS*$_{t-1}$			0.039	0.146	
			(0.079)	(0.080)†	
employer rank outside 50 × *DNS*$_{t-1}$			0.176	0.242	
			(0.061)**	(0.064)**	
female × *DNS*$_{t-1}$					0.117
					(0.054)*
Scientific field fixed effect	Yes	Yes	Yes	Yes	Yes
Institutional fixed effect	Yes	Yes	Yes	No	Yes
Year fixed effect	Yes	Yes	Yes	Yes	Yes
Log pseudo-likelihood	−80,936.8	−80,913.9	−80,860.5	−83,174.3	−80,896.8
df_m	316	316	318	51	318

Notes. Number of observations = 32,398; number of researchers = 3,114; number of institutions = 314. Robust standard errors are in parentheses, clustered around scientists.
†Significant at 10%; *significant at 5%; **significant at 1%.

collaboration network. We first focus on BITNET in Table 5(a). The baseline model shows that women are less likely to gain a coauthor, as are more experienced scientists (the term on experience-squared is negative and significant). An increase in coauthors is related to the number of jobs one has held, publication stock, and past coauthoring ties. Consistent with the finding of Bozeman and Corley (2004), collaboration is positively related to the presence of research funding.

Model 2 indicates that, holding constant the control variables, BITNET leads to a 12.5% (= exp[0.118]) increase in the number of new coauthors. There are also substantial and significant differences in the effect of BITNET on collaboration across subsets of the scientists in our sample. First, a positive and significant difference in the effect of BITNET for lower-tier-university scientists is found in model 3, which includes institutional fixed effects in the estimation. The effect is larger in model 4 (which does not include institutional fixed effects) and shows that scientists employed at lower-tier (outside top 50) universities with access to BITNET have 27.9% (= exp[0.246]) more new coauthors than do those working at elite universities with access to BITNET. Second, as in the case of research productivity, the effect of BITNET is significantly higher for women than for men (model 5 of Table 5(a)). Women gain 12.3% (= exp[0.116]) more coauthors than men do when their university has access to BITNET in the previous year.[24]

Table 5(b) reports estimations using DNS as the IT variable. We find no significant coauthorship gain associated with the availability of DNS (see model 2 of Table 5(b)). Again, the lack of a DNS effect may be because DNS was a successor technology. However, results in models 3 and 4 show that lower-tier-university (outside top 50) scientists gain significantly more coauthors from DNS than their counterparts at higher-tier universities gain from DNS. With regard to gender, access to DNS expands the collaboration network for women considerably more than it does for men.

6.3. Robustness Checks

We conducted several alternative estimations to test the robustness of our results. Table 6 reports these for models estimated using BITNET.[25] For ease of comparison we also include results from the corresponding models previously reported in the main tables (Tables 4(a) and 5(a)) at the top of Table 6.

In panel A, we estimate research productivity models using an author-count-deflated publication measure of productivity. This measure differs from the count measure used in the main estimations because it discounts publications that have multiple authors on the production team, implicitly assuming that single-authored publications are worth more (in terms of scientific output) than multiauthored publications. Though we believe this to be inferior to the publication count measure we currently use,[26] it is of interest to know how IT affects a scientist's *net* publication count after accounting for coauthors. The results show that the direct impact of BITNET on this publication measure is not significant (column (1) of panel A). This suggests that the observed gain in research productivity in our main model comes primarily through collaboration. Nevertheless, the interaction effects in the next two columns indicate that the BITNET effect varies among scientists. Scientists at mid- to lower-tier universities and female scientists gained more than top-tier and male scientists in terms of *net* publication count when BITNET became available on their campuses.

Panel B tests the IT-collaboration hypothesis with an alternative measure of scientists' collaboration networks. Rather than using a measure of collaboration that tracks the addition of new coauthors and hence captures the breadth of a scientist's collaboration network, we use an alternative approach that measures the number of coauthoring ties (or incidences) found in a scientist's published papers in a given year (i.e., tie flow). Because this measure counts each coauthoring tie reflected in the papers without eliminating repeated collaborations with the same persons, it might be argued that it does a better job of capturing the intensity of collaboration than our original measure does. We find that estimations with this collaboration variable yield results similar to those reported in our main models, except that the gender-IT interaction effect becomes slightly weaker, falling just below the 95% confidence level.

In panel C, we experiment with an alternative specification of professional experience. Specifically, instead of using a linear and a quadratic form of professional experience, we use 26 annual professional experience dummies (26 years is the maximum number of professional years of experience in our data). These dummies offer a more flexible way of specifying professional-experience effect. The results remain similar to those in our main models.

In panels D and E, we use alternative groupings of the ranking of the employer university. Panel C uses ranking groups 1–20, 21–50, and outside 50. Panel D uses ranking groups 1–10, 11–50, and outside 50. The magnitudes of ranking-related effects change somewhat, particularly with the ranking specifications in panel D. However, the conclusions remain the same: women and nonelite scientists benefit more from BITNET.

[24] Although the coefficients and standard errors of the female–BITNET interactions appear the same (model 5, Tables 4(a) and 5(b)), this is a result of rounding. Also, note that the main effects of IT (captured by the coefficients for BITNET) differ.

[25] Results for DNS-related models differ minimally from those for BITNET. These results are available upon request.

[26] The assumption that single-authored publications are worth more than multiauthored publications could be problematic as a priori there is no reason to expect that single-authored papers are more valuable than multiauthored ones, or that each author's time input into the production of a paper is strictly proportional to the total number of coauthors on a team.

Table 6 Robustness Test Using Alternative Specifications

	Research productivity			Collaboration network		
Correspond to main T(able)–M(odel)	T4a-M2	T4a-M4	T4a-M5	T5a-M2	T5a-M4	T5a-M5
Baseline results reported in Tables 4(a) and 5(a)						
$BITNET_{t-1}$	0.073	−0.080	0.056	0.118	−0.021	0.100
	(0.035)*	(0.049)	(0.036)	(0.042)**	(0.057)	(0.043)*
employer rank 26–50 × $BITNET_{t-1}$		0.193			0.124	
		(0.072)**			(0.078)	
employer rank 26–50 × $BITNET_{t-1}$		0.240			0.246	
		(0.060)**			(0.063)**	
female × $BITNET_{t-1}$			0.116			0.116
			(0.057)*			(0.057)*
Panel A: Using author-count-deflated publication flow as measure for research productivity						
$BITNET_{t-1}$	0.035	−0.099	0.013			
	(0.039)	(0.057)†	(0.040)			
employer rank 26–50 × $BITNET_{t-1}$		0.171				
		(0.078)*				
employer rank 26–50 × $BITNET_{t-1}$		0.215				
		(0.067)**				
female × $BITNET_{t-1}$			0.155			
			(0.066)*			
Panel B: Using coauthoring tie (incidence) flow as the measure for collaboration network						
$BITNET_{t-1}$				0.013	−0.040	0.114
				(0.037)**	(0.055)	(0.037)*
employer rank 26–50 × $BITNET_{t-1}$					0.178	
					(0.078)*	
employer rank 26–50 × $BITNET_{t-1}$					0.241	
					(0.062)**	
female × $BITNET_{t-1}$						0.110
						(0.057)†
Panel C: Using annual professional age (experience) dummies to replace linear and squared terms of experience						
$BITNET_{t-1}$	0.072	−0.081	0.055	0.120	−0.020	0.101
	(0.035)*	(0.049)†	(0.036)	(0.042)**	(0.057)	(0.043)*
employer rank 26–50 × $BITNET_{t-1}$		0.193			0.123	
		(0.072)**			(0.078)	
employer rank 26–50 × $BITNET_{t-1}$		0.242			0.246	
		(0.060)**			(0.063)**	
female × $BITNET_{t-1}$			0.116			0.120
			(0.056)*			(0.057)*
Panel D: Using alternative specification of university ranking—Ranking grouped as 1–20, 21–50, and outside 50						
$BITNET_{t-1}$	0.072	−0.092	0.054	0.117	−0.031	0.099
	(0.035)*	(0.053)†	(0.036)	(0.042)**	(0.061)	(0.043)*
employer rank 21-50 × $BITNET_{t-1}$		0.186			0.126	
		(0.070)**			(0.076)†	
employer rank outside 50 × $BITNET_{t-1}$		0.253			0.257	
		(0.064)**			(0.068)**	
female × $BITNET_{t-1}$			0.117			0.117
			(0.056)*			(0.057)*
Panel E: Using alternative specification of university ranking—Ranking grouped as 1–10, 11–50, and outside 50						
$BITNET_{t-1}$	0.072	−0.193	0.055	0.118	−0.112	0.100
	(0.035)*	(0.082)*	(0.036)	(0.042)**	(0.093)	(0.043)*
employer rank 11–50 × $BITNET_{t-1}$		0.278			0.212	
		(0.107)**			(0.114)†	
employer rank outside 50 × $BITNET_{t-1}$		0.360			0.345	
		(0.100)**			(0.105)**	
female × $BITNET_{t-1}$			0.117			0.117
			(0.057)*			(0.058)*

Table 6 (Continued)

	Research productivity			Collaboration network		
Correspond to main T(able)–M(odel)	T4a-M2	T4a-M4	T4a-M5	T5a-M2	T5a-M4	T5a-M5
Panel F: Using "total number of BITNET adopters by year" to replace calendar year dummies						
$BITNET_{t-1}$	0.084	−0.062	0.067	0.110	−0.019	0.093
	(0.033)*	(0.049)	(0.033)*	(0.040)**	(0.059)	(0.041)*
employer rank 26–50 × $BITNET_{t-1}$		0.187			0.117	
		(0.072)**			(0.079)	
employer rank outside 50 × $BITNET_{t-1}$		0.236			0.246	
		(0.060)**			(0.064)**	
female × $BITNET_{t-1}$			0.114			0.112
			(0.057)*			(0.058)†

†Significant at 10%; *significant at 5%; **significant at 1%.

In panel F, we replace calendar-year dummies with the total number of institutions that have adopted BITNET. The rationale is that the effect of BITNET on collaboration and productivity is realized only when other scientists' universities have adopted it, because it takes two scientists on different ends of the BITNET terminals to collaborate. Hence, we include the number of BITNET adopters by a given year as an indication of the extent of the diffusion of BITNET in academe. This specification does not change the findings of our main models.

7. Conclusion and Discussion

The Internet and other advancements in IT are changing the practice of science. Yet our knowledge concerning how advancements in IT have affected research productivity over time is limited. In large part this is because of the absence of longitudinal data linking information on the availability of IT technology to the productivity of scientists. Here we remedy this situation by creating a database that combines information on the institutional diffusion of BITNET and the DNS with career history data on the publishing patterns of research-active academic life scientists. The data also contain information on the scientists such as gender, professional age, quality of doctoral program, field of expertise, collaboration network, and NIH funding. In addition, characteristics of the employing institution are controlled for in the analysis.

Two dimensions of scientific research are measured: counts of publications and increase in coauthorship. We test whether the adoption of IT by an institution enhanced the research of individual scientists at that institution, and whether the enhancement effect is stronger for two specific subgroups of the scientific labor force: faculty at nonelite institutions and female faculty members.

We find some support for H1 that access to IT increases research productivity. We also find support for H2 that access to IT enhances collaboration. The latter finding is consistent with the frequent inference that IT has been a major contributing factor to the increase in the number of coauthors in science observed since the 1980s.

Our findings also support the hypotheses that the availability of IT has differential effects on productivity depending on a scientist's individual characteristics and position in academe. Specifically, women scientists benefit more than their male colleagues in terms of overall output and an increase in new coauthors. This is consistent with the idea that IT is especially beneficial to individuals who face greater mobility constraints. We also find that the tier of the research organization matters. The availability of IT has a greater effect on the productivity of scientists at nonelite institutions than it does for scientists at elite institutions. The finding is consistent with the idea that faculty at nonelite institutions have relatively more to gain, having fewer in-house colleagues and resources. The gender and research tier results suggest that IT has been an equalizing force, at least in terms of the number of publications and gain in coauthorship, enabling scientists outside the inner circle to participate more fully.

Our findings have implications for the management of scientific research, suggesting that innovations in IT technology contribute to scientific productivity. This is important given that the IT revolution continues. Whereas early investments in IT focused on connectivity, later investments have focused on enhancing the quality and speed of connectivity and the resources available online. For example, NIH, along with several other government agencies, has invested millions in building the Protein Data Bank, which is available online;[27] Google has set the goal of digitizing a number of major libraries; and JSTOR has

[27] The Protein Data Bank website (http://www.pdb.org) is accessed by about 140,000 unique visitors per month from nearly 140 different countries. Approximately 500 gigabytes of data are transferred each month (http://www.rcsb.org/pdb/static.do?p=general_information/about_pdb/index.html).

made a wide array of journals available to institutions that might otherwise not be able to afford a subscription. Moreover, in recent years, equipment, such as synchrotrons and telescopes, has become available for online use. Our research suggests that such investments can lead to an increase in productivity.

Our findings also provide support for the concept that technology can be especially enabling for those situated at the margin. This finding suggests that administrators at lower-tier universities can use IT as one tool for catching up with more elite universities. Our findings also suggest that universities committed to increasing equality among their labor force can use IT as an instrument in narrowing the gap among their scientists, at least between men and women.

Our democratizing findings are consistent with research on two policy innovations that have increased accessibility to research material and have had a similar democratizing effect on the practice of science. Murray et al. (2008), for instance, studied the impact of two Memoranda of Understanding (MoU) between DuPont and the National Institutes of Health that removed many of the restrictions related to working with certain genetically engineered mice. They found post-MoU citations to the original mouse articles to grow at a faster rate from institutions that had previously not done research with the mice than from institutions that had previously done research using the mice. The logic for their finding is that prior to the MoUs, accessibility to mice was considerably more restricted by intellectual property protection.[28] As a second example, Furman and Stern (2009) studied the effect of biological research centers (BRCs) by examining citations in articles written post-deposit to articles associated with materials that had been exogenously shifted to a research center. Consistent with a democratizing effect, they found the rate of citations from new institutions, new journals, and new countries to increase post deposit. They also found that researchers at institutions outside the top-50 U.S. research universities benefited more than those at the top 50 in terms of a post-deposit citation boost to papers that used materials that had subsequently been transferred to a BRC. The policy implication of this research on mice and BRCs, as well as of our own research on IT, is clear: innovations that promote accessibility level the playing field and broaden the base of individuals doing science.

We would be remiss if we did not point out several limitations of our study that we hope future research can address. First, although tracking institutional adoption of BITNET and DNS is an improvement over previous studies, allowing one to date when an individual was able to first access IT, we cannot determine whether and when an individual actually used the technology. As a result, our estimates may be biased, although the direction of bias is unclear without empirical investigation of scientists' IT adoption behavior. We attempted to address the question of actual use by determining whether and when the scientists in our sample used an e-mail or other electronic address. But because the inclusion of an e-mail or other electronic address only became widely accepted in the scientific community in the mid-to-late 1990s (outside our period of analysis), this approach is not viable.[29] We hope future research can address the problem with data that captures IT use at the individual level.

Second, the magnitude and nature of IT effects may differ somewhat for other technologies that have come online, such as those that enhance access to journals, materials, and equipment. It is even possible that certain technologies such as e-mail may create "information overload" and reduce productivity, as researchers spend increasing amounts of what would have been productive time managing their virtual mail. The effect of more contemporary IT and how it varies across subsets of the scientific labor force is a topic we hope to investigate in future research.

Third, the "outer circle" is fairly narrowly defined in our research, excluding in particular underrepresented minorities in science and engineering. This exclusion was data driven: we do not know minority status and, given the underlying distribution of the life-science workforce, we would not have had a sufficient number of observations for analysis if we did know the minority status. Nevertheless, whether IT provides differential advantage to underrepresented minorities is an important question that deserves attention.

Acknowledgments
Ding acknowledges support from the Haas School of Business, Lester Center for Entrepreneurship and Innovation, and the Ewing Marion Kauffman Foundation, Kansas City, Missouri. Levin, Stephan, and Winkler acknowledge support from the Andrew W. Mellon Foundation, New York.

[28] Researchers at institutions where a colleague had either engineered a mouse or accessed a mouse were likely to share the benefits whereas researchers at institutions that did not have a mouse found access more difficult. Furthermore, agreements made prior to the MoU allowed follow-on research for all faculty at the institution.

[29] We examined the inclusion of an electronic address in seven scientific journals relevant to the life sciences: *Cell, Science, Journal of Biological Chemistry, Virology, Biochemistry, Biochemical and Biophysical Research Communications*, *Brain Research*. We search for the terms "BITNET," "Email," "E-mail," or "electronic Mail" (case insensitive). We find 119 mentions of such terms during the period 1982–1990; only 9 of these were part of an electronic address. In 1995 we find 1,495 instances of an electronic address; in 2000 we find 12,274 instances.

Stephan acknowledges support from the International Centre for Economic Research, Turin, Italy. The authors thank Kelly Wilken and Erin Coffman for data assistance. They also thank Fiona Murray, Bill Amis, and three anonymous referees for helpful comments.

Appendix A. Sample Representativeness

We randomly drew scientists' names from UMI's ProQuest Dissertations database, which includes the names, fields, and degree-granting institutions for almost all doctoral degree recipients from accredited U.S. universities. The field and degree years sampled were chosen in proportions that matched the distribution of Ph.D. fields and graduation years for faculty serving on the Scientific Advisory Board (SAB) of biotechnology companies that made initial public offerings between the years 1970 and 2002. The sampling frame was structured in this way because the initial research project was designed to study university faculty members' engagement in the commercialization of academic science.

Table A.1 reports, in order of their representation in the sample, the 15 scientific disciplines that appear most frequently in our data. To determine the degree to which our sample reflects the underlying population of life scientists, we compare these fields to NSF data (see column (4)). We first note that with the exception of organic chemistry and psychobiology, all fields are considered to be in the life, or related sciences, according to NSF's standardized codes used in SESTAT (see http://sestat.nsf.gov/docs/educode2.html for detailed information of the fields in this group). We also compare the distribution of degrees in our sample to the classification and distribution of degrees awarded at U.S. universities between 1965 and 1990 as measured in the Survey of Earned Doctorates (SED). See column (5). We find that, with the exception of organic chemistry, psychobiology, and health sciences/pharmacy, our fields are classified by the SED as part of the life sciences. Column (6) of the table reports the ranking of a field in terms of degrees awarded during the period as reported by the SED. We find considerable, although not complete overlap, between our top fields and SED's top fields. For example, biochemistry contributes the largest number of cases to our sample (23%), and it is also the field in the life sciences with the largest number of doctoral awards in the SED data from 1965 to 1990. The second largest group of doctoral awards in the SED data is microbiology, which ranks third in our data. We conclude that our sample does not differ markedly from the underlying population of life scientists working in academe. To the extent that there is a bias, it is toward fields that are at the forefront of technological developments. We see this as an advantage in our study because scientists in such fields tend to have more extensive information about new technologies (including innovations in IT) and thus may be more disposed to put them to use.

Appendix B. Computation of Average Citation Measure

Following Stuart and Ding (2006), we use the average citation count to control for professional recognitions received by a scientist. Average citation is measured by predicted

Table A.1 Top 15 Scientific Disciplines in the Sample

(1) UMI Ph.D. subject code	(2) UMI subject description	(3) Frequency and share in sample	(4) Classified as "life and related sciences" by NSF in SESTAT[a]	(5) Classified as "life sciences" in SED[b]	(6) Rank in SED based on representation
487; 303	Biochemistry	711 (22.8%)	Yes	Yes	1
306	Biology, General	471 (15.1%)	Yes	Yes	7
410	Biology, Microbiology	379 (12.2%)	Yes	Yes	2
419	Health Sciences, Pharmacology	193 (6.2%)	Yes	Yes	5
786	Biophysics, General	180 (5.8%)	Yes	Yes	16
490	Chemistry, Organic	178 (5.7%)	No	No	—
369	Biology, Genetics	151 (4.8%)	Yes	Yes	15
433	Biology, Animal Physiology	148 (4.8%)	Yes	Yes	3
982	Health Sciences, Immunology	112 (3.6%)	Yes	Yes	18
307	Biology, Molecular	48 (1.5%)	Yes	Yes	6
287	Biology, Anatomy	47 (1.5%)	Yes	Yes	14
301	Bacteriology	45 (1.4%)	Yes	Yes	42
571	Health Sciences, Pathology	42 (1.3%)	Yes	Yes	24
349	Psychology, Psychobiology	31 (1.0%)	No	No	—
572	Health Sciences, Pharmacy	30 (1.0%)	Yes	No	—

[a] *Source*. "Education Codes and Groups," http://sestat.nsf.gov/docs/educode2.html; NSF codes are more broadly defined, and some of the subfields are grouped into the "other" category, e.g., bacteriology (301), anatomy (287) health sciences–immunology (982), and health sciences–pathology (571).

[b] *Source*. Statistical tables based on NSF's Survey of Earned Doctorates (1965–1990) reported in *Science and Engineering Doctorates: 1960–1986* (NSF 88-303) and *Selected Data on Science and Engineering Doctorate Awards: 1994* (NSF 95-337). There is no "health sciences" category in SED, but some of the fields listed under "health sciences" in our (UMI ProQuest) sample do correspond to SED's subfields under "biological sciences." For example, health sciences–pharmacology (419) in our data corresponds to SED's human/animal pharmacology; health sciences–immunology (982) corresponds to SED's biological immunology; health sciences–pathology (571) corresponds to SED's human/animal pathology.

number of citations received per paper for the papers published by a scientist up through a given year. The total citation count for each published article at the time we assembled our database (2002) was obtained from the Web of Science. Because we wish to estimate a scientist's annually updated, cumulative citation count, we distribute each paper's total citation count as of 2002 back through time, assuming that citations arrive according to an exponential distribution with hazard rate equal to 0.1. Our logic is based on the bibliometric literature (for example, Redner 1998), showing that citations follow an exponential distribution, and we find this to be true for the typical paper in our study. We identified the specific parameter, 0.1, by manually coding 50 randomly selected papers in each of three publication years: 1970, 1980, and 1990, and then choosing the parameter that yielded the best fit to the actual time path of citations to these randomly chosen papers. The predicted, cumulative citation count is then divided by the publication stock to obtain the average citation count per paper.

References

Adams, J. 1990. Fundamental stocks of knowledge and productivity growth. *J. Political Econom.* **98**(4) 673–702.

Adams, J. D., G. C. Black, J. R. Clemmons, P. E. Stephan. 2005. Scientific teams and institutional collaborations: Evidence from U.S. universities, 1981–1999. *Res. Policy* **34**(3) 259–285.

Agrawal, A., A. Goldfarb. 2008. Restructuring research: Communication costs and the democratization of university innovation. *Amer. Econom. Rev.* **98**(4) 1578–1590.

Alford, R. D. 1988. *Naming and Identity*. HRAF Press, New Haven, CT.

Anft, M. 2008. Of mice and medicine. *Johns Hopkins Magazine* **60**(4), http://www.jhu.edu/jhumag/0908web/mice.html.

Azoulay, P., W. W. Ding, T. E. Stuart. 2009. The impact of academic patenting on the rate, quality, and direction of (public) research output. *J. Indust. Econom.* **57**(4) 637–676.

Azoulay, P., J. S. Zivin, J. Wang. 2010. Superstar extinction. *Quart. J. Econom.* **125**(2) 549–589.

Barley, S. R. 1986. Technology as an occasion for structuring: Evidence from observations of CT scanners and the social order of radiology departments. *Admin. Sci. Quart.* **31**(1) 78–108.

Bozeman, B., E. Corley. 2004. Scientists' collaboration strategies: Implications for scientific and technical human capital. *Res. Policy* **33**(4) 599–616.

Brynjolfsson, E. 1993. The productivity paradox of information technology. *Comm. ACM* **36**(12) 67–77.

Burkhardt, M. E., D. J. Brass. 1990. Changing patterns or patterns of change: The effects of a change in technology on social network structure and power. *Admin. Sci. Quart.* **35**(1) 104–127.

Butler, D. M., R. J. Butler, J. T. Rich. 2008. The equalizing effect of the Internet on access to research expertise in political science and economics. *PS: Political Sci. Politics* **41**(3) 579–584.

Cohen, J. 1996. Computer mediated communication and publication productivity among faculty. *Internet Res.* **6**(2–3) 41–63.

Corporation for Research & Educational Networking (CREN). 1997. CREN history and future. June 1997. http://www.cren.net/cren/cren-hist-fut.html.

DeLone, W. H., E. R. McLean. 2003. The Delone and Mclean model of information systems success: A ten-year update. *J. Management Inform. Systems* **19**(4) 9–30.

Dewan, S., K. L. Kraemer. 2000. Information technology and productivity: Evidence from country-level data. *Management Sci.* **46**(4) 548–562.

Ding, W. W., F. Murray, T. E. Stuart. 2006. Gender differences in patenting in the academic life sciences. *Science* **313**(5787) 665–667.

Fox, M. F. 1983. Publication productivity among scientists: A critical review. *Soc. Stud. Sci.* **13**(2) 285–305.

Fox, M. F., S. Mohapatra. 2007. Social-organizational characteristics of work and publication productivity among academic scientists in doctoral-granting departments. *J. Higher Ed.* **78**(5) 542–571.

Furman, J. L., S. Stern. 2009. Climbing atop the shoulders of giants: The impact of institutions on cumulative research. Working paper, Boston University, Boston.

Ginther, D. K., S. Kahn. 2009. Does science promote women? Evidence from academia 1973–2001. R. B. Freeman, D. L. Goroff, eds. *Science and Engineering Careers in the United States*. University of Chicago Press, Chicago, 163–194.

Gourman, J. 1980–1997. *The Gourman Report: A Rating of Graduate and Professional Programs in American and International Universities*. 8 editions. Random House, New York.

Gourieroux, C., A. Monfort, A. Trognon. 1984. Pseudo maximum-likelihood methods—Applications to Poisson models. *Econometrica* **52**(3) 701–720.

Hamermesh, D. S., S. M. Oster. 2002. Tools or toys? The impact of high technology on scholarly productivity. *Econom. Inquiry* **40**(4) 539–555.

Hesse, B. W., L. S. Sproull, S. B. Kiesler, J. P. Walsh. 1993. Returns to science—Computer-networks in oceanography. *Comm. ACM* **36**(8) 90–101.

Humphrys, M. n.d. The Internet in the 1980s. http://computing.dcu.ie/~humphrys/net.80s.html.

Jones, B. F. 2009. The burden of knowledge and the "death of the renaissance man": Is innovation getting harder? *Rev. Econom. Stud.* **76**(1) 283–317.

Jones, B. F., S. Wuchty, B. Uzzi. 2008. Multi-university research teams: Shifting impact, geography, and stratification in science. *Science* **322**(5905) 1259–1262.

Kaminer, N., Y. M. Braunstein. 1998. Bibliometric analysis of the impact of Internet use on scholarly productivity. *J. Amer. Soc. Inform. Sci.* **49**(8) 720–730.

Kelley, M. R. 1994. Productivity and information technology—The elusive connections. *Management Sci.* **40**(11) 1406–1425.

Kim, E., A. Morse, L. Zingales. 2009. Are elite universities losing their competitive edge? *J. Financial Econom.* **93**(3) 353–381.

Lee, S., B. Bozeman. 2005. The impact of research collaboration on scientific productivity. *Soc. Stud. Sci.* **35**(5) 673–702.

Levin, S. G., P. E. Stephan. 1991. Research productivity over the life cycle—The evidence for academic scientists. *Amer. Econom. Rev.* **81**(1) 114–132.

Levin, S., P. Stephan, A. E. Winkler. 2010. Innovation in academe: The diffusion of Internet technologies. Working paper, University of Missouri–St. Louis, St. Louis.

Levin, S., W. Glanzel, P. Stephan, A. E. Winkler. 2009. The diffusion of IT and the increased propensity of teams to transcend institutional boundaries. Working paper, University of Missouri–St. Louis, St. Louis.

Lieberson, S., E. O. Bell. 1992. Children's 1st names—An empirical study of social taste. *Amer. J. Sociol.* **98**(3) 511–554.

Long, J., M. F. Fox. 1995. Scientific careers: Universalism and particularism. *Annual Rev. Sociol.* **21** 45–71.

Marwell, G., R. Rosenfeld, S. Spilerman. 1979. Geographical constraints on women's careers in academia. *Science* **205**(4412) 1225–1231.

Murray, F., P. Aghion, M. Dewatripoint, J. Kolev, S. Stern. 2008. Of mice and academics: Examining the effects of openness on innovation. Working paper, Massachusetts Institute of Technology, Cambridge.

National Science Foundation (NSF). 2009. America's investment in the future. http://www.nsf.gov/about/history/nsf0050/internet/launch.htm.

Orlikowski, W. J. 1992. The duality of technology: Rethinking the concept of technology in organizations. *Organ. Sci.* 3(3) 398–427.

Redner, S. 1998. How popular is your paper? An empirical study of the citation distribution. *Eur. Physical J. B* **4**(2) 131–134.

Rogers, E. 2003. *Diffusion of Innovations*, 5th ed. Free Press, New York.

Rosenblat, T. S., M. M. Mobius. 2004. Getting closer or drifting apart? *Quart. J. Econom.* **119**(3) 971–1009.

Santos Silva, J., S. Tenreyro. 2006. The log of gravity. *Rev. Econom. Statist.* **88**(4) 641–658.

Shauman, K. A., Y. Xie. 1996. Geographical mobility of scientists: Sex differences and family constraints. *Demography* **33**(4) 455–468.

Sproull, L., S. Kiesler. 1992. *Connections: New Ways of Working in the Networked Organization*. MIT Press, Cambridge, MA.

Stephan, P. 2010. The economics of science. B. Hall, N. Rosenberg, eds. *Handbook of the Economics of Innovation*, Vol. 1. Elsevier, Amsterdam, 218–273.

Stephan, P., S. Levin. 1992. *Striking the Mother Lode in Science: The Importance of Age, Place and Time*. Oxford University Press, New York.

Stuart, T. E., W. W. Ding. 2006. When do scientists become entrepreneurs? The social structural antecedents of academic entrepreneurship in the academic life sciences. *Amer. J. Sociol.* **112**(1) 97–144.

Van Alstyne, M., E. Brynjolfsson. 2005. Global village or cyber balkans? Modeling and measuring the integration of electronic communities. *Management Sci.* **51**(6) 851–868.

Walsh, J. P., T. Bayma. 1996. Computer networks and scientific work. *Soc. Stud. Sci.* **26**(3) 661–703.

Walsh, J. P., S. Kucker, N. G. Maloney, S. Gabbay. 2000. Connecting minds: Computer-mediated communication and scientific work. *J. Amer. Soc. Inform. Sci.* **51**(14) 1295–1305.

Winkler, A. E., S. G. Levin, P. Stephan. 2010. The diffusion of IT in higher education: Publishing productivity of academic life scientists. *Econom. Innovation New Tech.* **19**(5) 481–503.

Wooldridge, J. M. 2002. *Econometric Analysis of Cross Section and Panel Data*. MIT Press, Cambridge, MA.

Wuchty, S., B. F. Jones, B. Uzzi. 2007. The increasing dominance of teams in production of knowledge. *Science* **316**(5827) 1036–1039.

Xie, Y., K. Shauman. 1998. Sex differences in research productivity. *Amer. Sociol. Rev.* **63**(6) 847–870.

Zakon, R. 2005. Hobbes' Internet timeline v8.1. http://www.zakon.org/robert/internet/timeline.