NBER WORKING PAPER SERIES

THE IMPACT OF INFORMATION TECHNOLOGY ON SCIENTISTS' PRODUCTIVITY, QUALITY AND COLLABORATION PATTERNS

Waverly W. Ding
Sharon G. Levin
Paula E. Stephan
Anne E. Winkler

Working Paper 15285
http://www.nber.org/papers/w15285

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2009

Ding acknowledges support from the Haas School of Business, Lester Center for Entrepreneurship and Innovation, and the Ewing Marion Kauffman Foundation, Kansas City, MO. Levin, Stephan and Winkler acknowledge support from the Andrew W. Mellon Foundation, NY. We thank Kelly Wilken for data assistance and Fiona Murray for helpful comments. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2009 by Waverly W. Ding, Sharon G. Levin, Paula E. Stephan, and Anne E. Winkler. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

The Impact of Information Technology on Scientists' Productivity, Quality and Collaboration Patterns
Waverly W. Ding, Sharon G. Levin, Paula E. Stephan, and Anne E. Winkler
NBER Working Paper No. 15285
August 2009
JEL No. J16,J44,O33

ABSTRACT

This study advances the prior literature concerning the impact of information technology on productivity in academe in two important ways. First, it utilizes a dataset that combines information on the diffusion of two noteworthy and early innovations in IT -- BITNET and the Domain Name System (DNS) -- with career history data on research-active life scientists. This research design allows for proper identification of the availability of access to IT as well as a means to directly identify causal effects. Second, the fine-grained nature of the data set allows for an investigation of three publishing outcomes: counts, quality, and co-authorship. Our analysis of a random sample of 3,771 research-active life scientists from 430 U.S. institutions over a 25-year period supports the hypothesis of a differential return to IT across subgroups of the scientific labor force. Women scientists, early-to-mid-career scientists, and those employed by mid-to-lower-tier institutions benefit from access to IT in terms of overall research output and an increase in the number of new co-authors they work with. Early-career scientists and those in top-tier institutions gain in terms of research quality when IT becomes available at their campuses.

Waverly W. Ding
Haas School of Business
University of California, Berkeley
Berkeley, CA 94720-1900
wding@haas.berkeley.edu

Sharon G. Levin
University of Missouri-St. Louis
St. Louis, MO 63121
slevin@umsl.edu

Paula E. Stephan
Department of Economics
Andrew Young School of Policy Studies
Georgia State University
Box 3992
Atlanta, GA 30302-3992
and NBER
pstephan@gsu.edu

Anne E. Winkler
University of Missouri-St. Louis
St. Louis, MO 63121
awinkler@umsl.edu

I. Introduction

The Internet and other advancements in information technology (IT) have changed the workplace (e.g., Brynjolfsson 1993, 1998, Kelley 1998, Dewan and Kraemer 2000). The impact of such changes is of particular importance to the production of knowledge given that scientific inquiry is highly dependent on collaboration and access to information. However, up to now we have only limited knowledge concerning how advancements in IT have affected the research patterns of scientists over time. One drawback to previous studies is that they typically compare research patterns before and after an IT-innovation became widespread, attributing differences to the new technology without knowing when (and sometimes whether) the new technology actually was available to the individual scientist (e.g., Hamermesh and Oster 2002, Rosenblat and Mobius 2004, Kim, Morse and Zingales 2006, Wuchty, Jones and Uzzi 2007, Butler, Butler and Rich 2008). An important exception is Agrawal and Goldfarb (2008). A further limitation of such studies, which applies to the latter paper as well, is that they rely on aggregated data at the journal article or institutional level, making it difficult to accurately estimate the effect of IT diffusion on an *individual* scientist's knowledge production process.

The few studies that have investigated the impact of IT on productivity and collaboration patterns of *individual* scientists (Hesse, Sproull, Kiesler and Walsh 1993, Cohen 1996, Barjak 2006, Winkler, Levin and Stephan forthcoming) have, with the exception of Winkler et al., relied on self-reported data on IT usage. A weakness of this approach is that it is almost impossible to accurately date the initial adoption of the information technologies investigated in the studies. Winkler et al. overcome this difficulty by appending information on the date of institutional adoption of IT to individual-level data on scientists. Nevertheless, a limitation of all of these studies is that they use cross-sectional data to identify the impact of IT, which raises concerns about proper identification of causal effects.

This study advances the prior literature in two important ways. First, it utilizes a dataset that combines information on the diffusion of two noteworthy and early innovations in IT -- BITNET and the Domain Name System (DNS) -- with career history data on research-active life scientists. This research design allows for proper identification of the availability of access to IT as well as a means to directly identify causal effects. Second, the data set is extremely rich, allowing for an investigation of three publishing outcomes: counts, quality, and co-authorship.

Our research design permits us to test whether the adoption of IT by an institution enhances the research of three specific subgroups of the scientific labor force: (1) female faculty members, who often face greater mobility constraints than their male colleagues; (2) faculty early in their careers, who are likely more willing and able to take advantage of the new technology than more established scientists; and (3) faculty at lower-tier institutions, who are more likely to have fewer in-house colleagues and resources than faculty at top-research universities. Implicit in our analysis is the assumption that faculty members take advantage of the latest and best technology, especially when they have more to gain than others from the new technology.

Our empirical work largely supports the three hypotheses. First, we find that women scientists benefitted more than their male colleagues from the availability of IT in terms of overall output and an increase in the number of new co-authors they acquire. Second, while later-career stage scientists did not benefit from the adoption of IT by their institutions, early-to-mid-career stage scientists did. Finally, where one works mediates the effects: the availability of IT increased productivity of scientists at mid-tier (and in some instances at lower-tier) institutions.

The plan of this paper is as follows. In Section II we review the literature concerning the effect of information technology on scientists' research patterns. Section III summarizes our research design. Section IV introduces BITNET and DNS, two IT innovations that have had an impact on scientists' research and in which we are interested. Section V describes our data, variables and models. Section VI presents the results. Conclusions are drawn in Section VII.

II. Literature

Our three hypotheses regarding the impact of IT on the productivity of scientists have received some previous attention. Specifically, research has looked at how advancements in IT enhance scientists' productivity and connectivity regardless of their "location" in the profession[1] and the extent to which IT enhances the productivity and connectivity of some subgroups (e.g., women, junior faculty members, or those employed by lower-tier institutions) more than others. Below, we review prior findings regarding the IT-research productivity relationship and the IT-collaboration relationship.

IT and Research Productivity

Investigations of the relationship between IT and research productivity generally find support for the view that IT enhances productivity. Hesse et al. (1993) surveyed oceanographers and found a positive relationship between oceanographers' use of computer networks and their publication counts as well as professional recognitions. In a survey of scientists from four

[1] "Location" here refers to the geographical location of a scientist's employment setting as well as social standing in the academic labor force. For example, women scientists are reported to occupy a more disadvantaged position in science. Junior faculty members and those employed by lower-tier institutions also have relatively less resources to support their work.

disciplines--chemistry, philosophy, political science and sociology--and 26 institutions, Cohen (1996) similarly found that scientists who reported using computer-mediated communication tools reported higher numbers of publications and more professional recognition. Winkler et al. (forthcoming) 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 adoption of various indicators of IT. Evidence of a positive IT-productivity relationship is also reported in Kaminer and Braunstein (1998), Walsh, Kucker and Gabby (2000), and Barjak (2006).

The hypothesis of differential IT effects has been tested along three dimensions: institutional status, professional age and gender. Hesse and colleagues (1993) used geographical location to proxy for institutional status because 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 receive a higher productivity gain from IT. Cohen's (1996) study of scientists from a broader set of disciplines found no support for the hypothesis of disproportionate benefits for scientists employed at lower-tiered institutions, and Winkler et al. (forthcoming) found limited support for this hypothesis is their study of life scientists. With regard to seniority, Hesse et al. (1993) reported that junior researchers gained more professional recognition than did their senior colleagues when they engaged in more intensive use of IT. Winkler et al. (forthcoming) examined whether IT access benefits women relative to men, but they found no support for this hypothesis.

Overall, these studies provide mixed empirical evidence with regard to the view that IT differentially affects subsets of the scientific labor force. The data or methodology employed in these studies, however, is sufficiently problematic to lead one to conclude that IT has weak or mixed effects. What is needed is a dataset such as the one examined here – one that combines longitudinal data on career scientists with institutional-level variables reflecting timing of IT adoption – to provide more compelling evidence about the presence and magnitude of causal effects.

IT and Research Collaborations

In recent years, there has also been growing research interest in how IT affects the collaborations of scientists. This research interest has been prompted by the increase in the number of co-authored papers by individuals at different academic institutions and in different countries, as well as in the number of co-authors 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, Jones and Uzzi 2006). Team size even increased in mathematics, generally seen as the domain of individuals working alone and the field least dependent on capital equipment. 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 numbers 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 which had received 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 1975-2005, reaching approximately 40 percent by 2005 (Jones, Wuchty and Uzzi 2008).

Collaboration has increased internationally as well. Levin, Glanzel, Stephan and Winkler's (2009) study of authorship patterns across a wide array of four-year colleges and universities in the U.S. found that the percent 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. For example, Hamermesh and Oster (2002) compared publishing activity in three economics journals over the period 1970-1979 with that in the same journals over 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. Adams et al. (2005) also identified a growing mean distance between coauthors in their analysis of 2.4 million scientific papers, going from 77.7 miles in 1981 to 159.4 miles in 1999. Both studies noted that the rate of increase in collaboration distance was greatest during a period of rapid diffusion of e-mail and an associated drop in communication costs. A recent survey of U.S.-trained doctorates further points out the importance of e-mail to international collaboration: 98% of all scientists and engineers in the 2006 Survey of Doctorate Recipients who reported having an international collaborator used either the phone or e-mail to collaborate with their co-author(s) (National Science Board, forthcoming).

While the research, as outlined above, provides compelling evidence that IT enhances collaboration, empirical findings are more mixed with regard to which subgroup(s) of scientists benefit relatively more from the expansion of collaboration networks. Agrawal and Goldfarb (2008) found that faculty at medium-ranked research universities benefitted the most from the adoption of BITNET in the form of increased collaborations in the field of electrical engineering.

Butler et al. (2008) studied the impact of IT on collaboration in economics and political science using publication data from the top three journals in each field. In their paper they measure the availability of IT based on a review of NBER working papers published during the 1990s. They find that prior to January 1997, an e-mail address was never included; since January 1999, almost all papers had an e-mail address. Using this as an indicator of IT access, they find that collaboration increased, especially at lower-ranked institutions. With regard to differential effects of IT by gender on collaboration, Walsh et al. (2000) find limited evidence and Butler et al. (2008) find no significant differences. None of the three studies examine whether effects vary with professional age.

The mixed findings of these studies are due in part to the empirical approaches taken. As noted, with the exception of Agrawal and Goldfarb (2008), prior studies of the IT-collaboration relationship infer the role of IT rather than explicitly measure the presence of IT. Without the actual date or year of institutional or individual adoption of IT innovations, such an estimation strategy cannot account for the variations in the timing of IT adoptions across institutions or individuals. Moreover, such studies look at collaboration at the journal or institutional level. While these types of analyses are useful in identifying the increasing trend in collaboration, they are not suitable for estimating the effect of IT on an *individual* scientist's collaboration network, lacking adequate controls for possible confounding factors (e.g., Ph.D. cohort, career stage, gender and institutional prestige, etc.).

III. Research Design

We investigate the effects of IT on productivity and collaboration patterns of research-active academic life scientists. We also include in our analysis a measure of the quality of scientists' research. Our goal is to assess whether specific subgroups of the scientific labor force gain more than others in terms of productivity, quality and collaborative patterns after an IT innovation is made available to them. Again, following prior literature, we group the scientific labor force along three dimensions: gender, professional age and employer prestige.

Compared with previous work, our empirical design has three advantages. First, we used a dataset of a random sample of approximately 4,000 scientists. These scientists are drawn from multiple scientific disciplines and a wide range of academic institutions with different levels of institutional prestige and span a 25-year period. Detailed career history information as well as research productivity, quality and collaboration information on these scientists is collected from archival data sources. The advantage of such a dataset is that it allows for thorough analysis of each individual scientist's research patterns and adequate control for possible confounding factors. Because the data are fine-grained at an individual level, it is possible for us to hypothesize about and test whether and how the impact of IT is contingent on scientists' individual characteristics. Second, instead of inferring the effect of IT from differences in outcome variables across periods, we directly measure the time that the scientist's employing institution adopted IT innovations. Because we know the initial date when an IT innovation is available to a scientist, we can thus take advantage of the longitudinal nature of our dataset and use a proper lag structure to better establish the effect of IT on research outcomes. Third, prior literature has examined changes in the quantity of scientists' research publications after a change in available IT, but none has investigated the impact of IT on the quality of research. We include a measure for the quality of scientists' research as one of our outcome measures. We believe attention to both the quantity and quality of research allows us to conduct a more comprehensive assessment of the effect of IT on scientists' research.

IV. IT Innovations: BITNET and DNS

This study analyzes the research impact of two indicators of IT connectivity: date of institutional adoption of BITNET and date of institutional adoption of a domain name (DNS).[2] Previous work by Agrawal and Goldfarb (2008) used BITNET data and work by Winkler et al. (forthcoming) examined both indicators.

Although the IT revolution can be dated to the creation of ARPANET by the Department of Defense in 1969, restricted access to ARPANET led others to develop their own networks (NSF 2009). Among these, BITNET was an early leader in electronic communications across a range of scientific disciplines and universities. Conceptualized by the Vice Chancellor of University Systems at the City University of New York (CUNY), BITNET's first adopters were CUNY and Yale in May 1981 (*Bitnet history).* At its peak in 1991-1992, BITNET connected about 1,400 organizations (almost 700 academic institutions) in 49 countries (CREN). By the mid-1990s BITNET was eclipsed by the Internet as we know it today and began to fade away.[3]

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. 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), military (mil), or other (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 of every other

[2] 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. (In process).

[3] 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* available at www.livinginternet.com/u/ui_bitnet.htm).

system on the network, nor would it need to continuously update the file containing this information as the number of hosts on the Internet grew exponentially.

Figure 1 shows the diffusion of BITNET over the period 1981-1990 and the diffusion of DNS over the period 1985-1993 at the 430 academic institutions where the research-active life scientists in our study are located.[4] Both figures exhibit the typical S-curve associated with diffusion of an innovation over time (Rogers 2003)—especially among the non-top institutions; adoption first rises at an increasing rate and then levels off.

Diffusion patterns very considerably by tier. Among the top 25 research institutions, BITNET and DNS diffused rapidly; in the case of DNS approximately all 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. Indeed, by the end of the period, only 50% of these institutions had access to BITNET and approximately 60% had access to DNS.

V. Data, Variables and Model

Data

To determine whether the adoption by institutions of BITNET and DNS led to systematic differences in scientists' research outcomes, we begin with a random sample of 12,000 life scientists in the U.S., drawn from the *UMI Proquest Dissertations* database.[5] We restrict our

[4] Data on the adoption dates of BITNET beyond 1990 are not available. See *Atlas of cyberspaces,* available at wwwlib.umi.com/dissertations.

[5] 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

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. We use the Web of Science's *Science Citation Index* to collect the publications, co-authors, 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 dataset of scientist-year observations from 1969 to 1993 with annually updated covariates for the individual and the employer institution.[6]

Each scientist begins in the data the year he or she receives a Ph.D. and continues until (i) there is a 5-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. These restrictions result in a dataset of 3,771 scientists with 46,301 scientist-year observations.

Variables

For each scientist-year observation, we created covariates including the individual's gender, professional experience, number of jobs, publication and citation counts, co-authorship patterns, employer ranking and the availability of the two IT-related technologies, namely BITNET and DNS. Detailed definitions of these variables are given in Table 1 and descriptive statistics are provided in Table 2.

---

research project was designed to study university faculty members' engagement in the commercialization of academic science. Despite the sampling method, our sample is highly representative of the underlying population of academic life scientists. See Appendix 1 for more information on how our sample compares the NSF and SESTAT's definition of life sciences.

[6] 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 hasn't adopted it.

The three dependent variables in the analysis are (i) research productivity, defined as the number of research papers published by a scientist in a given year, (ii) research quality, defined as the average journal impact factor for papers published by a scientist in a given year (following Azoulay, Ding, and Stuart forthcoming), and (iii) co-authorship gain, which measures the increase in the number of new coauthors found in a scientist's publications in a given year. An example is provided in Table 1.

Our coding of gender is primarily based on the first name of a scientist.[7] When a first name is androgynous, we search the web for the scientist's vitae, bio-sketch or pictures, and code gender accordingly. This strategy permits us to confidently identify gender for 98 percent 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).

Other individual variables included in the models are professional age, defined as the number of years lapsed since a scientist obtained his or her Ph.D. degree, and the number of jobs a scientist has held to proxy the range of his professional network. 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 number of citations received per paper for the papers published by a scientist up through a given year. The total citation count for each

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

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. The predicted, cumulative citation count is then divided by the publication stock to obtain the average citation count per paper.

Two variables are included in the analysis regarding the scientist's institution: employer ranking and whether the institution has adopted BITNET and DNS. The former come from the Gourman Report, which issued rankings for graduate schools beginning in 1980. We assigned universities their original rating for all years prior to 1980 and updated them every other year for the subsequent period. Continuous ranking measure is not informative in our models, so we group institutions into three categories--top 25, between 26 and 50, and below 50. Information on if and when the scientist's university adopted BITNET and DNS come from the Atlas of Cyberspaces and the ALLWHOIS website, respectively, and is entered in the models as a dichotomous variable with one indicating a scientist's university has already adopted the IT technology (at least one year prior) and zero otherwise.

Model

We use a pooled cross-sectional Poisson Quasi-Maximum Likelihood Estimator (PQMLE) to examine the effects of BITNET and DNS on the three outcomes of scientific research noted above: productivity, quality, and collaboration patterns. The choice of a pooled cross-sectional model over a fixed effect model is driven largely by our interest in how some important population characteristics (e.g., gender and employer ranking) affect the outcome variables.[8] 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 PQML estimator does not impose an equi-disperson condition as in a standard Poisson estimator (Wooldridge, 2002). Further, "robust" standard errors are consistent even if the underlying data generating process is not Poisson. In fact, the PQML estimator can be used for any non-negative dependent variable, whether integer or continuous. In our research quality models, even though the outcome variable contains non-integer values, the QML estimator should still be consistent as long as our conditional mean is correctly specified (Gourieroux et al. 1984, Santos Silva and Tenreyro 2006).

VI. Results

Effect of IT on Research Productivity. The regression results are reported in Table 3. The baseline model includes a set of control variables: calendar-year dummies, Ph.D. subject-field dummies, professional experience, gender, number of jobs held, lagged publication stock, lagged average citation count, lagged past-5-year co-authoring ties, and the employer's ranking category.

As expected, research productivity is a concave function of a scientist's professional age.[9] Women scientists have lower productivity than men. Number of jobs held, publication stock, average citation and past co-authoring ties all show a positive association with the current year's publication count. Finally, scientists working at lower-tier institutions (ranked outside the top-

[8] Though not reported in the paper, we have run fixed effect Quasi-Maximum-Likelihood Poisson regressions for models in which time-invariant variables are not our main interest. The fixed effect PQML estimates on the technology variables are largely consistent with the results of pooled cross-sectional PQML. The fixed effect PQML results are available upon request.
[9] Professional age (experience) is highly correlated with a scientist's actual age (Stephan and Levin 1992).

50) appear to publish less than those employed by the base group (institutions ranked in the top-25), although the statistical power of the relationship is weak.

Holding these baseline factors constant, we find that neither the availability of BITNET nor DNS at a scientist's institution, measured with a one-year lag, was associated with a significant increase in research productivity (see models 2 and 3 in Table 3).[10] This finding may, however, mask variations in the effect of IT on different groups within the sample. To test the three hypotheses set forth at the outset of the paper, we interact the IT variables with three covariates: gender, professional experience and employer ranking. The results of IT-gender interactions are reported in columns 4 and 5 of Table 3. For female scientists the availability of BITNET and DNS increases research productivity by 13% (=exp[0.018+0.102]) and 9% (=exp[-0.015+0.099]), respectively. Both coefficients are significantly higher than the effects found for male scientists, measured by the exponent of the coefficient on the non-interacted IT variable.

We next group a scientist's professional experience into five categories (1-4, 5-8, 9-14, 15-20 and 21-26 years after the Ph.D. was earned) and interact each with the two IT variables. Results are reported in columns 6 and 7 of Table 3, and also illustrated in the top panel of Figure 2. We find that the effect of IT changes as a scientist's career progresses. IT has no significant impact at an early-career stage (e.g., 1-8 years after the Ph.D. was earned); it has a stronger and more positive effect on research productivity for mid-career stage scientists (with 9-14 years of professional experience), increasing output by approximately 16%. The effect on research productivity decreases significantly for more mature scientists (with 21 or more years of professional experience).

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

Last, we interact the IT variables with the employer's rank (top 25, 26-50, and outside top 50); these results are reported in columns 8 and 9 of Table 3 and are also shown in the top panel of Figure 2. We find that scientists employed at low-to-mid-tier institutions (outside top 50 and 26-50, respectively) benefit from IT. Those employed by top-25 institutions do not. For scientists at mid-ranked institutions, access to BITNET boosts output by 17%; access to DNS boosts output by 15%. For scientists who are employed by universities ranked outside the top-50, access to BITNET is associated with a 6% increase in the current year's publication count; no significant effect is found for DNS.

Effect of IT on Research Quality. Our second set of models (see Table 4) focuses on the effect of IT on research quality. The estimations control for a set of individual and institutional factors similar to those controlled for in the estimations described above, with the exception of the addition of a control for past research quality (measured by the average journal impact factor for all papers published by a scientist up through the previous year). Holding these factors constant, we observe no significant main effect of BITNET or DNS on quality.

To assess whether the IT effect varies across subgroups of scientists, we once again interact the IT variables with gender, professional experience and employer ranking. We find some notable differences between the quality results and those found in the earlier productivity models. First, there is no evidence that women scientists benefit (in term of improved publication quality) from either BITNET or DNS significantly more than men (see columns 4 and 5 of Table 4). Second, early-career-stage scientists (those with 1-4 years of experience) gain most in terms of research quality from the availability of these technologies: an 11% increase in the average journal impact factor when BITNET is available, and a 14% increase when DNS is available (see columns 6 and 7 of Table 4 and the middle panel of Figure 2). Third, and contrary to what we observed for research output, the quality-boosting effect of BITNET and DNS exists only for scientists at top-ranked (within top-25) universities, who enjoy an 11% gain when their universities adopt BITNET or DNS (see columns 8 and 9 of Table 4 and the middle panel of Figure 2).

Effect of IT on Research Collaboration. Our last set of models (see Table 5) concerns the effect of IT on collaboration patterns of scientists. For this purpose, we examine how the availability of IT changes the number of new co-authors on a project as evidenced by publications in a given year.[11] Holding constant our standard control variables, we find that BITNET and DNS do not lead to significant changes in the number of new co-authors. There are, however, substantial and significant differences in the effect of IT across subsets of the scientists in our sample.

First, as in the case of research productivity, the effect of IT is significantly higher for women than for men (see columns 4 and 5 of Table 5). In term of co-authors, women gain 14% (=exp[0.135]) more than men when their university has access to BITNET in the previous year, and 16% (=exp[0.150]) more than men, with access to DNS. Second, breaking down the effects of BITNET and DNS by career stage, we find that early-to-mid-career scientists (with 1-14 years of experience) benefit from the adoption of these technologies at their campuses, but the IT effects on collaboration decrease significantly during the latter half of the career (see columns 6 and 7 of Table 5 and the illustration in the bottom panel of Figure 2). The effects are highly significant in the case of BITNET and weakly significant in the case of DNS. IT is particularly instrumental in helping newly-minted Ph.D. researchers (with 1-4 years of experience) expand

[11] Although the maximum number of coauthor gain in a year (=170) may seem high, it is not unusual for life scientists to be part of a research team consisting of a few hundred people. Nonetheless, we tested our models in Table 5 by excluding the observations at the top 1% of the "coauthor gain" variable. We found no meaningful difference from the results reported in Table 5.

their collaboration networks. Lastly, scientists who are employed by mid-ranked universities enjoy a 13% boost in the number of new co-authors when BITNET becomes available at their universities. No statistically significant effect is found for DNS (see columns 8 and 9 of Table 5 and the illustration in the bottom panel of Figure 2). Scientists employed by universities outside the top 50 experience approximately a 7% gain in new co-authors when their employer adopted BITNET; this effect just misses being significant at the 5% level.

VII. Conclusion

The internet and other advancements in information technology (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 due to the absence of data linking information on the adoption of IT technology to the productivity of scientists. Here we remedy this situation by creating a database that combines information on the diffusion of BITNET and the Domain Name System (DNS) with career history data on the publishing patterns of research-active academic life scientists. Three characteristics of publishing patterns are measured: counts of publications, quality of publications and an increase in co-authorship. We test whether the adoption of IT by an institution enhanced the research of specific subgroups of the scientific labor force: (1) female faculty members, (2) faculty early in their careers, and (3) faculty at lower-tier institutions.

Our novel approach involves determining the date that BITNET and DNS were adopted by each of the 430 institutions our sample of research-active life scientists worked at and relating their availability to three outcomes of the research process. While we find no direct effect of IT, our findings support the hypotheses that the adoption of IT had differential effects on productivity depending on a scientist's individual characteristics and location. Specifically, women scientists benefitted more than their male colleagues in terms of overall output and an increase in new co-authors. This is consistent with the idea that IT is especially beneficial to individuals who face greater mobility constraints. Despite the plusses, IT was not able to boost the quality of women's publications, suggesting that their new coauthors were located in lower-tier institutions. Second, late-career stage scientists did not benefit from the adoption of IT by their institutions while early-to-mid-career stage scientists—who likely were more willing and able to take advantage of the new technology—did in terms of research quantity, quality and collaboration networks. Third, the tier of the research organization matters. The availability of IT increased the productivity of scientists at mid-tier (and in some instances lower-tier) institutions. This is consistent with the idea that faculty at mid-and-lower-tier institutions had relatively more to gain, having fewer in-house colleagues and resources, although some scientists at the lowest-tier institutions may be too isolated and resource-poor for IT to have made a significant difference, particularly in terms of the quality of research. Finally, the most notable effects of IT are on collaboration and are consistent with the frequent inference that IT has been a major contributing factor to the increase in the number of co-authors in science observed since the 1980s.

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 co-authorship, enabling scientists outside the inner circle to participate more fully. We would be remiss if we did not point out research on policy innovations that have increased accessibility in science and thus had a similar effect on the practice of science. Murray and her colleagues (2008), for instance, studied the impact of two Memoranda of Understanding (MOU) between Dupont and the National Institutes of Health which 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.[12] 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 which had been exogeneously 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 benefitted more than those at the top 50 in terms of a post-deposit citation boost to papers which 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.

[12] Researchers at institutions where a colleague had either engineered a mouse or accessed a mouse were likely to share the benefits while 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.

References

Adams, J.D., G.C. Black, J.R. Clemmons, P.E. Stephan. 2005. Scientific Teams and Institutional Collaborations: Evidence from US Universities, 1981-1999. *Research Policy* 34(3) 259-285.

Agrawal, A., A. Goldfarb. 2008. Restructuring Research: Communication Costs and the Democratization of University Innovation. *American Economic Review* 98(4) 1578-1590.

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

Azoulay, P., W.W. Ding, T.E. Stuart. 2008. The Impact of Academic Patenting on the Rate, Quality, and Direction of (Public) Research Output. *Journal of Industrial Economics* forthcoming.

Barjak, F. 2006. The Role of the Internet in Informal Scholarly Communication. *Journal of the American Society for Information Science and Technology* 57(10) 1350-1367.

Bitnet history. (http://www.livinginternet.com/u/ui_bitnet.htm).

Brynjolfsson, E. 1993. The Productivity Paradox of Information Technology. *Communications of the Acm* 36(12) 67-77.

Brynjolfsson, E., L.M. Hitt. 1998. Beyond the Productivity Paradox. *Communications of the Acm* 41(8) 49-55.

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 Science & Politics* 41(3) 579-584.

Cohen, J. 1996. Computer Mediated Communication and Publication Productivity among Faculty. *Internet Research-Electronic Networking Applications and Policy* 6(2-3) 41-&.

CREN History and Future. (http://www.cren.net/cren-hist-fut.html.)

Dewan, S., K.L. Kraemer. 2000. Information Technology and Productivity: Evidence from Country-Level Data. *Management Science* 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.

Furman, J. L. and S. Stern. 2009. Climbing Atop the Shoulders of Giants: The Impact of Institutions on Cumulative Research. Working Paper.

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. *Economic 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. *Communications of the Acm* 36(8) 90-101.

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. *Journal of the American Society for Information Science* 49(8) 720-730.

Kelley, M.R. 1994. Productivity and Information Technology - the Elusive Connections. *Management Science* 40(11) 1406-1425.

Kim, E., A. Morse and L. Zingales. 2006. Are Elite Universities Losing Their Competitive Edge? *Journal of Financial Economics.* Forthcoming.

Levin, S.G., P.E. Stephan, M.B. Walker. 1995. Plancks Principle Revisited - A Note. *Social Studies of Science* 25(2) 275-283.

Levin, S., W. Glanzel, P. Stephan, A. Winkler. 2009. The Diffusion of IT and the Increased Propensity of Teams to Transcend Institutional Boundaries *Working Paper*.

Levin S., P. Stephan, and A.Winkler. In Process. Innovation in Academe: The Diffusion of Internet Technologies.

Lieberson, S., E.O. Bell. 1992. Children's 1st Names - an Empirical Study of Social Taste. *American Journal of Sociology* 98(3) 511-554.

Murray, F., P.Aghion, M. Dewatripoint, J. Kolev and S. Stern, 2008. Of Mice and Academics: Examining the Effects of Openness on Innovation. Working Paper.

National Science Board, forthcoming, *Science and Engineering Indicators 2010*.

National Science Foundation. 2009. *America's Investment in the Future*. Available at (www.nsf.gov/about/history/nsf0050/internet/launch.htm).

Redner, S. 1998. How Popular Is Your Paper? An Empirical Study of the Citation Distribution. *European Physical Journal B* 4:131–34.

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

Rosenblat, T.S., M.M. Mobius. 2004. Getting closer or drifting apart? *Quarterly Journal of Economics* 119(3) 971-1009.

Silva, J., S. Tenreyro. 2006. The Log of Gravity. *Review of Economics and Statistics* 88(4) 641-658.

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

Stuart, T.E., W.W. Ding. 2006. When Do Scientists Become Entrepreneurs? The Social Structural Antecedents of Academic Entrepreneurship in the Academic Life Sciences. *American Journal of Sociology* 112(1): 97-144.

Walsh, J.P., T. Bayma. 1996. Computer Networks and Scientific Work. *Social Studies of Science* 26(3) 661-703.

Walsh, J.P., S. Kucker, N.G. Maloney, S. Gabbay. 2000. Connecting Minds: Computer-Mediated Communication and Scientific Work. *Journal of the American Society for Information Science* 51(14) 1295-1305.

Winkler, A.E., S.G. Levin, P. Stephan. Forthcoming. The Diffusion of IT in Higher Education: Publishing Productivity of Academic Life Scientists. *Economics of Innovation and New Technology.*

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

Wuchty, S., B.F. Jones, B. Uzzi. 2007. The Increasing Dominance of Teams in Production of Knowledge. *Science* 316(5827) 1036-1039.

Table 1
Variable Definitions 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
Research Quality	Average journal impact factor (JIF) for papers published by a scientist in a given year	Web of Science
Co-authorship 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 co-authorship 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 co-authorship gain for year *t*	Web of Science
Professional Experience	Number of years elapsed from the year a scientist receives his Ph.D. degree	UMI Proquest Dissertations
Female	1 = Yes; 0 = No	Naming convention
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
Avg. JIF of all past publications	Average journal impact factor (JIF) for the papers published by a scientist between his Ph.D. grant year and the current year	Web of Science; Journal Citation Reports
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 on page 13)	Web of Science
Past 5-year co-authoring ties	Number of co-authorship dyads in papers published by a scientist between *t*-5 and *t*; for a paper written by a scientist with two coauthors, two co-authorship 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
Employer rank	Ranking categories of employer university (1-25, 26-50 or outside of top 50)	The Gourman Reports
BITNET	1 = Employer university has adopted Bitnet; 0=otherwise	Atlas of Cyberspace
DNS	1 = Employer university has adopted DNS; 0 = otherwise	ALLWHOIS
Ph.D. subject field	Field in which a scientist is awarded his Ph.D. degree	UMI Proquest Dissertation
Year	Calendar year	

Table 2
Descriptive Statistics

Variable Name	Mean	Standard Deviation	Min	Max
Research Productivity (Publication flow)	1.587	2.208	0	35
Research Quality (Avg journal impact factor)	2.503	3.148	0	24.59
Co-authorship Gain	2.308	4.273	0	170
Professional Experience	8.344	5.762	1	26
Female	0.179	0.383	0	1
Number of Jobs	1.335	0.573	1	5
Publication Stock	13.037	19.785	0	298
Avg. JIF of all past publications	3.325	2.773	0	24.59
Average Citation Count	12.202	18.805	0	1270.1
Past 5-year co-authoring ties	18.201	29.165	0	688
Employer rank 1-25	0.273	0.445	0	1
Employer rank 26-50	0.144	0.351	0	1
Employer rank outside 50	0.583	0.493	0	1
BITNET	0.332	0.471	0	1
DNS	0.240	0.427	0	1
Year	1983.3	6.266	1968	1993

Table 3
Poisson Quasi-Maximum-Likelihood Estimate of Effect of Information Technology on Research Productivity

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(8)
Experience = 5-8 years	0.055 (0.021)**	0.053 (0.021)*	0.053 (0.022)*	0.052 (0.021)*	0.053 (0.022)*			0.049 (0.021)*	0.048 (0.022)*
Experience = 9-14 years	0.120 (0.037)**	0.118 (0.037)**	0.118 (0.037)**	0.117 (0.037)**	0.117 (0.037)**			0.109 (0.037)**	0.108 (0.037)**
Experience = 15-20 years	-0.064 (0.044)	-0.067 (0.045)	-0.067 (0.045)	-0.066 (0.045)	-0.067 (0.045)			-0.080 (0.044)†	-0.081 (0.044)†
Experience = 21-26 years	-0.237 (0.058)**	-0.240 (0.058)**	-0.240 (0.058)**	-0.237 (0.058)**	-0.237 (0.058)**			-0.256 (0.058)**	-0.258 (0.057)**
Female	-0.155 (0.034)**	-0.155 (0.034)**	-0.155 (0.034)**	-0.197 (0.041)**	-0.185 (0.039)**	-0.157 (0.034)**	-0.155 (0.034)**	-0.154 (0.033)**	-0.154 (0.033)**
Number of jobs	0.157 (0.026)**	0.158 (0.026)**	0.158 (0.025)**	0.158 (0.026)**	0.158 (0.025)**	0.162 (0.028)**	0.162 (0.027)**	0.155 (0.025)**	0.154 (0.025)**
Publication Stock$_{t-1}$	0.006 (0.001)**	0.006 (0.001)**	0.006 (0.001)**	0.006 (0.001)**	0.006 (0.001)**	0.006 (0.001)**	0.006 (0.001)**	0.006 (0.001)**	0.006 (0.001)**
Average Citation Count$_{t-1}$	0.001 (0.001)*	0.001 (0.001)*	0.001 (0.001)*	0.001 (0.001)*	0.001 (0.001)*	0.001 (0.001)*	0.001 (0.001)*	0.001 (0.001)*	0.001 (0.001)*
Past 5-year co-authoring ties$_{t-1}$	0.007 (0.001)**	0.007 (0.001)**	0.007 (0.001)**	0.007 (0.001)**	0.007 (0.001)**	0.007 (0.001)**	0.007 (0.001)**	0.007 (0.001)**	0.007 (0.001)**
Employer rank 26-50	0.043 (0.037)	0.044 (0.037)	0.045 (0.036)	0.044 (0.037)	0.045 (0.036)	0.046 (0.037)	0.045 (0.037)		
Employer rank outside top 50	-0.057 (0.030)†	-0.049 (0.029)†	-0.049 (0.028)†	-0.049 (0.029)	-0.049 (0.028)†	-0.049 (0.030)	-0.046 (0.029)		
BITNET$_{t-1}$		0.034 (0.030)	0.034 (0.027)	0.018 (0.033)	0.034 (0.027)		0.037 (0.027)		0.048 (0.028)†
DNS$_{t-1}$			0.0004 (0.037)		-0.015 (0.039)				
Female × BITNET$_{t-1}$				0.102 (0.048)*					
Female × DNS$_{t-1}$					0.099 (0.048)*				
Exp 1-4 years × BITNET$_{t-1}$						0.022 (0.049)			
Exp 5-8 years × BITNET$_{t-1}$						0.054			

						(0.036)			
Exp 9-14 years × $BITNET_{t-1}$						0.147			
						(0.034)**			
Exp 15-20 years × $BITNET_{t-1}$						-0.046			
						(0.046)			
Exp 21-26 years × $BITNET_{t-1}$						-0.190			
						(0.078)*			
Exp 1-4 years × DNS_{t-1}							0.016		
							(0.061)		
Exp 5-8 years × DNS_{t-1}							0.023		
							(0.046)		
Exp 9-14 years × DNS_{t-1}							0.145		
							(0.034)**		
Exp 15-20 years × DNS_{t-1}							-0.071		
							(0.057)		
Exp 21-26 years × DNS_{t-1}							-0.196		
							(0.087)*		
Employer rank 1-25 × $BITNET_{t-1}$								0.011	
								(0.051)	
Employer rank 26-50 × $BITNET_{t-1}$								0.162	
								(0.040)**	
Employer outside 50 × $BITNET_{t-1}$								0.061	
								(0.031)†	
Employer rank 1-25 × DNS_{t-1}									-0.051
									(0.064)
Employer rank 26-50 × DNS_{t-1}									0.140
									(0.042)**
Employer outside 50 × DNS_{t-1}									0.041
									(0.034)
Constant	-0.666	-0.672	-0.672	-0.668	-0.670	-0.678	-0.680	-0.695	-0.694
	(0.208)**	(0.208)**	(0.208)**	(0.208)**	(0.208)**	(0.208)**	(0.208)**	(0.205)**	(0.205)**
Log Pseudo-Likelihood	-77781.3	-77776.7	-77776.7	-77765.2	-77767.1	-77853.4	-77866.4	-7777.2	-77765.7
Wald Chi2	2758.9	2808.4	2809.0	2791.0	2804.8	2188.9	2304.4	2930.9	3018.2
Number of covariates	43	44	45	45	46	44	45	44	45

Notes: (1) Number of observations = 46,301; number of researchers = 3,771; number of institutions = 430. (2) All models control for calendar year dummies and Ph.D. subject field dummies. (3) Robust standard errors in parentheses, clustered around scientists. (4) † significant at 10%; * significant at 5%; ** significant at 1%.

Table 4
Poisson Quasi-Maximum-Likelihood Estimate of Effect of Information Technology on Research Quality

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Experience = 5-8 years	-0.185	-0.184	-0.184	-0.184	-0.184			-0.191	-0.192
	(0.018)**	(0.018)**	(0.018)**	(0.018)**	(0.018)**			(0.018)**	(0.018)**
Experience = 9-14 years	-0259	-0257	-0257	-0257	-0257			-0269	-0271
	(0.024)**	(0.024)**	(0.024)**	(0.024)**	(0.024)**			(0.024)**	(0.024)**
Experience = 15-20 years	-0.346	-0.344	-0.344	-0.344	-0.344			-0.357	-0.359
	(0.032)**	(0.032)**	(0.032)**	(0.032)**	(0.032)**			(0.032)**	(0.033)**
Experience = 21-26 years	-0.420	-0.418	-0.418	-0.418	-0.418			-0.431	-0.432
	(0.051)**	(0.051)**	(0.051)**	(0.051)**	(0.051)**			(0.051)**	(0.051)**
Female	-0.046	-0.046	-0.046	-0.046	-0.047	-0.036	-0.034	-0.046	-0.045
	(0.021)*	(0.021)*	(0.021)*	(0.026)†	(0.024)*	(0.022)†	(0.022)	(0.021)*	(0.021)*
Number of jobs	0.108	0.107	0.107	0.107	0.107	0.076	0.073	0.107	0.106
	(0.015)**	(0.015)**	(0.015)**	(0.015)**	(0.015)**	(0.015)**	(0.015)**	(0.015)**	(0.015)**
Publication Stock$_{t-1}$	-0.001	-0.001	-0.001	-0.001	-0.001	-0.002	-0.003	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)**	(0.001)**	(0.001)	(0.001)
Avg. JIF of all past publications$_{t-1}$	0.087	0.087	0.087	0.087	0.087	0.087	0.087	0.088	0.088
	(0.003)**	(0.003)**	(0.003)**	(0.003)**	(0.003)**	(0.003)**	(0.003)**	(0.003)**	(0.003)**
Average Citation Count$_{t-1}$	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002
	(0.0004)**	(0.0004)**	(0.0004)**	(0.0004)**	(0.0004)**	(0.0003)**	(0.0003)**	(0.0004)**	(0.0004)**
Past 5-year co-authoring ties$_{t-1}$	0.005	0.005	0.005	0.005	0.005	0.006	0.006	0.005	0.005
	(0.001)**	(0.001)**	(0.001)**	(0.001)**	(0.001)**	(0.001)**	(0.001)**	(0.001)**	(0.001)**
Employer rank 26-50	-0.082	-0.084	-0.084	-0.084	-0.084	-0.088	-0.090		
	(0.024)**	(0.024)**	(0.024)**	(0.023)**	(0.024)**	(0.024)**	(0.024)**		
Employer rank outside top 50	-0.154	-0.160	-0.161	-0.160	-0.161	-0.176	-0.179		
	(0.019)**	(0.019)**	(0.019)**	(0.019)**	(0.019)**	(0.019)**	(0.019)**		
BITNET$_{t-1}$		-0.027	-0.023	-0.027	-0.023		-0.036		0.012
		(0.021)	(0.022)	(0.022)	(0.022)		(0.022)		(0.022)
DNS$_{t-1}$			-0.009		-0.010				
			(0.024)		(0.025)				
Female × BITNET$_{t-1}$				-0.001					
				(0.035)					
Female × DNS$_{t-1}$					0.004				
					(0.037)				
Exp 1-4 years × BITNET$_{t-1}$						0.101			
						(0.036)**			

Exp 5-8 years × $BITNET_{t-1}$						-0.031			
						(0.029)			
Exp 9-14 years × $BITNET_{t-1}$						-0.045			
						(0.024)†			
Exp 15-20 years × $BITNET_{t-1}$						-0.107			
						(0.030)**			
Exp 21-26 years × $BITNET_{t-1}$						-0.186			
						(0.045)**			
Exp 1-4 years × DNS_{t-1}							0.135		
							(0.044)**		
Exp 5-8 years × DNS_{t-1}							-0.031		
							(0.037)		
Exp 9-14 years × DNS_{t-1}							-0.011		
							(0.028)		
Exp 15-20 years × DNS_{t-1}							-0056		
							(0.032)†		
Exp 21-26 years × DNS_{t-1}							-0.124		
							(0.045)**		
Employer rank 1-25 × $BITNET_{t-1}$								0.107	
								(0.026)**	
Employer rank 26-50 × $BITNET_{t-1}$								0.037	
								(0.031)	
Employer outside 50 × $BITNET_{t-1}$								-0.056	
								(0.026)*	
Employer rank 1-25 × DNS_{t-1}									0.100
									(0.030)**
Employer rank 26-50 × DNS_{t-1}									0.041
									(0.037)
Employer outside 50 × DNS_{t-1}									-0.044
									(0.029)
Constant	0.490	0.494	0.4994	0.494	0.494	0.529	0.533	0.381	0.382
	(0.199)*	(0.199)*	(0.199)*	(0.199)*	(0.199)*	(0.199)**	(0.198)**	(0.199)†	(0.199)†
Log Pseudo-Likelihood	-115335.3	-115330.9	-115330.6	-115330.9	-115330.6	-115737.2	-115763.3	-115483.9	-115528.0
Wald Chi2	2059.7	2077.6	2078.8	2079.7	2080.5	2049.5	2045.6	1977.6	1959.9
Number of covariates	44	45	46	46	47	45	46	45	46

Notes: (1) Number of observations = 46,301; number of researchers = 3,771; number of institutions = 430. (2) All models control for calendar year dummies and Ph.D. subject field dummies. (3) Robust standard errors in parentheses, clustered around scientists. (4) † significant at 10%; * significant at 5%; ** significant at 1%.

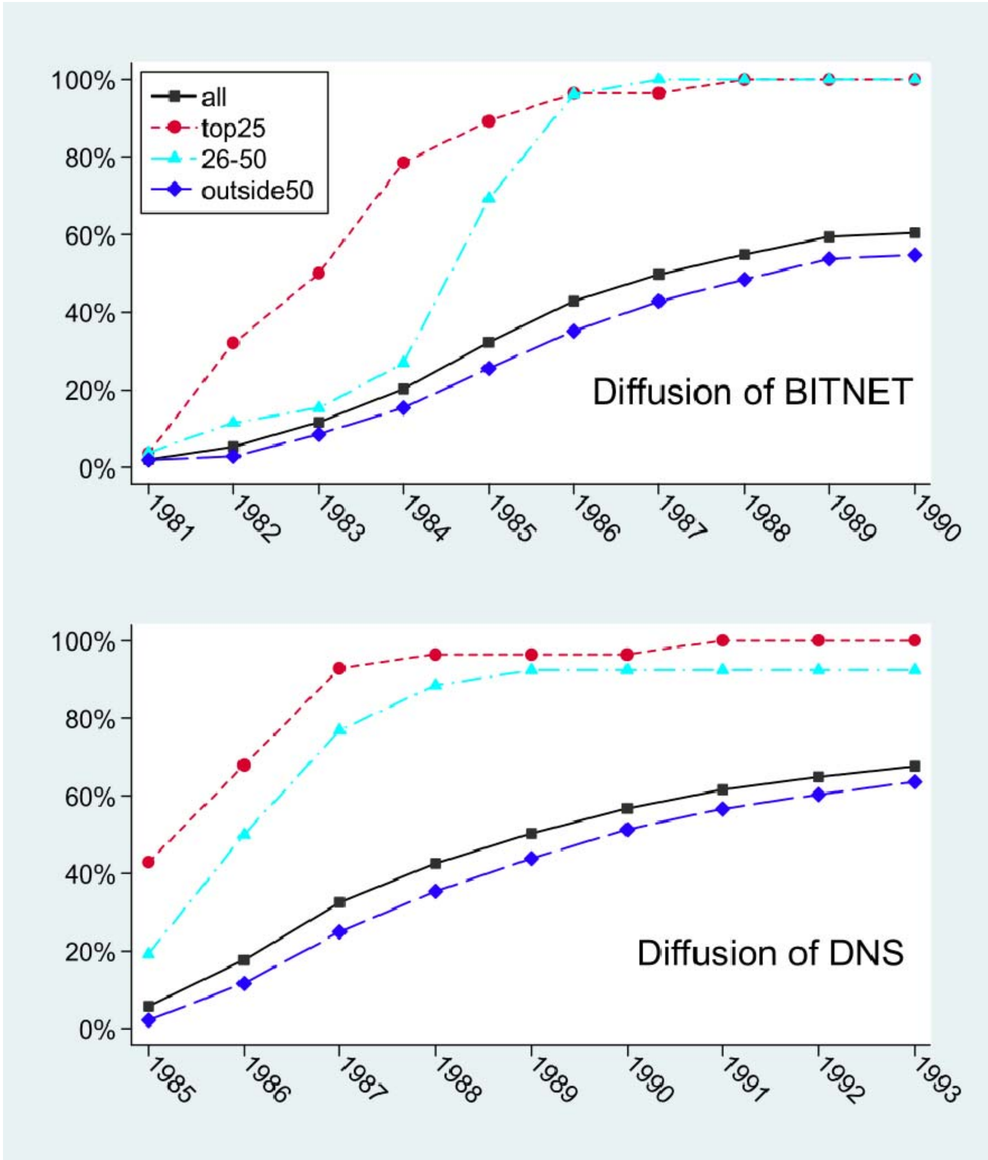
Table 5
Poisson Quasi-Maximum-Likelihood Estimate of Effect of Information Technology on Co-authorship Gain

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Experience = 5-8 years	0.100	0.096	0.097	0.094	0.096			0.094	0.094
	(0.027)**	(0.027)**	(0.027)**	(0.027)**	(0.027)**			(0.027)**	(0.027)**
Experience = 9-14 years	0.133	0.127	0.129	0.126	0.128			0.124	0.124
	(0.042)**	(0.043)**	(0.044)**	(0.043)**	(0.043)**			(0.043)**	(0.043)**
Experience = 15-20 years	-0.040	-0.046	-0.045	-0.046	-0.045			-0.051	-0.053
	(0.047)	(0.048)	(0.048)	(0.048)	(0.048)			(0.047)	(0.047)
Experience = 21-26 years	-0.207	-0.213	-0.212	-0.209	-0.207			-0.220	-0.222
	(0.072)**	(0.072)**	(0.072)**	(0.072)**	(0.073)**			(0.072)**	(0.072)**
Female	-0.087	-0.087	-0.087	-0.150	-0.141	-0.093	-0.091	-0.087	-0.086
	(0.038)*	(0.038)*	(0.038)*	(0.049)**	(0.046)**	(0.039)*	(0.039)*	(0.038)*	(0.038)*
Number of jobs	0.172	0.174	0.173	0.173	0.173	0.186	0.181	0.172	0.171
	(0.031)**	(0.030)**	(0.030)**	(0.030)**	(0.030)**	(0.033)**	(0.032)**	(0.030)**	(0.029)**
Publication Stock$_{t-1}$	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002
	(0.001)*	(0.001)*	(0.001)*	(0.001)*	(0.001)*	(0.001)*	(0.001)*	(0.001)*	(0.001)*
Average Citation Count$_{t-1}$	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Past 5-year co-authoring ties$_{t-1}$	0.009	0.009	0.009	0.009	0.009	0.009	0.009	0.009	0.009
	(0.001)**	(0.001)**	(0.001)**	(0.001)**	(0.001)**	(0.001)**	(0.001)**	(0.001)**	(0.001)**
Employer rank 26-50	0.025	0.028	0.027	0.028	0.027	0.034	0.032		
	(0.042)	(0.042)	(0.041)	(0.042)	(0.041)	(0.043)	(0.042)		
Employer rank outside top 50	-0.031	-0.016	-0.021	-0.016	-0.021	-0.012	-0.013		
	(0.035)	(0.032)	(0.032)	(0.032)	(0.032)	(0.033)	(0.032)		
BITNET$_{t-1}$		0.058	0.074	0.035	0.074		0.082		0.081
		(0.039)	(0.033)*	(0.042)	(0.034)*		(0.034)*		(0.033)*
DNS$_{t-1}$			-0.037		-0.062				
			(0.041)		(0.044)				
Female × BITNET$_{t-1}$				0.135					
				(0.057)*					
Female × DNS$_{t-1}$					0.150				
					(0.055)**				
Exp 1-4 years × BITNET$_{t-1}$						0.205			
						(0.059)**			
Exp 5-8 years × BITNET$_{t-1}$						0.132			

						(0.043)**			
Exp 9-14 years × $BITNET_{t-1}$						0.140			
						(0.042)**			
Exp 15-20 years × $BITNET_{t-1}$						-0.055			
						(0.062)			
Exp 21-26 years × $BITNET_{t-1}$						-0.185			
						(0.077) *			
Exp 1-4 years × DNS_{t-1}							0.116		
							(0.067)†		
Exp 5-8 years × DNS_{t-1}							0.094		
							(0.051)†		
Exp 9-14 years × DNS_{t-1}							0.069		
							(0.042)		
Exp 15-20 years × DNS_{t-1}							-0.137		
							(0.068)*		
Exp 21-26 years × DNS_{t-1}							-0.291		
							(0.078)**		
Employer rank 1-25 × $BITNET_{t-1}$								0.043	
								(0.058)	
Employer rank 26-50 × $BITNET_{t-1}$								0.125	
								(0.048)**	
Employer outside 50 × $BITNET_{t-1}$								0.067	
								(0.040)†	
Employer rank 1-25 × DNS_{t-1}									-0.067
									(0.064)
Employer rank 26-50 × DNS_{t-1}									0.038
									(0.051)
Employer below 50 × DNS_{t-1}									-0.015
									(0.040)
Constant	-2.133	-2.144	-2.140	-2.138	-2.135	-2.158	-2.153	-2.149	-2.148
	(0.571)**	(0.571)**	(0.571)**	(0.571)**	(0.571)**	(0.571)**	(0.571)**	(0.570)**	(0.570)**
Log Pseudo-Likelihood	-116978.4	-116957.5	-116952.3	-116925.6	-116915.3	-117020.1	-117015.5	-116949.5	-116939.5
Wald Chi2	3612.6	3702.6	3709.9	3690.0	3720.6	2921.0	3044.0	3771.8	3788.2
Number of covariates	43	44	45	45	46	44	45	44	45

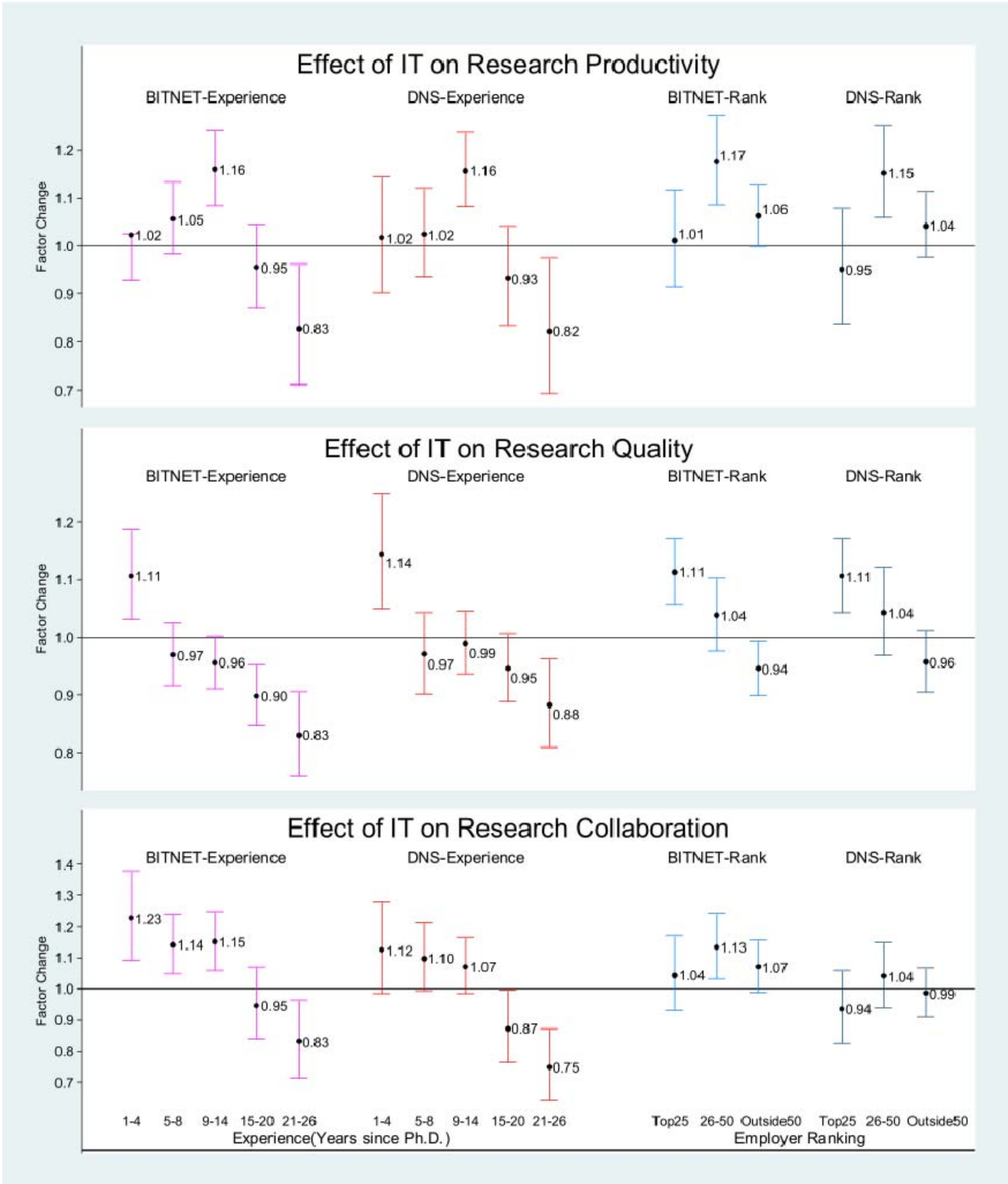
Notes: (1) Number of observations = 46,301; number of researchers = 3,771; number of institutions = 430. (2) All models control for calendar year dummies and Ph.D. subject field dummies. (3) Robust standard errors in parentheses, clustered around scientists. (4) † significant at 10%; * significant at 5%; ** significant at 1%.

Fig.1 Diffusion of BITNET and DNS



Legend. Cumulative percentage of institutions adopting BITNET and DNS. Diffusion patterns are graphed for all institutions (black with squares), institutions ranked in top-25 (red with circles), institutions ranked between 26 and 50 (light blue with triangles), and institutions not ranked (dark blue with diamonds), by the Gourman Report.

Fig.2. Effects of BITNET and DNS on Scientists' Research Patterns



Legend. Effect *g* implies predicted count of publication flow (productivity), average journal impact factor for publication flow (quality) and gain in new co-authors (collaboration pattern) changes by a factor of *g* (or [g-1]*100 percent) when BITNET or DNS becomes available to a scientist in a specified group. 95% confidence intervals of the estimates are indicated by the caps on the spikes. Predictors are statistically significant at the 5% level if 1.0 falls outside the confidence interval. (Full regression results are provided in models 6-9 of tables 2, 4 and 5).

Appendix 1. 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 A1 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 towards fields that are at the forefront of technological developments. We see this as an advantage in our study since 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.

Table A1
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†	(5) Classified as "Life Sciences" in SED††	(6) Rank in SED Based on Representation
487; 303	Biochemistry	845	(22.8%)	Yes	Yes	1
306	Biology, General	563	(15.2%)	Yes	Yes	7
410	Biology, Microbiology	455	(12.3%)	Yes	Yes	2
419	Health Sciences, Pharmacology	234	(6.3%)	Yes	Yes	5
786	Biophysics, General	208	(5.6%)	Yes	Yes	16
490	Chemistry, Organic	198	(5.3%)	No	No	---
369	Biology, Genetics	175	(4.7%)	Yes	Yes	15
433	Biology, Animal Physiology	166	(4.5%)	Yes	Yes	3
982	Health Sciences, Immunology	152	(4.1%)	Yes	Yes	18
307	Biology, Molecular	67	(1.8%)	Yes	Yes	6
301	Bacteriology	61	(1.6%)	Yes	Yes	42
287	Biology, Anatomy	53	(1.4%)	Yes	Yes	14
571	Health Sciences, Pathology	51	(1.4%)	Yes	Yes	24
349	Psychology, Psychobiology	36	(1.0%)	No	No	---
572	Health Sciences, Pharmacy	33	(0.9%)	Yes	No	---

† Source: "Education Code 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).††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.