PRIPE
7/24/06
10:30 AM

Restructuring Research:
Communication Costs and the Democratization of University Innovation

January 2006

Ajay Agrawal and Avi Goldfarb [*]

Abstract
We report evidence indicating that Bitnet adoption facilitated increased research collaboration between US universities; however, not all institutions benefited equally. Using panel data from seven top engineering journals, Bitnet connection records, and a variety of institution ranking data, we find that medium-ranked universities were the primary beneficiaries; they benefited largely by increasing their collaboration with top-ranked schools. In addition, although Bitnet adoption increased collaboration between distant institutions, it had its greatest effect on co-located universities. Thus, our findings suggest the advent of Bitnet – and likely subsequent versions, including the Internet – increased the role of second-tier universities in the national innovation system, particularly those co-located with top-tier institutions.

(JEL O33, R11, Z13)

Keywords: collaboration, Internet, communication technology, productivity, university research, specialization

---

[*]University of Toronto (both authors). Corresponding author: ajay.agrawal@rotman.utoronto.ca. We thank Pierre Azoulay, Eric Brynjolfsson, Iain Cockburn, Jeff Furman, Brent Goldfarb, Shane Greenstein, Bronwyn Hall, Ig Horstmann, Bhaven Sampat, Tim Simcoe, Paula Stephan, Scott Stern, Jerry Thursby, and John Walsh as well as seminar participants at the National Bureau of Economic Research, Georgia Tech, Ecole Polytechnique Fédérale de Lausanne, the Canadian Economics Association meetings, and the University of Toronto for useful comments. We also thank Raghav Misra, Swapnil Kotecha, Cara Saunders, and Alex Oettl, all of whom provided excellent research assistance. Errors and omissions are our own. This research was funded by the Social Sciences and Humanities Research Council of Canada (Grant Nos. 410-2004-1770 and 538-02-1013). Their support is gratefully acknowledged.

1. Introduction

The objective of this research is to further our understanding of knowledge production. How specialized are inputs? Are universities organized to efficiently produce knowledge? To what extent is the knowledge production system sensitive to communication costs? Specifically, we ask: Did researchers at US universities connecting to Bitnet, an early version of the Internet, increase multi-institutional collaboration? If so, by how much? Did all universities that connected benefit equally?

Exploiting the variation in year of adoption and publication output over time in the 289 universities that published in seven top electrical engineering journals, we find that although a Bitnet connection did seem to facilitate a general increase in multi-institutional collaboration (by 85%, on average), all adopters did not benefit equally. Overall, the asymmetric effect of Bitnet seems to have increased the role of second-tier universities, particularly those co-located with top-tier institutions, in the national innovation system. By lowering the costs of communication and the sharing of data, Bitnet facilitated a more efficiently functioning market for inputs into the production of knowledge (both labor and equipment) and thereby contributed to a broadening of the institutions that participated in the production of high-quality electrical engineering research.

To begin, why are we interested in studying research collaboration? Collaboration is important for a number of reasons. First, under certain conditions, inter-organization collaboration is more efficient than single-organization research since collaborators are able to draw upon a wider set of expertise, specialized equipment, and other resources (Cockburn and Henderson, 1998).[1] Second, collaboration facilitates knowledge flows, and knowledge flows, particularly knowledge spillovers, are central to economic growth as characterized in

[1] Cockburn and Henderson measure co-authorships across institutions and report results suggesting that, at least in the pharmaceutical industry, firms that are more open to collaboration with universities are more productive in terms of innovation.

endogenous growth models (Romer, 1986 and 1990).[2] Third, collaboration is an important feature of the national innovation system (Nelson and Rosenberg, 1993). Fourth, collaboration patterns may be partly responsible for the geographic localization of knowledge flows (Jaffe *et al*, 1993; Audrestch and Feldman, 1996; Zucker *et al*, 1998; Agrawal *et al*, 2003; Thompson and Fox Kean, 2005). Finally, infrastructure investments with precisely the aim of fostering research collaboration have been the focus of important public policy objectives, such as the development of the Internet. We examine the actual effects of such an infrastructure in the case of Bitnet.

From a researcher's perspective there are benefits and costs associated with collaboration. So how much will researchers collaborate? In equilibrium, we expect they will collaborate up to the point where the marginal benefit from collaboration equals the marginal cost. But what types of collaborations are on the margin? Observing the changes in collaboration patterns that result from a change in marginal costs offers some insight. Since key costs associated with collaboration include communication, coordination, and the sharing of data, we expect the level of collaboration to increase after Bitnet adoption since electronic networking facilitates these activities at a significantly lower cost.

Interestingly, however, the impact of this cost reduction may vary across pairs of potential collaborators at the margin. If so, adoption will result in changes in collaboration patterns in addition to simply increasing overall collaboration. We explore this possibility along two dimensions: 1) institution quality and 2) distance between collaboration pairs. We next offer intuition regarding how university-pair heterogeneity in these dimensions may be related to the impact of Bitnet with respect to collaborative behavior.

---

[2] Singh (2005) provides evidence suggesting that network ties between inventors, as inferred from their past collaborations, are important determinants of knowledge diffusion patterns.

Beginning with quality, consider universities of two research quality types: high and low. The former type has a strong orientation towards research, which is reflected in large resource allocations to research activities and a broad range of doctoral programs, while the latter does not. Researchers at high-quality universities therefore have greater access to capital-intensive equipment (lasers, robots, simulators, etc.), research grants, and other resource endowments. It also may be the case that more productive researchers select into these resource-rich institutions. Since Bitnet adoption lowers the cost of collaboration, it may benefit high-high quality pairs most since individually these institutions produce the highest volume of research and thus may have the most on which to collaborate; as we will discuss below, top-tier institutions produce approximately six times as many publications as their medium-tier counterparts. A reduction in communication costs may simply serve as a general reduction in the cost of research, thereby facilitating more research by those institutions that already produce the most: top-tier universities. This would reinforce the existing innovation production structure.

At the same time, however, there may be gains from specialization. Researchers in high-quality schools may focus on winning grants, supervising the use of specialized equipment, attending international conferences to present results, and other such activities. Although researchers at lower-quality institutions may not have access to the resources necessary for running certain types of experiments entirely on their own, they may indeed have the expertise and equipment necessary for certain steps in the research process. Using Bitnet, data can be transferred to researchers at lower-quality institutions who are talented and have access to the equipment necessary for data analysis and computing. In other words, a step function reduction in the cost of communication and data sharing may benefit high-low institution pairs most because the technology facilitates specialization and gains from trade in knowledge production.[3] This would lead to a broadening of the institutions that participate in the production of high-quality research.

Considering the above two possibilities, the way in which quality mediates the effect of Bitnet adoption in terms of collaborative research output is ambiguous. We address this empirically below, and our findings suggest that the benefits of Bitnet are greatest for high-medium pairs and that it is the medium-quality universities in these pairings that are the main beneficiaries of Bitnet. In other words, there seems to be a disproportionate number of high-medium collaborations at the margin such that the primary effect of a drop in costs is to facilitate specialization and enhance trade between these particular institutions for inputs in the knowledge production process.

Next, we turn to distance. We assume that the cost of collaboration increases with distance. This is due to costs related to communication (phone and fax), travel, shipping, and convenience caused by differences in time zones. As such, pairs that are furthest apart might benefit most from Bitnet adoption since their collaboration costs are reduced the most. That is, to the extent that Bitnet communication can substitute for at least some phone communication, travel, and shipping (via electronic file transfers) as well as facilitate asynchronous communication (making interaction across time zones easier), the greatest beneficiaries may be those that are furthest apart.

At the same time, however, researchers need to interact in order to collaborate. Collaborations are predicated on shared ideas, which are often the unplanned output of direct interaction (Crane, 1969; Merton, 1973; Mairesse and Turner, 2005). Direct interaction is more

[3] Indeed, several sociologists have characterized the manner in which university researchers utilized early electronic networks as a mechanism for facilitating a division of labor leading to a greater involvement of researchers at "peripheral" institutions (Walsh and Bayma, 1996; Hesse et al, 1993).

likely between individuals who are closer together. More generally, people are more likely to know people who live nearby. So, although the cost reduction per collaboration is greater for pairs that are further apart, pairs that are closer together interact more. As such, the way in which distance mediates the effect of Bitnet is ambiguous. We address this empirically below, and our findings suggest that the benefits of Bitnet are indeed greatest for pairs that are close together. In summary, then, our primary findings suggest that the main effects of the drop in communication and data-sharing costs associated with Bitnet adoption are to facilitate greater participation in the production process of high-quality research by including schools with fewer resources but at the same time to further accentuate the geographic agglomeration of research activities.

Our paper proceeds as follows. In Section 2 we describe how our work builds on the prior literature, and in Section 3 we provide a brief description of Bitnet and how it facilitated research collaboration. In Section 4 we describe the Bitnet adoption data and engineering publication data that form the basis of our key measures. In Section 5 we present our empirical framework, and in Section 6 we present the results, including the estimated Bitnet effect and the estimated degrees to which the effect is mediated by the quality of and distance between potential collaborators. Finally, in Section 7 we discuss the implications of these findings.

2. Literature

Although there is a significant literature concerning the economics of scientific collaboration (Crane, 1969; Beaver and Rosen, 1978, 1979; Barnett, Ault, and Kaserman, 1988; Cockburn and Henderson, 1998; Mairesse and Turner, 2005) and also concerning the social impact of lowering communication costs via the Internet (Gaspar and Glaeser, 1998; Wellman and Gulia, 1997; Smith, 1999; Olson and Olson, 2003; Van Alstyne and Brynjolfsson, 2005), there has been very little prior research examining the effect of electronic network adoption on research collaboration.

A notable exception is Hamermesh and Oster (1998), who explore the effect of the 1980s "communications revolution" on collaborative research in economics. While this prior work is similar to our study in that both are interested in the effect of lowered communication costs on research collaboration, our work differs in three fundamental ways. First, we directly measure the relationship between institution-level Bitnet adoption and collaboration. They instead show that general collaboration patterns changed from the 1970s to the 1990s without showing a direct link to communications technology. Second, we focus on changes in the propensity to collaborate, using institution-pairs as the unit of analysis. They focus on how citation counts of collaborative work changed, using co-authored papers as the unit of analysis. Third, and perhaps most importantly, we examine how the change in collaboration is mediated by the quality of and distance between potential collaborators.

A second exception is Gaspar and Glaesar (1998). While collaboration is not the main focus of their article, they show that there has been a rapid growth in local co-authorships in economics since the 1960s. They use this to support their argument that information technology is a complement to face-to-face interaction. Using more rigorous econometric analysis, we confirm their conjecture. Unlike Gaspar and Glaesar, however, we examine the mediating role of university quality and show that adoption leads to increased productivity among a particular set of schools and results from their interaction with more research-oriented local institutions. Furthermore, our focus is on how communication technology affects collaboration rather than agglomeration.

The final exception is Hesse *et al* (1993). The authors present a cross-sectional analysis that relates individual publication propensity to network usage.[4] They provide evidence that oceanography researchers who use the network more also publish more. Interestingly, they also show that the correlation between high network users and high productivity is greater for "peripheral" scientists, meaning those who are younger or at non-coastal research institutions and thus are likely to have less immediate access to ocean-related, capital-intensive equipment and data.[5] While the authors suggest that collaboration is a likely explanation for the higher levels of productivity, they do not actually measure it. In addition, they are cautious about interpreting their findings as evidence that the network facilitates increased productivity because they do not have time series data and because everyone in their data has adopted.

3. A Brief Description of Bitnet

Our research examines whether and how Bitnet adoption facilitates collaboration. As such, it is important to offer some context in terms of what Bitnet was used for, its relation to other networks, why we focus on this particular network, and what was involved in the adoption process.

Bitnet was an early leader in network communications for the research and education community. It allowed communication via email, access to remote file archives, use of Listserv, and compatibility with other operating systems such as UNIX.[6] Most importantly, however, the network enabled the exchange of data through file transfer protocols. This was particularly useful

[4] Although the authors of this 1993 study cite Bitnet as an important network for research oceanographers, they focus on their survey respondents' use of SCIENCEnet, a similar but smaller electronic network system.
[5] In a related paper, Walsh and Bayma (1996) also document online collaboration between core and peripheral institutions for four other disciplinary areas including mathematics, physics, chemistry, and experimental biology.
[6] http://computing.dcu.ie/~humphrys/net.80s.html (Mark Humphrys, *The Internet in the 1980s*).

for researchers who wished to collaborate on projects that involved large datasets, such as those generated from experiments using capital-intensive equipment.

Our study spans the period 1981-1990. The first Bitnet adopters were the City University of New York (CUNY) and Yale University in May 1981. By the end of the 1980s, Bitnet had become the largest academic network in the world for computer-based communications.[7] The World Wide Web was not invented until 1989 (the end of our study period), and the first mass-market browser, Mosaic, was not developed until 1993.[8] Thus, the version of the network we examine in this paper predates the Internet as it is known today.

Still, other networks did exist at the time. In the late 1960s and early 1970s, the Advanced Research Projects Agency (ARPA) founded the first Internet-like network, ARPANET. This network was primarily used by defense research labs.[9] It facilitated exchange of computer data across North America amongst ARPA-funded researchers for ARPA projects. ARPANET's restriction to only ARPA-funded researchers led to the development of several other networks, including CSNET, USENET, EDUNET, and Bitnet. CSNET was founded in 1981 to serve the needs of computer scientists, while USENET was created by computer science graduate students to help exchange information via newsgroups. EDUNET was established in the mid-1970s to allow sharing of mainframe computing resources across universities. Like ARPANET, it was limited to a small fraction of research institutions.

Bitnet, on the other hand, was a network created to promote the tools of computer networking for all scholars, not just select ARPA-funded researchers or computer scientists or

[7] In 1986, NSFNET was founded for connecting supercomputer centers across North America. It created the backbone of what would become the Internet, leading to the rapid growth of Bitnet over the next few years. In 1987, the Bitnet Information Center introduced Listserv, an automatic mailing list server for Bitnet. Listserv remains Bitnet's lasting legacy and is still employed by many universities around the world. Gurbaxani (1990) provides a detailed account of the diffusion of Bitnet.

[8] The invention of the Web is commonly credited to Tim Berners-Lee at CERN in Switzerland. The browser was developed at the University of Illinois National Center for Supercomputing Applications.

[9] There were only 19 universities of the 62 institutions that were connected to ARPANET by 1979.

individuals at top schools.[10] It was intended as a low-cost and easy-to-setup computer network for communication between university scholars. The network was restricted to academia, and no commercial communication was allowed.[11]

We focus on Bitnet for a number of reasons. First, the academic focus of the network and its broad application allow for a richer understanding of how adoption changes collaboration patterns across a diverse set of institutions.[12] Second, Bitnet adoption was carefully documented; data exist on the exact date of adoption by every institution in the network up to 1990. This is not the case for other networks. Finally, Bitnet's ability to exchange data through file transfer protocols as opposed to certain other networks that only allowed bulletin board postings and text messages is particularly important for studying collaboration in areas that benefit greatly from data sharing, such as electrical engineering.

In order to connect to Bitnet, each institution required a 9600 baud leased line between its computer facility and another institution that was linked with the network as well as IBM networking software.[13] Bitnet setup costs were reasonably low because mainframes already existed on many university campuses and communication relied on dial-up technology. Administratively, however, creating an initial connection to the network seems to have been a reasonably involved process.[14] The decision to adopt was generally made by a university's

---

[10] Ira Fuchs, the Vice Chancellor of University Systems at CUNY, conceptualized Bitnet (Because It's There NETwork) to take advantage of the existing supply of IBM mainframes on many university campuses.
[11] The Bitnet executive committee organized the effort and established regulations and bylaws.
[12] Of the 289 US universities that published in the journals we examine, 225 adopted Bitnet by 1990. Far fewer were connected to the other networks.
[13] Bitnet was a "store-and-forward" network. Information originating at a given Bitnet-connected computer (node) was received by intermediate nodes and forwarded to its destination.
[14] In a rich description of the National University of Singapore's decision to connect, the Dean of Science at that institution describes the bureaucratic steps that he had to take, including obtaining the assistance and approval of a number of senior university administrators (who, in this case, were very supportive) (http://www.physics.nus.edu.sg/~phytanb/bitnet4.htm).

computing center managers in connection with the computer science department and approved by higher-level administrators.

4. Data

We employ four types of data in our empirical work: 1) publication data, 2) Bitnet connection data, 3) institution quality data, and 4) distance data. We describe these below.

4.1 Publication Data

Since we are interested in identifying the Bitnet effect on collaboration, we use publication data from researchers in technical areas that are likely to be early adopters of this communication technology. We do this because we exploit the variation in connection years for identification. The publishing behavior of these technical researchers is most likely to reflect the time variation in adoption. Thus, we select a variety of electrical engineering research topics that appeared in journals published by the Institute of Electrical & Electronic Engineers (IEEE).

Specifically, we collect publication data from seven journals over the 15-year period of 1977-1991.[15] Each of these journals is considered among the top outlets for research in the specified field. Since we focus only on these seven top journals, the total number of publications in our analysis does not rise over time.[16]

[15] The journals are: 1) IEEE Transactions on Aerospace and Electronic Systems, 2) IEEE Transactions on Nuclear Science, 3) IEEE Transactions on Biomedical Engineering, 4) IEEE Journal of Quantum Electronics, 5) IEEE Transactions on Electron Devices, 6) IEEE Transactions on Communications, and 7) IEEE Transactions on Education.

[16] There are a total of 1702 papers published in the first year of observation (1977) and 1646 papers published in the last year (1991). The total number of publications fluctuates from year to year due to the publication of special issues and occasional conference proceedings.

There are 28,312 papers published in these seven journals during the time period under investigation.[17] We parse out all unique author-affiliated institutions from each paper. Papers are categorized as either single-institution or multi-institution (i.e., collaborative). To be clear, papers with multiple authors may still be classified as single-institution if all authors are from the same university.

After extracting the institutional information from the set of publications, we identify 739 unique institutions, of which 289 are US universities – our institution type of interest. These form the basis of our unit of analysis. We focus on US universities because many of the international institutions and US non-university research labs used other networks besides Bitnet. US universities are particularly likely to have Bitnet as the first widely spread data communication network they adopted.[18]

These data allow us to construct two datasets, each of which is focused on the measurement of the number of collaborative (multi-institution) papers. Our main data set consists of 41,616 institution-pairs over the 15 years (1977-1991) of publishing data from the specified journals, resulting in a balanced panel with 624,240 observations. For use in Table 4, we also construct a single-institution dataset that includes the same 15 years of publishing from the specified journals by the 289 institutions of interest. This is therefore a balanced panel dataset that consists of 4335 observations.

4.2 Bitnet Connection Data

[17] The distribution of publications across journals is not uniform. In the order listed in footnote 13, the number of publications per journal is 6174, 8505, 3418, 4976, 827, 3585, and 827, respectively.
[18] http://computing.dcu.ie/~humphrys/net.80s.html (Mark Humphrys, *The Internet in the 1980s*).

We use an online reference, Cyber Geography Research, for a record of Bitnet connections.[19] This archive lists the 1054 institutions worldwide that connected to Bitnet by the end of 1990 as well as their connection date.[20] Of the 289 US universities that published at least once in the seven IEEE journals we examine, 225 connected to Bitnet during this time. In other words, 64 US universities in our publishing sample had not connected to Bitnet by the end of 1990. Figure 1 illustrates the connection rate. The variation in connection years is important for our econometric analysis since we exploit this to identify the effect of Bitnet.

Figure 2 illustrates the geographic distribution of connected universities over the first three years. The first three institutions connected were all on the east coast. By the second year, some universities further inland were connected (such as Ohio State) while two (UC Berkeley and UC San Francisco) were on the west coast. By the third year, universities up and down the east coast were connected as well as more inland institutions (such as the University of Missouri-Columbia) and a third institution on the west coast (Stanford).

4.3 Quality Data

We use the 1987 Carnegie Foundation classification system to classify the research "quality" of each university in our dataset.[21] We classify universities as Carnegie Type 1, 2, or 3 (CT1, CT2, CT3). Carnegie Type 1 (CT1) is an aggregate of the Carnegie Foundation's categories "Research University 1 and 2." Thus, institutions with our CT1 classification offer a full range of baccalaureate programs, are committed to graduate education through the doctorate degree, and

---

[19] http://www.cybergeography.org/atlas/bitnet_topology.txt.

[20] In our analysis, we use the year following the technical connection as the first year Bitnet is available at the university. In the journals examined here, six months is a typical publication lag from manuscript submission to publication. All results are robust to using the year of adoption.

[21] *A Classification of Institutions of Higher Education* (1987 Edition), A Carnegie Foundation Technical Report, Princeton University Press, Lawrenceville, NJ.

give high priority to research. They receive annually at least $12.5 million in federal support and award at least 50 Ph.D. degrees each year.[22]

Institutions with a CT2 classification are an aggregate of the Carnegie Foundation's categories "Doctorate-Granting Universities 1 and 2." These institutions offer a full range of baccalaureate programs, and their mission includes at least some commitment to graduate education through the doctorate degree, such that they award annually 20 or more Ph.D. degrees in at least one discipline or 10 or more Ph.D. degrees in three or more disciplines; however, they do not meet the requirements for CT1. All other universities are classified as CT3.

To ensure robustness, we examine two other definitions of institution quality. First, we use the total number of publications by institutions in the seven IEEE journals in our data from 1972 to 1979, before Bitnet was adopted at any school. We split these by quartile and group the bottom two quartiles together. Those in the top quartile have at least 15 publications over this period. Those in the second quartile have 3 to 14 publications. Those in the remaining two quartiles have two or fewer publications. The bottom two quartiles are grouped together because the total number of publications for these two quartiles over the entire data period (1972 to 1991) is similar. Second, we examine the total number of federal grants received by electrical engineering departments from 1986 to 1992.[23] We again split schools into three groups: those above the median of schools that received grants, those below the median of schools that received grants, and schools that did not receive federal grants in this period.

4.4 Distance Data

[22] The years used in calculating average federal support were 1983, 1984, and 1985.
[23] These data are from the dataset associated with the National Research Council's publication *Research-Doctorate Programs in the United States: Continuity and Change* (http://books.nap.edu/catalog/5305.html).

We generate distance data as follows. First, we search the Internet for the official website of each university in our dataset to establish the primary location (city, state) of its research campus. Then, we obtain latitude and longitude measures from the US Geological Survey based on the city-state data.[24] Finally, we determine the distance between each university pair by employing the great circle method to calculate the distance in kilometers (km) between the two sets of geographic coordinates.[25]

4.5 Descriptive Statistics

This section examines the basic distributional properties of our key measures: 1) research publications, 2) Bitnet adoption, 3) institution quality, and 4) distance between potential collaborators. Table 1a presents descriptive statistics with institution-years as the unit of analysis, and Table 1b presents descriptive statistics with institution-pair-years as the unit of analysis.

The average university in our dataset publishes 2.51 papers per year in our specified set of journals over the particular time period under investigation (Table 1a). Of these, 1.36 are multi-institutional and 1.15 are single-institutional. In other words, on average these institutions publish approximately 18% more multi-institutional papers per year. Overall, 37% of the institution-year observations have at least one multi-institutional paper.

With respect to Bitnet adoption, while the first three institutions connect to Bitnet in 1981, the average university in the dataset, conditional on being connected before or during 1990, is not connected until halfway through 1985. Figure 1 illustrates the significant variation in adoption years across universities, which is central to our identification strategy.

[24] The US Geological Survey can be accessed at http://geonames.usgs.gov/, and a web query application exists at http://geonames.usgs.gov/pls/gnis/web_query.gnis_web_query_form.
[25] The great circle formula used is: acos(cos(lat1)*cos(long1)*cos(lat2)*cos(long2)
+ cos(lat1)*sin(long1)*cos(lat2)*sin(long2) + sin(lat1)*sin(lat2))*earthRadius.

The first hint of a Bitnet effect is reflected in the comparison of multi-institutional papers per year by institutions with and without Bitnet: 2.47 and 0.92, respectively. However, this difference could be caused by other factors such as time (multi-institution papers are more likely in later years, perhaps for reasons that are not related to Bitnet adoption) and quality (higher-quality institutions may be more likely to publish a greater number of multi-institutional papers and also may be more likely to adopt Bitnet earlier).

Next, we compare average paper production by institutions in the three Carnegie research quality categories. Overall, the research output from these institutions corresponds with what one would expect. CT1 institutions produce more than six times as many multi-institutional papers as CT2, and CT2 institutions produce almost 1.5 times as many multi-institution papers as CT3. The ratios for single-institution papers are similar.

Table 1b examines the basic properties of our institution-pairs. On average, only 0.107% of institution-pairs collaborate per year. The average institution-pair is separated by a distance of 1742 km and produces 0.00132 multi-institution papers per year. Again, we see suggestive evidence of a Bitnet effect since university-pairs generate, on average, more than four times as many multi-institution papers if both are connected to Bitnet. However, for the same reasons as described above, other factors such as time and quality could be confounding this relationship.

We now move on to a more rigorous examination of the effect of Bitnet. That is, following up on the suggestive evidence presented in the descriptive statistics, we seek, after controlling for likely confounding effects, to estimate the extent to which Bitnet adoption facilitates collaboration and publication.

5. Empirical Strategy

Our estimation strategy is based on difference-in-differences identification. Using the paired institution data, we examine how collaboration between institution-pairs that both adopt Bitnet changes relative to other institution-pairs in which one or both have not adopted. We observe the number of collaborative multi-institution publications at the institution and the institution-pair levels.

We are interested in Bitnet as a communications technology. Since Bitnet is only effective if both institutions have the technology, we focus on the effect of both institutions adopting Bitnet on collaboration. We interpret evidence that just one institution adopting Bitnet does not lead to significantly more collaboration as support for our main findings. We label the first institution in the pair *i*, the second institution *j*, and the year *t*.

We run linear regressions on the data using the following equation:

(1) $$Collaboration_{ijt} = \alpha X_{ijt} + \beta Both\ Have\ Bitnet_{ijt} + \mu_t + \phi_{ij} + \varepsilon_{ijt}$$

where the key explanatory variable, $Both\ Have\ Bitnet_{ijt}$ is a dummy that equals one if both institution *i* and *j* have connected to Bitnet by year *t*.[26] In addition, ϕ_{ij} measures institution-pair fixed effects, μ_t measures year fixed effects, and X_{ijt} is a vector of observable institution-pair-year characteristics. This vector contains our proxy for observed pair quality in year *t*, the total number of single-institution papers published by both universities in that year.[27] In some specifications, it also contains: 1) the distance between institutions interacted with Bitnet adoption, and 2) a dummy indicating whether either of the pair has adopted.

In our main models, we treat $Collaboration_{ijt}$ as a dummy variable for whether institutions *i* and *j* have any collaborations in year *t*. We estimate equation (1) using a fixed effects linear probability (OLS) regression with the fixed effects differenced out using average

[26] We also examine time since Bitnet adoption, the effect of which is most clearly illustrated in Figure 3.
[27] We also show that qualitative results do not change if the product of the single-institution publications at the two universities is used instead of the sum.

values. We treat collaboration as a dummy variable because 85% of all institution-pair-years with at least one collaboration have only one collaboration. We also show results for a fixed effects OLS regression on total number of collaborations, a fixed effects probit regression on only those pairs with at least one publication, a fixed effects negative binomial regression on total number of publications on only those pairs with at least one publication, and a fixed effect zero inflated poisson regression where the number of single-institution publications defines the first stage.[28, 29]

For this linear equation to identify the average effect of Bitnet adoption on collaboration between two given institutions, we implicitly assume that unobserved institution-pair quality can be decomposed into an additively separable fixed component and a time varying component. The time component is constant across institution-pairs (Athey and Stern, 2002). Recall that the vector X_{ijt} does contain a proxy for observed pair quality in year t, the total number of single-institution papers published by both universities in that year. We explore this assumption in a number of robustness checks in Section 6.1.

6. Results

In this section, we examine whether Bitnet adoption influences research collaboration among university scientists. Using a difference-in-difference type identification, Section 6.1 shows that Bitnet adoption facilitates multi-institution collaboration. Section 6.2 demonstrates that only

---

[28] We focus on the linear results for two reasons. First, OLS allows coefficients to be easily compared across models and interpreted. Second, linear regression allows for differencing out the mean fixed effects and using the full data set. We limit the fixed effects probit, negative binomial, and zero-inflated poisson regressions to only those pairs with at least one collaboration in order to overcome computational difficulties in estimating thousands of dummy variables. We also focus on the linear model due to the large number of zeros in our dependent variable. To ensure that it is not the large number of zeros driving our results, we also estimate a zero-inflated poisson regression.

[29] The R^2 for the regressions in this paper are very low. Given the large number of observations, the differencing out of institution fixed effects, and the small number of explanatory variables, we do not feel this is surprising. Our estimates are statistically significant and economically important.

collaborations between medium-ranked (CT2) and top-ranked (CT1) schools increase significantly, and it provides evidence that only medium-ranked schools have a significant increase in productivity as a consequence of Bitnet adoption. Top- and bottom-ranked schools do not experience a significant increase in collaboration or in total publications. Section 6.3 shows that Bitnet especially facilitates collaboration between nearby institutions.

6.1 The Effect of Bitnet on Collaboration across Institutions

In this section, we present evidence that Bitnet adoption facilitates academic collaboration. The linear regression in the first column of Table 2 provides the main result. Collaborative publications increase significantly when both universities in the pair are connected to Bitnet (*Both have Bitnet),* controlling for year and institution-pair fixed effects as well as for within-pair quality changes over time (measured by single-institution publications). This relationship is economically large: The rate of collaboration between institutions increases by approximately 85% if both institutions have connected. In addition, these data indicate that although Bitnet amplifies existing collaborations, it also facilitates new ones. In fact, 81% of the collaborations that occur after Bitnet adoption are between institutions that did not collaborate between 1977 and the year of adoption (1985, on average).

This type of empirical research inevitably draws two types of criticism regarding causality interpretations: 1) spurious correlation of institution-pair collaboration propensity and adoption and 2) endogeneity. Perhaps certain universities made sudden policy shifts to increase their performance, which resulted in both Bitnet adoption and increased research output? Or perhaps certain universities recruited young new faculty who were both keen on electronic networking and had a taste for collaboration? The remainder of this section addresses these concerns.

To fully dispel spurious correlation and endogeneity concerns underlying the relationship between adoption and propensity to collaborate, we would need a strong instrument that is correlated with adoption but not with the propensity to collaborate. We try two potential instruments, and while our main results are robust to these instruments (see Appendix Table A1 columns 1 and 2), the instruments themselves are poor.[30]

In the absence of a strong instrument, we use a variety of methods that, combined, strongly suggest that spurious correlation is not a concern and that adoption facilitates collaboration. First, our main results partially control for observable changes in pair quality over time by including a measure of single-institution publications on the right hand side.

Second, we show that a control for whether only one of the schools in the pair has adopted is not significant and does not change the main result (Table 2 column (2)). We observe that the coefficient on *Both have Bitnet* remains virtually unchanged when we control for whether only one of the members of the pair has connected. We find this a particularly interesting result. There is no increase in the likelihood of collaboration when only one joins the network (the coefficient on *One or more has adopted Bitnet* is not significant), but there is a significant effect when both connect. We interpret this result as suggesting that it is both universities in the pair being connected that drives the increase in collaboration.

[30] In particular, the instruments are distance from the nearest adopter and inclusion on a list of likely early adopters compiled by Ira Fuchs of CUNY in 1981. Distance is meant as a proxy for costs of adoption. The first-stage regression suggests that this is a poor proxy: Distance is *positively* correlated with adoption. Inclusion on Mr. Fuchs' list indicates whether the school received his original invitation letter to join the network, which was at least partly predicated on whether the campus had an IBM mainframe compatible with Bitnet and thus a proxy for set-up costs. This measure does not vary over time and therefore has to be interacted with a time trend to work with fixed effects regressions. Furthermore, this list contains almost exclusively CT1 (top-tier) institutions.

Third, we verify that the measured impact of Bitnet does not begin prior to adoption. To explore this possibility, we substitute the *Both have Bitnet* variable for a dummy variable for each of the five years before adoption and each year after adoption. Figure 3 shows the predicted collaboration rates by year prior to and after adoption. Appendix Table A1 column 3 shows the regression coefficients. There is clearly no increase in collaborations in the years preceding Bitnet adoption. Collaboration rates begin to rise in the year following adoption and then rise substantially two and three years after adoption.[31] They then remain at a higher rate.

Fourth and finally, we conduct a variety of other robustness checks on our institution-pair results to check for misspecification of the variables or functional forms. Column 3 of Table 2 employs a count measure rather than a dummy for our dependent variable. In this specification, both institutions having Bitnet results in a 75% increase in the number of collaborations. Columns 4 through 6 of Table 2 employ non-linear techniques (probit, negative binomial, and zero-inflated poisson) to show the results are not a consequence of the linear specification.[32] In appendix Table A1, we drop the fixed effects and conduct OLS with clustered standard errors by institution-pairs in column 4, include a variation of the quality control measure in column 5,[33] and omit the control measure in column 6. Still, the main result persists.

Regarding endogeneity, we acknowledge that it is possible that individuals influenced their universities to adopt Bitnet precisely because they wanted to collaborate. To address this, we purposefully choose to study electrical engineers rather than computer scientists since it is much less likely that the former were directly involved in adoption decisions. Also, we conduct our analysis such that pair adoption is the key explanatory variable, which is less susceptible to

[31] It is important to note that, unlike in economics, the average publication lag for the engineering journals examined here is less than 6 months.
[32] The zero-inflated poisson regression uses single-institution papers in the first stage.
[33] We employ the product of the count of single-institution papers from each school in the pair rather than the sum.

endogeneity concerns than single-institution adoption. Finally, we rely on the Figure 3 results discussed above showing there was no systematic rise in collaboration prior to adoption, which would be a likely sign of reverse causality.

Despite these efforts, we interpret our results as evidence that Bitnet adoption facilitates, rather than causes, multi-institutional collaboration. Researchers only will collaborate if they want to. Thus, even if one remains concerned that the researchers under investigation in this study did influence their university's decision to adopt Bitnet so they could collaborate, the network clearly succeeded in facilitating this collaboration.

We now turn our attention to how the relationship between Bitnet adoption and collaboration is mediated by institution quality and distance. Are the benefits from Bitnet adoption spread uniformly across all adopters, or do they vary with quality and distance?

6.2 Does the Bitnet Effect Vary with Institution Quality?

We examine the relationship between university research quality and the Bitnet effect more closely in Table 3. Specifically, we examine whether Bitnet serves to further accentuate the concentration of innovation amongst top schools by amplifying advantages through cost reductions or whether, in contrast, the network serves to allow more universities to innovate through gains from specialization. Here, using our institution-pair data, we divide our sample into six groups, reflecting all possible combinations of quality types as measured by the Carnegie Foundation.[34] While the coefficients on *Both Have Bitnet* are positive except for the low-low pairs,[35] only the coefficient on high-medium pairs is significantly positive. For this sub-sample,

[34] In appendix Table A2, we show that the Table 3 results are robust to using two alternative measures of university research quality: 1) the number of publications in the seven journals in our data from 1972 to 1979 (before Bitnet) and 2) the number of federal grants to the electrical engineering departments from 1986 to 1992.
[35] The small number of collaborations among low-low pairs (21) is likely responsible for this counterintuitive result.

both universities in the pair being connected increases the likelihood of collaboration by over 150%.

We next seek to better understand who benefits from these collaborations between high-medium pairs. In Table 4, we present regressions on single-institution level data to provide suggestive evidence that it is medium-ranked institutions that benefit most from collaboration with top-tier institutions. Our data include 289 institutions over 15 years for 4335 total observations.

The first three columns of Table 4 examine whether CT1, CT2, or CT3 schools experience a greater increase in total collaborations after Bitnet adoption. These are OLS regressions of total multi-institution publications on whether the university has adopted Bitnet, the total number of single-institution publications, year fixed effects, and institution fixed effects (differenced out).[36] The results show that only medium-ranked schools collaborate more after adopting Bitnet. While Table 3 showed that collaborations increase between CT1 and CT2 universities, this does not seem to lead to an increase in the overall number of collaborations for CT1 schools. A simple interpretation of this result is that Bitnet adoption led CT1 schools to substitute from collaborations with other CT1 schools and from single-institution publications into collaborations with CT2 schools.

For CT2 schools, however, the collaborations facilitated by Bitnet do lead to an increase in overall collaborative research. In addition, columns (4) through (6) use total publications as the dependent variable. They show that Bitnet adoption is also associated with an increase in

[36] The qualitative results of this table do not change if fixed effect negative binomial regressions are used instead. OLS is used to be consistent with the rest of the paper. For the regressions in Table 5 to identify the relationship between adoption and research production, we assume that unobserved institution quality can be decomposed into an additively separable fixed component and a time varying component. The time component is constant across institutions. This assumption is questionable if Bitnet adoption is associated with an unobserved quality improvement. We cannot tease this out in this section. The robustness checks in section 6.1 show that the overall effect of Bitnet adoption on collaboration is unlikely to be a result of this type of spurious correlation.

total research productivity of the CT2 schools. This is not true of CT1 and CT3 schools. We find these quality-related results striking. CT2 schools increased their total publishing output in top journals after adoption. Descriptive statistics show that while CT2 schools experienced an increase in share of publications over the 1980s, CT1 schools experienced a nearly equivalent decline in their share.[37] Overall, the benefits of Bitnet adoption, measured by an increase in publications, accrue primarily to medium-ranked schools (Table 4) due to collaboration with top-ranked schools (Table 3). The reduction in communication costs associated with Bitnet leads to a broadening of the institutions that participate in high-quality research. Thus, consistent with descriptions in the sociology literature of how university researchers used early electronic networks for collaboration and the increased involvement of individuals at peripheral institutions, our empirical findings indicate that Bitnet led to specialization and gains from trade in the production of knowledge.

6.3 Does the Bitnet Effect Vary with Distance?

Here we examine whether the Bitnet effect varies with distance. As discussed above, if Bitnet adoption is a substitute for face-to-face interaction, we would expect the benefits of adoption to be greatest for universities that are furthest apart. On the other hand, if Bitnet adoption is a complement to face-to-face interaction, we would expect the benefits of adoption to be greatest for universities that are co-located.

To address this issue, we employ a spline regression, grouping together universities that are: (1) within 100km, (2) between 100 and 1000km apart, (3) between 1000 and 3000km apart,

---

[37] CT2 schools' share increased from 5.8% of publications in 1980 and 1981 to 7.8% in 1988 and 1989. CT1 schools' share fell over the same period from 84.0% to 81.8%. However, since the proportion of multi-institutional publications increased over the same period, it is not necessarily the case that the absolute contribution from CTI schools decreased over this time. It is certainly the case though that the absolute contribution as well as the relative contribution from CT2 schools increased.

and (4) further than 3000km apart. Our results (Table 5), which again include year and institution-pair fixed effects as well as a control for within-pair productivity changes over time, suggest that Bitnet adoption is likely a complement to face-to-face interactions. Bitnet adoption has the greatest effect on university-pairs that are co-located (within 100km). We also find that collaboration increases regardless of distance, suggesting that Bitnet does facilitate collaboration, even for distant institutions. Overall, the Bitnet effect is nearly three times greater for co-located universities as it is for those that are in the next category (100-1000km apart). For robustness, we report various groupings using different distance ranges, but the main result persists: Co-located universities benefit most from Bitnet adoption.

Finally, we examine the interaction of quality and distance on the Bitnet effect (Table 6). Overall, our prior results persist: The greatest effect on multi-institutional paper production occurs for medium-ranked universities that collaborate with co-located, top-ranked universities. Indeed, medium-ranked universities also increase their collaboration with non-co-located top-ranked universities, but the effect of Bitnet is almost 10 times greater for those that are co-located.[38]

The results in Table 6 also suggest that low-ranked universities (CT3) benefit from Bitnet adoption through collaboration with co-located top-ranked institutions as well. Interestingly, pairs of top-ranked, co-located universities seem to reduce their level of collaboration upon adopting Bitnet. We speculate that this may be due to substitution for collaboration with medium-ranked schools.

---

[38] Examples of collaborating, co-located, CT1-CT2 pairs include the California Institute of Technology and the University of California (Riverside), Columbia University and Polytechnic University, Harvard University and Northeastern University, University of Pennsylvania and Drexel University, Massachusetts Institute of Technology and Northeastern University, and Stanford University and the University of California (Santa Cruz).

We interpret the findings of this section to indicate that low-cost electronic communication, while a substitute for face-to-face interactions under certain conditions, is also a complement. Researchers communicate with people they know, and they are more likely to know those who work nearby.

7. Conclusions

What general conclusions can we draw from these findings? Exploiting the fact that universities vary in the year in which they adopt Bitnet and that both in a pair need to be connected to benefit from lower communication costs, we discover that, shortly after universities connect, they tend to experience an increase in their publication output, particularly in terms of multi-institution publications. Furthermore, the greatest boost from Bitnet adoption is achieved by medium-tier universities that are co-located with top-tier schools. However, we are unable to comment on whether Bitnet delivered an overall productivity increase due to the nature of our data. We have no data on input, and our output measure – publications from a fixed set of journals – remains reasonably constant over time. To be clear, what we observe is that Bitnet facilitated a change in the relative performance of certain types of university collaborations.

These findings have two primary implications for our understanding of the national innovation system. First, Bitnet facilitated greater participation by researchers at middle-tier U.S. institutions in the production of high-quality research. By lowering the cost of communication and data sharing, Bitnet widened the circle of institutions able to consistently innovate in electrical engineering. It did this by allowing middle-tier schools to collaborate with top-tier schools. We speculate that both types of institutions gained because Bitnet facilitated specialization.

Second, innovation systems have famously geographic features. Marshall (1920) and others (Krugman, 1991; Jaffe *et al*, 1993; Porter, 2003) have discussed and measured the striking concentration of innovative activity. Because of a significant reduction in the cost of communication and file sharing, it would be reasonable to expect that a widespread electronic network would temper the tendency towards agglomeration since pairs that are separated by the greatest distance would experience the greatest reduction in costs. But it seems this view under-appreciates the importance of direct interaction and social relationships in the collaboration process (Merton, 1973). Although Bitnet does facilitate an increase in distant collaboration, our findings indicate that its greatest effect is on pairs of universities that are close together, further accentuating the tendency for innovative activity to be concentrated and highlighting the importance of the social component of the innovation system.

Finally, this seeming importance of the social component leads to the still open question: How do researchers choose their collaborators? Effective research collaboration is predicated on familiarity, common knowledge, and trust (Crane, 1965 and 1969; Merton, 1973). Professor-student and graduate student cohort relationships often have these qualities, in addition to a common interest and expertise in similar research questions. These relationships are usually characterized by close interaction at the same institution followed by institutional and geographic dispersion. As such, we speculate that these types of relationships may play a very important role in influencing multi-institutional collaborative patterns. To the extent that this is true, Bitnet may have had its greatest effect in facilitating these types of collaborations across institution types and distance. If so, the advent of Bitnet increased the role recruitment patterns play in the architecture of innovation systems and also the sensitivity to shifts in these patterns on the structure by and geography in which research is performed. In addition, this suggests that in other national settings where graduate student dispersion patterns or the distribution of capital-intensive research equipment are markedly different, the introduction of electronic networks may not have the same effect. Alas, our data only set the stage for asking such questions but offer little for addressing them. We leave these puzzles for future research.

References

Agrawal, Ajay, Iain Cockburn, and John McHale (2003) "Gone But Not Forgotten: Labor Flows, Knowledge Spillovers, and Enduring Social Capital," National Bureau of Economic Research Working Paper #9950.

Audrestch, David B., and Maryann P. Feldman (1996) "R&D Flows and the Geography of Innovation and Production," *American Economic Review*, Vol. 86(3), 630-640.

Barnett, Andy H., Richard W. Ault, and David L. Kaserman (1988) "The Rising Incidence of Co-authorship in Economics: Further Evidence," *Review of Economics and Statistics,* 70(3), 539-543.

Beaver, D. and R. Rosen (1978) "Studies in Scientific Collaboration. Part 1. The Professional Origins of Scientific Co-authorship," *Scientometrics*, 1, 65-84.

Beaver, D. and R. Rosen (1979) "Studies in Scientific Collaboration. Part 2. Scientific Co-authorship, Research Productivity and Visibility in the French Scientific Elite," *Scientometrics,* 1, 133-49.

Cockburn, I., and R. Henderson (1998) "Absorptive Capacity, Coauthoring Behavior, and the Organization of Research in Drug Discovery," *Journal of Industrial Economics*, Vol. XLVI, No. 2, pp. 157-182.

Crane, D. (1965) "Scientists in Major and Minor Universities: A Study of Productivity and Recognition," *American Sociological Review*, XXX, 699-714.

Crane, D. (1969) "Social Structure in a Group of Scientists: A Test of the 'Invisible College' Hypothesis," *American Sociological Review*, XXXIV, 335-352.

Gaspar, J. and E. Glaeser (1998) "Information Technology and the Future of Cities," *Journal of Urban Economics*, 48(1), 136-156.

Gurbaxani, V. (1990) "Diffusion in Computing Networks: The Case of BITNET," *Communications of the ACM,* 33, 65-75.

Hamermesh, D.S., and S.M. Oster (1998) "Tools or Toys? The Impact of High Technology on Scholarly Productivity," National Bureau of Economic Research Working Paper #6761.

Hesse, B.W., L.S. Sproull, S.B. Kiesler, and J.P. Walsh (1993) "Returns to Science: Computer Networks in Oceanography," *Communications of the ACM*, 36, 8, 90-101.

Jaffe, A.B., M. Trajtenberg, and R. Henderson (1993) "Geographic Localization of Knowledge Flows as Evidenced by Patent Citations," *Quarterly Journal of Economics,* 108, 577-598.

Krugman, P. (1991) Geography and Trade, The MIT Press, Cambridge, MA.

Mairesse J. and L. Turner (2005) "Measurement and Explanation of the Intensity of Co-publication in Scientific Research: An Analysis at the Laboratory Level," in New Frontiers in the Economics of Innovation and New Technology: Essays in Honor of Paul David, eds. C. Antonelli, D. Foray, B. Hall, and E. Steinmueller, Edward Elgar Publishing.

Marshall, A. (1920) Principles of Economics, Macmillan, London, UK.

Merton, R.K. (1973) The Sociology of Science: Theoretical and Empirical Investigations, The University of Chicago Press, Chicago, IL.

Nelson, R. and N. Rosenberg (1993) "Technical Innovation and National Systems," in National Innovation Systems, ed. R. Nelson, Oxford University Press, Oxford, UK.

Olson, G.M., and J.S. Olson (2003) "Mitigating the Effects of Distance on Collaborative Intellectual Work," *Economics of Innovation and New Technology*, Vol. 12(1), pp. 27-42.

Porter, M. E. (2003) "The Economic Performance of Regions," *Regional Studies*, 37(6/7), 549-78.

Romer, P. M. (1986) "Increasing Returns and Long-Run Growth," *Journal of Political Economy*, Vol. 94, pp. 1001-37.

Romer, P. M. (1990) "Endogenous Technological Change," *Journal of Political Economy*, Vol. 98, supplement to No. 5, pp. 71-102.

Singh, J. (2005) "Collaborative Networks as Determinants of Knowledge Diffusion Patterns," *Management Science*, 51(5): 756-770.

Smith, M.A. (1999) "Invisible Crowds in Cyberspace: Mapping the Social Structure of the Usenet" in Communities in Cyberspace: Perspectives on New Forms of Social Organization, Routledge Press, London, UK.

Thompson, P. and M. Fox-Kean (2005) "Patent Citations and the Geography of Knowledge Spillovers: A Reassessment," *American Economic Review*, 95(1), 450-460.

Van Alstyne, M. and E. Brynjolfsson (2005) "Global Village or Cyber-Balkans? Modeling and Measuring the Integration of Electronic Communities," *Management Science*, 51, 851-868.

Walsh, J. and T. Bayma (1996) "The Virtual College: Computer-Mediated Communication and Scientific Work," *The Information Society*, 12, 343-363.

Wellman, B. and M. Gulia (1997) "Net Surfers Don't Ride Alone: Virtual Communities as Communities" in Communities in Cyberspace, eds. P. Kollock and M. Smith, University of California Press, Berkeley, CA.

Zucker, L., M. Darby, and M.B. Brewer (1998) "Intellectual Capital and the Birth of U.S. Biotechnology Enterprises," *American Economic Review*, 88, 290-306.

Table 1a: Descriptive Statistics—Institution-Level

Variable (by year)	Mean	Standard deviation	Minimum	Maximum	# of observations
Total papers	2.508	6.439	0	131	4,335
Multi-institution papers	1.360	3.147	0	39	4,335
Single-institution papers	1.149	3.739	0	92	4,335
Any multi-institution papers (dummy)	0.374	0.484	0	1	4,335
Average year adopting Bitnet*	1985.6	2.054	1981	1990	225
Has Bitnet	0.282	0.450	0	1	4,335
Total papers if do not have Bitnet	1.862	5.724	0	106	3,114
Total papers if have Bitnet	4.160	7.738	0	131	1,221
Multi-institution papers if do not have Bitnet	0.923	2.521	0	31	3,114
Multi-institution papers if have Bitnet	2.473	4.153	0	39	1,221
CMU Type 1					
Multi-institution papers	3.062	4.537	0	39	1,575
Single-institution papers	2.696	5.697	0	92	1,575
CMU Type 2					
Multi-institution papers	0.494	1.046	0	9	945
Single-institution papers	0.368	0.837	0	9	945
CMU Type 3					
Multi-institution papers	0.333	1.131	0	17	1,815
Single-institution papers	0.212	1.263	0	9	1,815

*Conditional on adopting Bitnet by the end of 1990

Table 1b: Descriptive Statistics—Institution-Pairs

	Mean	Standard deviation	Minimum	Maximum	# of observations
# collaborative papers between the pair	0.00132	0.0466	0	6	624,240
Dummy if any collaborative papers that year	0.00107	0.0328	0	1	624,240
# collaborative papers if at least one has not adopted Bitnet	0.000852	0.0379	0	6	516,635
# collaborative papers if both have adopted Bitnet	0.00360	0.0756	0	5	107,605
Distance	1,742.732	1,281.485	0	8,293.748	624,240
Sum of # of single-institution papers produced by the pair	2.297	5.295	0	122	624,240
Product of # of single-institution papers produced by the pair	1.362	15.250	0	3,496	624,240
Dummy if at least one of the pair has adopted Bitnet	0.391	0.488	0	1	624,240
Dummy if both institutions have adopted Bitnet	0.172	0.378	0	1	624,240

Table 2: Bitnet Adoption and Collaboration Using Institution-Pairs

	(1)	(2)	(3)	(4)	(5)	(6)
	Main specification: Linear regression with a dummy for any collaboration as the dependent variable	Includes variable if just one institution has adopted	Dependent variable is the total # of collaborations	Probit regression with only pairs with at least one collaboration	Negative binomial regression on total # of collaborations using only pairs with at least one collaboration	Zero inflated poisson regression on total # of collaborations using only pairs with at least one collaboration
Both have Bitnet	0.000917	0.000918	0.00102	0.270	0.409	0.460
	(0.000156)**	(0.000158)**	(0.000221)**	(0.113)*	(0.219)+	(0.255)+
One or more has adopted Bitnet		-0.00000500				
		(0.0001645)				
Sum of # of single-institution papers	0.0000544	0.0000544	0.0000399	0.00777	0.0190	-0.0242
	(0.0000128)**	(0.0000128)**	(0.0000181)*	(0.00425)+	(0.00829)**	(0.0253)
# of Observations	624,240	624,240	624,240	6,930	6,930	6,930
# of Groups	41,616	41,616	41,616	462	462	462
R^2 (overall)	0.002	0.002	0.001	N/A	N/A	N/A
Log Likelihood	N/A	N/A	N/A	-1,251.76	-1,373.50	-1,369.15

Unless otherwise specified, regressions include year and institution-pair fixed effects
Standard errors in parentheses
**significant at 99% level; *significant at 95% level; +significant at 90% level

Table 3: Bitnet Adoption, Collaboration, and Institution-Pair Quality

	(1)	(2)	(3)	(4)	(5)	(6)
	CT1 and CT1	CT1 and CT2	CT1 and CT3	CT2 and CT2	CT2 and CT3	CT3 and CT3
Both have Bitnet	0.000809	0.00149	0.000368	0.000620	0.000171	-0.000463
	(0.00102)	(0.000380)**	(0.000241)	(0.000391)	(0.000155)	(0.000213)*
Sum of # of single-institution papers	0.000214	-0.0000617	0.00000380	0.000214	0.00000630	-0.0000335
	(0.0000486)**	(0.0000270)*	(0.0000163)	(0.0000990)*	(0.0000282)	(0.0000273)
# of Observations	81,900	99,225	190,575	29,295	114,345	108,900
# of Groups	5,460	6,615	12,705	1,953	7,623	7,260
R^2 (overall)	0.006	0.001	0.001	0.001	0.001	0.001

Regressions include year and institution-pair fixed effects
Standard errors in parentheses
CT1, CT2, and CT3 define the Carnegie Foundation's rankings of research focus.
**significant at 99% level; *significant at 95% level; +significant at 90% level

Table 4: Bitnet Adoption, Total Publications, and Multi-Institution Publications in the Single-Institution Data

	Dependent Variable is Multi-Institution Publications (Collaborations)			Dependent Variable is Total Publications		
	(1)	(2)	(3)	(4)	(5)	(6)
	CT1	CT2	CT3	CT1	CT2	CT3
Has Bitnet	0.280	0.334	0.0710	-0.0745	0.487	0.206
	(0.228)	(0.110)**	(0.0633)	(0.520)	(0.156)**	(0.145)
Single-institution	0.219	0.170	0.240			
papers	(0.0156)**	(0.0395)**	(0.0852)**			
# of Observations	1,560	915	1,560	1,575	945	1,815
# of Groups	104	61	104	105	63	121
R^2 (overall)	0.519	0.129	0.341	0.011	0.036	0.013

Regressions include year and institution fixed effects
Standard errors in parentheses
**significant at 99% level; *significant at 95% level; +significant at 90% level

Table 5: Bitnet Adoption, Collaboration, and Institution-Pair Distance

	(1)	(2)	(3)	(4)
	Main specification	Alternative distance (1)	Alternative distance (2)	Main specification with total # of collaborations as the dependent variable
Distance is under 100 km and Both Adopted Bitnet	0.00203	0.00203	0.00203	0.00236
	(0.000867)*	(0.000867)*	(0.000867)*	(0.00123)+
Distance is between 100 km and 1000 km and Both Adopted Bitnet	0.000742		0.000743	0.000533
	(0.000225)**		(0.000225)**	(0.000319)+
Distance is between 1000 km and 3000 km and Both Adopted Bitnet	0.000754			0.00112
	(0.000201)**			(0.000284)**
Distance is over 3000 km and Both Adopted Bitnet	0.00155			0.00154
	(0.000293)**			(0.000415)**
Distance is between 100 km and 500 km and Both Adopted Bitnet		0.000998		
		(0.000327)**		
Distance is between 500 km and 1000 km and Both Adopted Bitnet		0.000559		
		(0.000282)*		
Distance is over 1000 km and Both Adopted Bitnet		0.000977	0.000977	
		(0.000179)**	(0.000179)**	
Sum of # of single-institution papers	0.0000555	0.0000547	0.0000546	0.0000409
	(0.0000128)**	(0.0000128)**	(0.0000128)**	(0.0000181)*
# of Observations	624,240	624,240	624,240	624,240
# of Groups	41,616	41,616	41,616	41,616
R^2 (overall)	0.003	0.003	0.003	0.001

Regressions include year and institution-pair fixed effects
Standard errors in parentheses
**significant at 99% level; *significant at 95% level; +significant at 90% level

Table 6: Bitnet Adoption, Collaboration, Institution-Pair Quality, and Distance

	(1)	(2)	(3)	(4)	(5)	(6)
	CT1 and CT1	CT1 and CT2	CT1 and CT3	CT2 and CT2	CT2 and CT3	CT3 and CT3
Distance is under 100 km and Both Adopted Bitnet	-0.0185585	0.0192371	0.0047773	-0.0001092	0.0017877	-0.0019598
	(0.0040994)**	(0.0017607)**	(0.0012771)**	(0.0020295)	(0.0009123)+	(0.0012405)
Distance is between 100 km and 1000 km and Both Adopted Bitnet	-0.0002927	0.0017980	0.0004862	0.0011414	-0.0002585	-0.0003163
	(0.0012254)	(0.0004926)**	(0.0003517)	(0.0005452)*	(0.0002358)	(0.0003274)
Distance is between 1000 km and 3000 km and Both Adopted Bitnet	0.0017114	0.0001902	0.0001236	0.0005301	0.0003123	-0.0005181
	(0.0011632)	(0.0004408)	(0.0003262)	(0.0004502)	(0.0002021)	(0.0003200)
Distance is over 3000 km and Both Adopted Bitnet	0.0027329	0.0030993	0.0001980	-0.0001052	0.0004064	-0.0004536
	(0.0014988)+	(0.0006476)**	(0.0004367)	(0.0007454)	(0.0003035)	(0.0003971)
Sum of # of single-institution papers	0.0002157	-0.0000576	0.0000042	0.0002141	0.0000058	-0.0000334
	(0.0000487)**	(0.0000270)*	(0.0000163)	(0.0000991)*	(0.0000282)	(0.0000273)
# of Observations	81,900	99,225	190,575	29,295	114,345	108,900
# of Groups	5,460	6,615	12,705	1,953	7,623	7,260
R^2 (overall)	0.002	0.004	0.001	0.001	0.001	0.001

Regressions include year and institution-pair fixed effects.
Standard errors in parentheses
CT1, CT2, and CT3 define the Carnegie Foundation's rankings of research focus.
**significant at 99% level; *significant at 95% level; +significant at 90% level

Figure 1
Cumulative Bitnet Adoption over Time

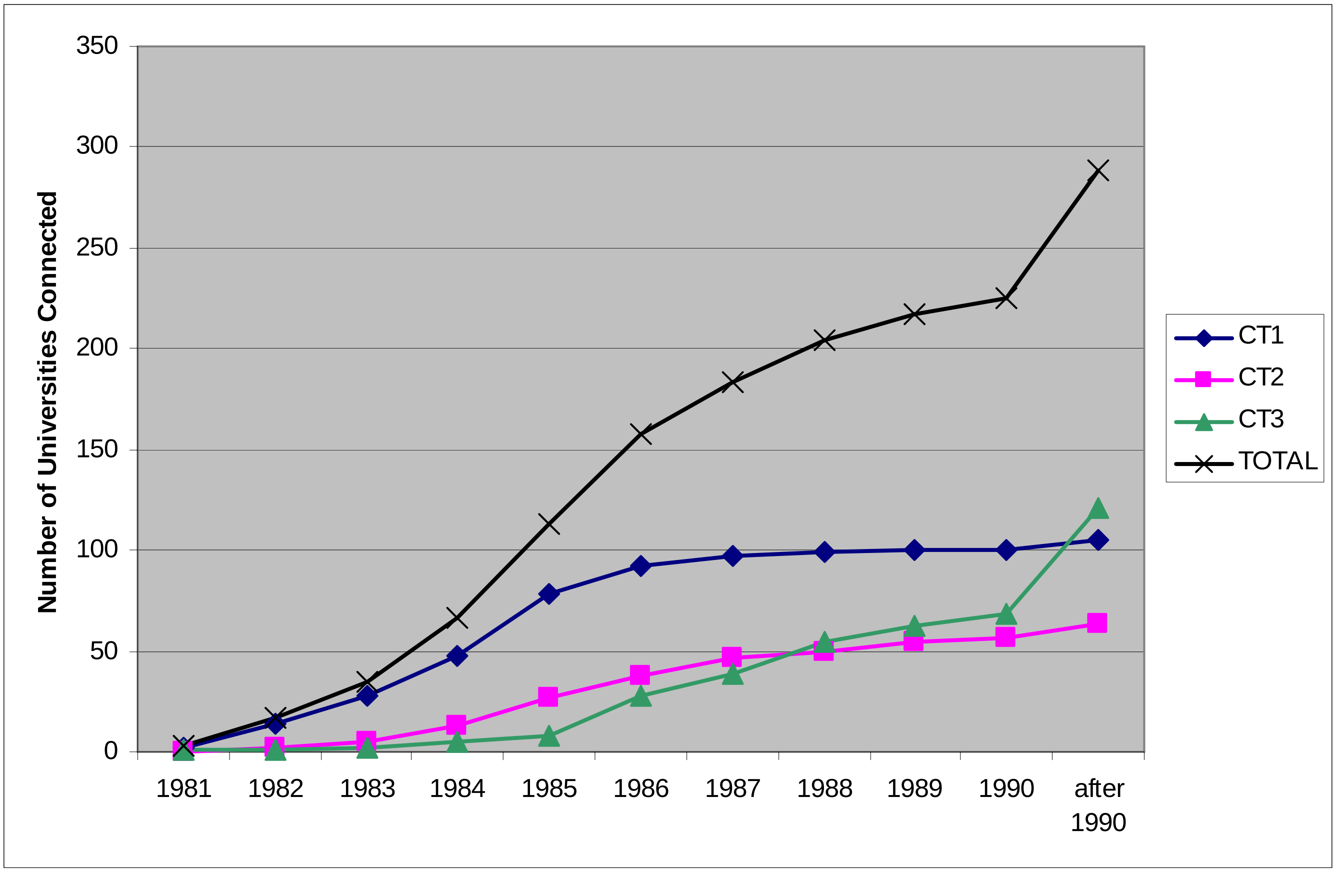


Figure 2
Bitnet Connections – 1981, 1982, and 1983

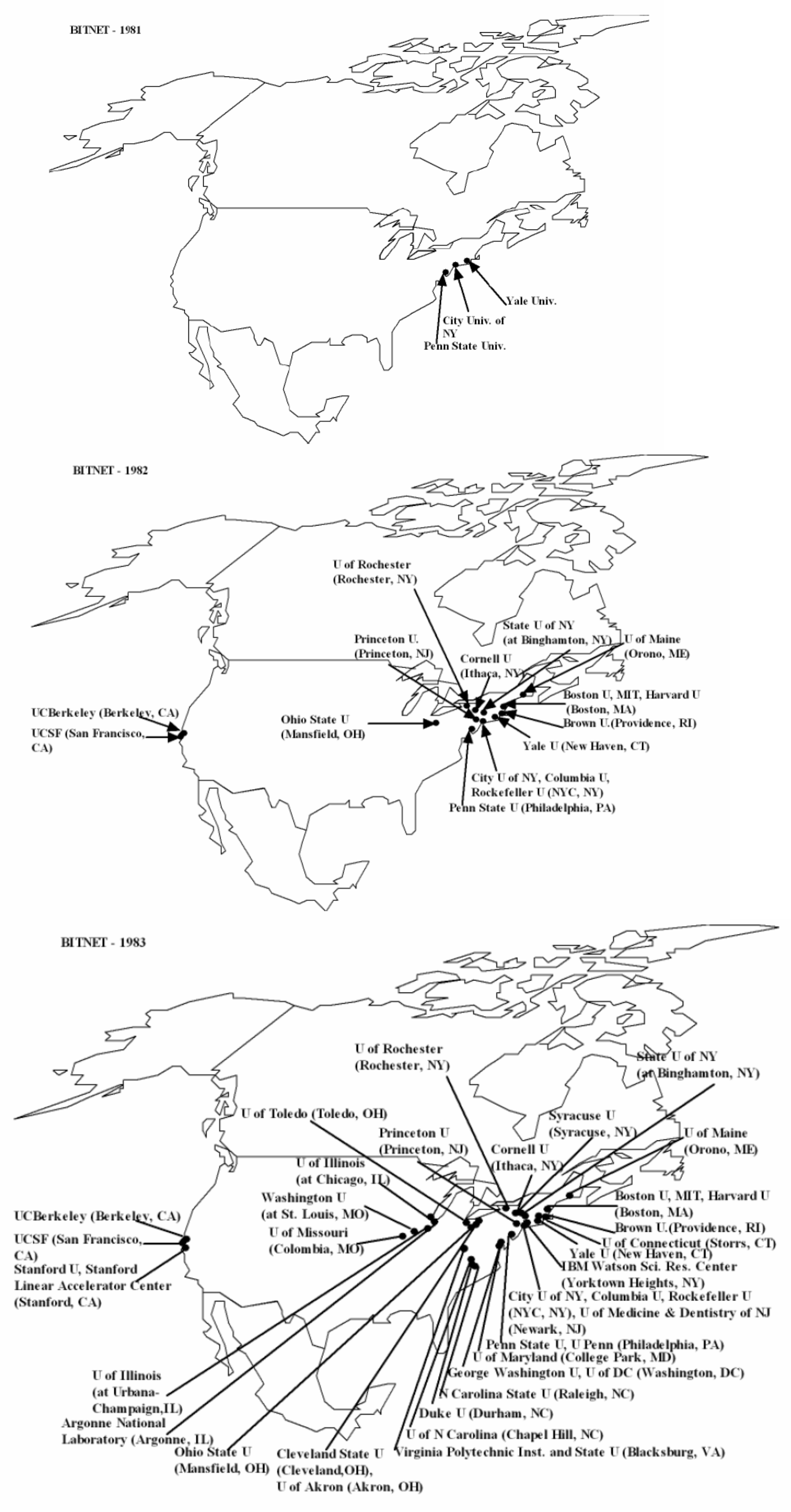
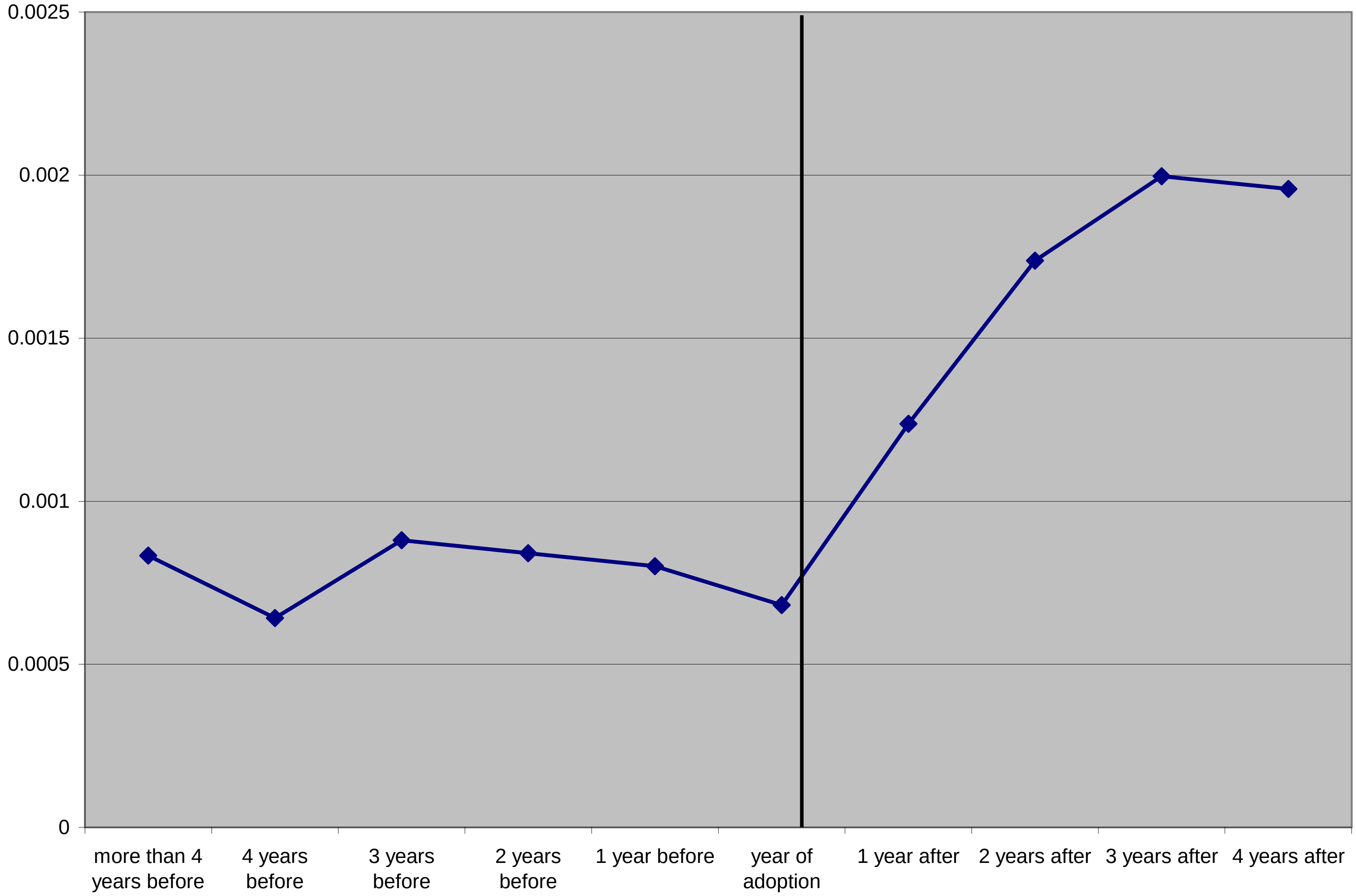


Figure 3: Predicted Collaboration Rates by Year Before and After Adoption



Appendix: Robustness Tables

Appendix Table A1: Robustness of the main (Table 2) effect to alternative specifications

	(1)	(2)	(3)	(4)	(5)	(6)
	Instrumental Variables: Fuchs' List, Distance from Adopter[&&]	Instrumental Variables using Fuchs' List[&&]	Years before and after adoption dummies (Figure 3)	No fixed effects linear regression; errors clustered by institution-pair	Alternative control for pair productivity	No control for pair productivity
Both have Bitnet	0.00248	0.00231		0.00247	0.000921	0.000907
	(0.00151)+	(0.00152)		(0.000335)**	(0.000156)**	(0.000156)**
One or more has adopted Bitnet						
Sum of # of single-institution papers	1.58E-05	1.58E-05	0.0000579	0.000430		
	(2.10E-05)	(2.10E-05)	(0.0000128)**	(0.0000542)**		
Product of # of single-institution papers					1.54E-05	
					(3.40E-07)**	
5 Years Before Adoption[&]			-0.000304			
			(0.000229)			
4 Years Before Adoption[&]			-0.000412			
			(0.000235)+			
3 Years Before Adoption[&]			-0.000206			
			(0.000241)			
2 Years Before Adoption[&]			-0.000271			
			(0.000246)			
1 Year Before Adoption[&]			-0.000312			
			(0.000250)			
Actual Year of Adoption[&]			-0.000490			
			(0.000253)+			
1 Year After Adoption[&]			0.000100			
			(0.000256)			
2 Years After Adoption[&]			0.000612			
			(0.000268)*			
3 Years After Adoption[&]			0.000827			
			(0.000286)**			
4 Years After Adoption[&]			0.000862			
			(0.000316)**			
5 Years After Adoption[&]			0.00134			
			(0.000360)**			
6 Years After Adoption[&]			0.00299			
			(0.000470)**			
7 Years After Adoption[&]			0.00521			
			(0.000744)**			
8 Years After Adoption[&]			0.00240			
			(0.00136)+			
9 Years After Adoption[&]			0.00891			
			(0.00279)**			
10 Years After Adoption[&]			0.00151			
			(0.0187)			
# of Observations	416,160	416,160	624,240	624,240	624,240	624,240
# of Groups	41,616	41,616	41,616	N/A	41,616	41,616
R^2 (Overall)	0.001	0.001	0.003	0.004	0.003	0.001

Unless otherwise specified, regressions include year and institution-pair fixed effects; standard errors in parentheses

**significant at 99% level; *significant at 95% level; +significant at 90% level

[&]Base is more than 5 years before

[&&]Instrumental variable regressions use 1982 to 1991 because year dummies perfectly predict adoption from 1977 to 1981.

Appendix Table A2: Robustness of the quality results (Table 3) to different quality definitions

		(1)	(2)	(3)	(4)	(5)	(6)
	Quality Definition	Tier 1 and Tier 1	Tier 1 and Tier 2	Tier 1 and Tier 3	Tier 2 and Tier 2	Tier 2 and Tier 3	Tier 3 and Tier 3
Coefficient on "Both have Bitnet"	IEEE publications in 7 journals 1972 to 1979	-0.00331	0.00116	0.000477	0.000889	0.000735	0.000164
		(0.00214)	(0.000615)+	(0.000201)*	(0.000507)+	(0.000154)**	(0.000094)+
	# of Groups	2,775	5,250	10,800	2,415	10,080	10,296
	Federal grants to electrical engineering depts. 1986 to 1992	-0.00383	0.00363	0.000775	0.000669	-0.000118	0.00000960
		(0.00296)	(0.000999)**	(0.000317)*	(0.000852)	(0.000210)	(0.000132)
	# of Groups	1,653	3,191	10,208	1,485	9,681	15,400

Regressions include year and institution-pair fixed effects. # Observations is 15 times the number of groups (there are 15 years of data)
Standard errors in parentheses
**significant at 99% level; *significant at 95% level; +significant at 90% level
For IEEE publications, Tier 1 includes the top quartile, Tier 2 includes the second quartile, and Tier 3 includes all schools below the median.
For federal grants, Tier 1 includes all schools above the median of those that received grants, Tier 2 includes all schools below the median who received grants, and Tier 3 includes all schools that did not receive a grant.