Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

From Wires to Partners: How the Internet Has Fostered R&D Collaborations Within Firms

Chris Forman, Nicolas van Zeebroeck,

MANAGEMENT SCIENCE
Vol. 58, No. 8, August 2012, pp. 1549–1568
ISSN 0025-1909 (print) | ISSN 1526-5501 (online)

informs

http://dx.doi.org/10.1287/mnsc.1110.1505

From Wires to Partners: How the Internet Has Fostered R&D Collaborations Within Firms

Chris Forman
College of Management, Georgia Institute of Technology, Atlanta, Georgia 30308, chris.forman@mgt.gatech.edu

Nicolas van Zeebroeck
European Center for Advanced Research in Economics and Statistics (ECARES), Solvay Brussels School of Economics and Management, Université libre de Bruxelles, 1050 Brussels, Belgium, nicolas.van.zeebroeck@ulb.ac.be

How did the diffusion of the Internet influence research collaborations within firms? We examine the relationship between business use of basic Internet technology and the size and geographic composition of industrial research teams between 1992 and 1998. We find robust empirical evidence that basic Internet adoption is associated with an increased likelihood of collaborative patents from geographically dispersed teams. On the contrary, we find no evidence of such a link between Internet adoption and within-location collaborative patents, nor do we find any evidence of a relationship between basic Internet and single-inventor patents. We interpret these results as evidence that adoption of basic Internet significantly reduced the coordination costs of research teams, but find little evidence that a drop in the costs of shared resource access significantly improved research productivity.

Key words: R&D organization; geography of innovation; Internet adoption; IT investments; collaborative work
History: Received January 31, 2011; accepted December 6, 2011, by Lee Fleming, entrepreneurship and innovation. Published online in *Articles in Advance* April 6, 2012.

1. Introduction

An increasing fraction of scientific research is no longer done by individual inventors but by collaborative research teams (e.g., Adams et al. 2005, Jones 2009, Wuchty et al. 2007). This shift toward collaborative research is thought to be caused in part by increasing incentives for researchers to specialize, because of the increasing knowledge burden faced by scientists as knowledge accumulates over time (Jones 2009). More broadly, increased specialization and division of labor among researchers may improve research productivity independent of the human capital investments of researchers.

Historically, collaborative work has been hampered by the existence of significant coordination costs that increase with team size, geographic dispersion, and heterogeneity of team composition (e.g., Becker and Murphy 1992). It is widely believed that by lowering these coordination costs, adoption of information technology (IT) such as the Internet may increase the returns to collaborative work (e.g., Cairncross 1997, Friedman 2005). However, although a small body of recent research has examined the implications of IT investment for collaborative academic research (Agrawal and Goldfarb 2008, Ding et al. 2010), to our knowledge there has been little systematic empirical work on the implications of IT investment for industrial research. This is a surprising gap in understanding. Collaborative research has not only been shown to be increasing in frequency but has also been shown to be more highly cited (e.g., Sauer 1988). Furthermore, because collaborative ties are known to increase the likelihood of knowledge flows (e.g., Singh 2005, Fleming et al. 2007), changes in collaboration patterns have important implications for the diffusion of knowledge within firms.

In this paper we take a first step toward empirically evaluating how IT investments shape research collaborations within firms.[1] We motivate our hypotheses using prior models of team-based knowledge work, in particular the models of Becker and Murphy (1992) and Adams et al. (2005) that view optimal team size as a trade-off between the benefits of specialization and division of labor versus increased coordination costs. We use these models to motivate a set of hypotheses about how a decline in coordination costs will lead to an increase in the incidence of collaborative research.

[1] Our focus on within-firm collaborations is particularly appealing to directly observe the coordination benefit of Internet, but it is also motivated by a data constraint. Cross-firm collaborations are measured using patent assignments and are extremely noisy. Collaborating firms would be coassigned their joint patents only if they agree to share their ownership, which is only one of many possibilities to compensate each party for its contribution to an invention. In contrast, by law, all inventors have to be listed on the front page of their patents.

To test these hypotheses, we focus on the role of investments by firms in a set of Internet technologies that lower communications costs. Our analysis focuses on basic Internet connectivity. Prominent examples of basic Internet include Internet access or an internal intranet. The set of technologies we examine require little adaptation or coinvention (Bresnahan and Greenstein 1996) to be used successfully, and so allow us to focus on the short-run changes to collaboration patterns that are made in response to a decline in communication costs. The main hypothesis of this paper is that *by reducing the coordination costs of collaborative work, investments in IT will be associated with an increase in the likelihood of geographically dispersed, multi-inventor collaborative research teams relative to other types of research teams (including output from single inventors).*

Our first set of results assumes that basic Internet adoption is exogenous to research collaborations. We examine collaborations within pairs of heterogeneous geographically distant firm locations, where coordination costs are likely to be greatest. Our results show that when two locations within a firm both adopt basic Internet technology, the likelihood of a collaborative patent invented by researchers in both locations increases significantly compared to an otherwise identical pair without basic Internet. We find that these results remain robust to numerous specifications and changes to controls.

We address the assumption that Internet adoption is exogenous. We first utilize the timing of Internet adoption as the source of a falsification exercise. We find no evidence that the incidence of cross-location research collaborations is correlated with a location pair's future adoption of Internet technology; that is, location pairs who adopt Internet technology experience no increase in the likelihood of a collaborative patent prior to adoption. We then demonstrate that our results are robust to the use of instrumental variables. We employ two sets of instruments that capture local variance in the costs to adopting Internet technology. The first addresses cross-sectional differences in local regulatory conditions that will shape the costs of purchasing Internet access. The second captures cross-sectional differences in familiarity and expertise with the Internet in the local regions where the establishments reside.

Last, we examine the implications of basic Internet adoption for two other types of research groups: multi-inventor collaborations with collocated inventors and single inventors. We find that adoption of basic Internet has no impact on the likelihood of a collaborative patent among researchers within a single firm location, nor does it have any impact on the likelihood of a patent by lone inventors. Together, we interpret this as evidence that, by lowering coordination costs, basic Internet has increased the productivity of larger, geographically dispersed research teams relative to other types of research collaborations. Although basic Internet technology may have increased researcher productivity in other ways—for example, by lowering access costs to shared resources—we find no evidence that these potential benefits resulted in an increase in the incidence of patenting among research teams (including lone inventors) where ex ante coordination costs were low.

Our research contributes to a better understanding of the costs and benefits of scientific research collaborations, and in particular the implications of the diffusion of IT for collaborative work. Some of our findings differ significantly from those of prior work on the implications of IT investment for academic research collaborations.[2] In particular, one paper related to ours is by Agrawal and Goldfarb (2008), who show that adoption of BITNET (a U.S. cooperative university computer network founded in 1981) facilitated cross-institution collaboration in the academe, particularly among researchers in the same geographic region. In contrast, we examine a different setting—industrial research collaborations—and find that adoption of basic Internet was associated with a disproportionate increase in cross-location collaborations, with little effect on within-location collaborations. We speculate that these results are due to differences in the way firm and academic research collaborations are formed, the nature of scientific and industrial research activities, and the functionalities of the two kinds of IT considered: BITNET and Internet.

More broadly, although our analyses examine collaborations among researchers in locations within the United States, our results speak to research on the benefits and costs of geographically dispersed collaborations that has usually been conducted on samples of multinational companies. As is well known, although geographically dispersed research organizations may be effective at assimilating local knowledge from outside of the firm (e.g., Kogut and Zander 1993, Frost et al. 2002), cross-regional transfer of knowledge is difficult and costly even within the boundaries of the firm (Teece 1977, Singh 2008, Sorenson et al. 2006). As a result, the evidence on whether geographic dispersion improves a firm's innovative capabilities remains mixed (e.g., Furman et al. 2005, Leiponen and Helfat 2011). However, it is well known that collaborative work is a powerful enabler of knowledge transfers (e.g., Singh 2005, Fleming et al. 2007). By suggesting a beneficial effect of Internet adoption on distant collaborations, our paper is therefore in the spirit of recent

[2] For examples of this work, see Agrawal and Goldfarb (2008), Ding et al. (2010), Rosenblat and Mobius (2004), Winkler et al. (2010), and Walsh and Bayma (1996). For an example of work that examines theoretically the role IT can play in linking dispersed communities, see van Alstyne and Brynjolfsson (2005).

work that has examined the implications of the use of coordinating mechanisms within firms to facilitate integration of knowledge across units (e.g., Singh 2008, Argyres and Silverman 2004). More broadly, there is increasing interest in measuring whether IT investments have in fact facilitated increasing dispersion of innovative activity (e.g., Macher and Mowery 2008). However, as yet there is little evidence on the link between IT investments and the organization of research activity within firms. This paper takes a first step toward presenting such evidence.

2. Research Framework

In this section we present a simple framework that will show how a reduction in coordination costs enabled by investment in IT will lead to increases in the productivity of geographically dispersed research teams relative to other types of research collaborations. This in turn will motivate a set of predictions on how adoption of basic Internet will be associated with a change in the likelihood of collaboration among different types of research teams.

Our focus on the likelihood of collaboration rather than a direct test of the productivity of different types of research collaborations reflects two types of data constraints. First, we do not possess direct measures of some of the inputs into the innovation production functions of these different types of research collaborations: project-level data on research and development (R&D) expenditures do not exist for the firms in our sample. Second, as is well known, patent-based measures of research outputs are imperfect. One potential proxy for research output is the number of collaborative patents. However, as is well known, although patents are commonly used as a measure of inventive output, they are also represent a right to exclude others from the invention that the patent incorporates. As a result, firms may patent even very marginal inventions to develop thickets of intellectual property rights (e.g., Hall and Ziedonis 2001, Ziedonis 2004). In other words, changes in the number of patents will reflect firm-level appropriability strategies in addition to coordination costs. A common alternative is to study citation-weighted patents; however, our research framework predicts how changes in coordination costs will influence the likelihood of a collaboration rather than the quality of collaborative output per se.

To mitigate these shortcomings of our data, we study the impact of IT adoption on the likelihood of observing a research collaboration among inventors rather than studying the productivity of research teams. Specifically, we compare the impact of Internet adoption on the likelihood of collaboration (as measured by the incidence of at least one granted patent) among research teams where ex ante coordination costs are high (e.g., geographically dispersed collaborative teams) to those where ex ante costs are low (e.g., collocated inventors). If we observe, for example, after adoption that the likelihood of a collaborative patent among geographically dispersed researchers increases while that of collocated researchers is unchanged, then our results will be informative about how the Internet influenced coordination costs and the relative productivity of different types of research collaborations.

Our framework and research design is motivated by Becker and Murphy's (1992) model of team formation (and the Adams et al. 2005 adaptation to a research context) in that we view decisions about team composition as shaped by the division of labor, task specialization, and coordination costs. In these models, research output is determined by factors such as the number of collaborators, their skill level, and a productivity shifter. Increases in the number of collaborators will increase gross output through task specialization and division of labor.[3] Furthermore, if specialized skills are geographically dispersed throughout the firm, then research output may be increasing in the geographic dispersion of researchers. For example, Adams et al. (2005) demonstrate this for the case where the average skill level of researchers is increasing with geographic dispersion.

However, increases in team size and dispersion are also likely to increase coordination costs. In particular, cross-regional transfer of complex or tacit knowledge is known to be difficult, even within firm boundaries (e.g., Teece 1977, Singh 2008, Sorenson et al. 2006). Furthermore, concerns of free riding and shirking may also be increasing in team size (e.g., Holmstrom 1982), and monitoring geographically dispersed team members may be particularly challenging.

By lowering communication costs, adoption of basic Internet can help to reduce coordination costs. For example, Internet technology can lower communication costs by providing access to Internet protocol (IP)-based e-mail, telephony, and other collaborative tools (Rice 1994, Lee and Choi 2003). This will facilitate lower access costs to others, especially to researchers in distant locations who have relatively

[3] Note that the productivity benefits from teamwork that derive from division of labor and task specialization can be moderated by the detrimental effects of imperfect coordination or by shirking and free riding (see, e.g., Hamilton et al. 2003). More generally Latané et al. (1979) observed that individuals tend to decrease their effort when performing in groups as compared to when performed alone. Karau and Williams (1993, pp. 696–697) showed that such "social loafing" generalizes across tasks and populations, but they nonetheless observed that it is moderated by the meaningfulness and complexity of the tasks and the uniqueness of individual inputs. These moderating factors seem particularly appealing in the context of technology-based R&D that we investigate.

few alternative means of communication available. In short, adoption of basic Internet will lower coordination costs, particularly among geographically dispersed researchers, and will increase the productivity of such teams relative to other types of collaborations.

We note that adoption of basic Internet has the potential to influence research output in other ways than through lower coordination costs. For example, Internet technology facilitates access to codified knowledge (e.g., Ding et al. 2010) by lowering the costs of accessing shared resources such as electronic databases for journals and online repositories for data. It also facilitates the development of more efficient processes for accessing knowledge, as when an institution sets up an online mechanism for accessing books from a library. In short, adoption of basic Internet is likely to increase the total factor productivity for all types of research collaborations. As a result of these declines to coordination costs and improvements to total factor productivity, adoption of basic Internet will lead to an increase in the likelihood of collaboration from geographically dispersed research teams.

The implications of basic Internet adoption for output from other types of research teams are more ambiguous. We consider the impact of basic Internet on two alternative types of teams: the case of collaborative teams within a geographic location and the case of lone inventors. In our setting, lone inventors are those who work within large firms and that are listed as single inventors on a patent. For both of these types of groups, coordination costs will fall by less than for geographically dispersed teams. As a result, productivity for geographically dispersed teams will rise by more than for other types of research groups, leading to a potential shift in research inputs toward geographically dispersed teams. This shift in resources may lead to a decline in research output for collocated and lone inventor teams. However, as noted above, total factor productivity for all types of teams may rise because of declines in the costs of accessing shared resources, so research output may also increase for these latter groups.

In short, it is difficult to determine ex ante whether Internet adoption will lead to an increase or fall in the incidence of patenting for single-location teams and lone inventors: the increase in total factor productivity from declines in shared resource access costs may be offset by a shift in resources toward multilocation collaborations. However, our framework does predict clearly that the postadoption increase in incidence of patenting for these groups will be lower than for geographically dispersed teams.

In short, our research framework implies three predictions:

Hypothesis 1. *Adoption of basic Internet will be associated with an increase in the likelihood of collaboration for multiinventor, geographically dispersed teams.*

Hypothesis 2. *Adoption of basic Internet will be associated with a smaller increase in the likelihood of collaboration for single-location multi-inventor teams than for geographically dispersed teams.*

Hypothesis 3. *Adoption of basic Internet will be associated with a smaller increase in the likelihood of collaboration for lone inventors than for geographically dispersed teams.*

3. Empirical Strategy

3.1. Adoption of Internet Technology and Collaborative Output

We argue that adoption of basic Internet will be associated with a decline in coordination costs for research teams. As a result, we expect an increase in the likelihood of research collaborations from geographically dispersed teams. To examine whether the empirical evidence is consistent with this hypothesis, we seek to measure the impact of Internet adoption on multiinventor collaborations in geographically dispersed firm-location pairs.

We use fixed effects panel data models to study whether adoption of basic Internet technology is associated with an increase in the likelihood of observing a granted patent application coinvented by researchers in a firm-location pair. In short, we use a difference-in-differences identification strategy, comparing the incidence of a collaborative patent in a firm-location pair prior to the treatment of basic Internet adoption to the incidence after treatment. This approach allows us to remove time-invariant unobserved firm-pair features that may be correlated with Internet adoption and patents. This yields the following estimating equation:

$$CollaborativePatent_{ijkt} = \alpha_1 X_{ijkt} + \alpha_2 Z_{ijkt} + \beta\, Internet_{ijkt} + \mu_{ijk} + \tau_t + \varepsilon_{ijkt}. \quad (1)$$

Here $CollaborativePatent_{ijkt}$ is a dummy variable for whether there is a patent coinvented by researchers in both locations j and k of a particular firm i at time t (dated by application year). $Internet_{ijkt}$ measures whether both establishments j and k had adopted basic Internet by time t. Internet technology had not diffused among firms prior to 1995 except in very rare cases, so the value of this variable will be equal to zero prior to this date. We have two types of controls: the variables in X_{ijkt} capture observable changes in firm-pair conditions for things like (the log of per-establishment) firm R&D expenditures and firm-location employment that may affect the volume of collaborations in a firm pair. The variables in Z_{ijkt} capture changes in local characteristics that may influence inventive output, μ_{ijk} measures location-pair fixed effects, and τ_t captures year fixed effects. We estimate our model over the period 1992–1998, using every other year of data.[4] Our hypothesis is that the adoption of basic Internet at both locations in the firm pair will be associated with an increase in the incidence of collaborative patents: a test of $\beta > 0$ against the null of $\beta = 0$. Our estimation approach shares similarities with that used by Agrawal and Goldfarb (2008) in their study on academic research collaborations.

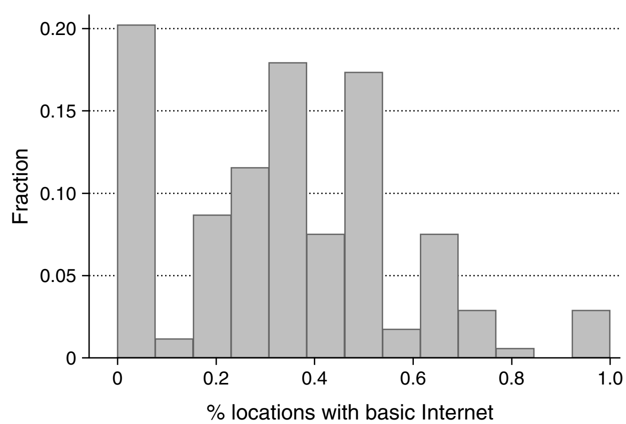
We estimate Equation (1) using a fixed effects linear probability model, and use robust standard errors, clustered over firm-location pairs. Our focus on linear probability models rather than nonlinear approaches reflects several considerations. First and foremost, the linear probability estimates will provide consistent estimates of the parameters of interest: one major drawback to their use will be the existence of heteroskedastic standard errors, which we adjust for using robust standard errors. Second, the linear probability model allows for differencing out the fixed effects without loss of any observations in the data set (as would be the case, for example, with a conditional fixed effects logit or Poisson model). Third, the linear model allows for more straightforward interpretation of the implied marginal effects from our parameter estimates. Last, although fixed effects logit and Poisson regressions also allow for conditioning on fixed effects, King and Zeng (2001) showed that nonlinear methods may be inconsistent when there are a large number of zeroes in the dependent variable, as there are in our sample. Our results are robust to the use of alternative nonlinear models such as conditional fixed effects logit and (unconditional) fixed effects probit models.

As noted above, our endogenous variable is $CollaborativePatent_{ijkt}$, which represents the incidence of a patent applied for in year t with inventors in both locations within the pair. Patents have been used extensively as a measure of research collaborations; however, there are, of course, significant limitations to their use in this way. As Jaffe and Trajtenberg (2002) note, not all inventions meet the U.S. Patent and Trademark Office (USPTO) criteria for patentability. Furthermore, inventors must make an explicit decision to patent an invention, as opposed to relying on some other method for intellectual property protection. In particular, there may be incremental inventive activity that is not patented and therefore is not reflected in patent statistics (e.g., Cohen et al. 2000). Firms may sometimes also choose to use trade secrecy rather than patenting to protect groundbreaking inventions because of incomplete enforcement of property rights. However, the incidence of patents has been shown to be correlated with a firm's stock market value, and thereby provides one useful measure of a firm's intangible stock of knowledge (Hall et al. 2005). Furthermore, as long as a firm location's patent propensity does not vary significantly over time in a way that is correlated with Internet adoption, this should not bias our estimates of the key parameters of interest.

Our estimation framework requires several additional assumptions to identify the parameters of interest. The first is that there exists significant within-firm variance in the adoption of basic Internet within firms. To probe this assumption further, we calculate for each *firm*-year in our sample the percentage of firm locations adopting basic Internet. Figures 1 and 2 present histograms of these percentages for 1996 and 1998. The horizontal axes represent the percentage of adopters within each firm, whereas the vertical axes shows the fraction of firms in our sample within each percentage group. These figures show significant variance in the penetration of basic Internet within firms, particularly in 1996, when the commercial Internet was still at early stages of diffusion. By 1998, almost half (49.1%) of the firms in our sample had 100% penetration, though this reflects the more advanced stage of diffusion of the type of IT investment that we examine more than efforts on the part of firms to coordinate their IT investments.

There are several reasons why we observe this within-firm variance in basic Internet adoption,

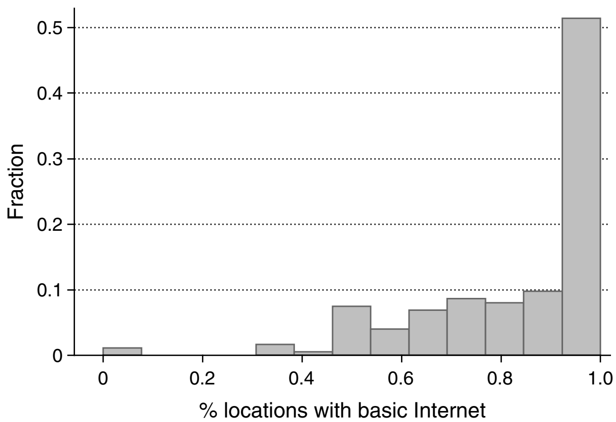
Figure 1 **Percentage of Locations Adopting Internet Within Firms in Estimation Sample, 1996**



Notes. This figure presents the distribution of the percentage of establishments adopting basic Internet by firm in 1996. The horizontal axis represents the percentage of adopters within each firm, whereas the vertical axis shows the fraction of firms in our sample within each percentage group.

[4] In other words, we use data from 1992, 1994, 1996, and 1998. This strategy reflects a resource constraint: these are the years for which we have IT data.

Figure 2 **Percentage of Locations Adopting Internet Within Firms in Estimation Sample, 1998**



Notes. This figure presents the distribution of the percentage of establishments adopting basic Internet by firm in 1998. The horizontal axis represents the percentage of adopters within each firm, whereas the vertical axis shows the fraction of firms in our sample within each percentage group.

despite the obvious benefits to firms of coordinating their Internet investments across locations. First, as has been suggested in prior work, there was significant geographic variance in the cost of business Internet adoption across our sample period (Forman et al. 2005). In particular, there were significant differences in the number of Internet service providers (ISPs) and the services they offered across locations (Downes and Greenstein 2007). Because Internet access provision is local, these differences in supply have the potential to affect the price of Internet access. Furthermore, not all ISPs and not all locations offered high-speed access over this period (Augereau and Greenstein 2001, Augereau et al. 2006): firm establishments in locations without high-speed access would have lower net benefits to adoption. The net benefits of adoption also depended on a firm location's legacy IT infrastructure. For example, locations with heavy investments in legacy mainframe infrastructure or platform-specific investments in prior generations of client–server IT would face considerable costs to adopt Internet technology. It has been shown that these costs shaped Internet adoption patterns (Forman 2005). Thus, if prior IT investments differed across firm locations, these differences likely shaped within-firm adoption patterns. Last, governance of IT functions within firms is frequently decentralized (e.g., Sambamurthy and Zmud 1999, McElheran 2011), and such decentralized governance will lead to investment decisions maximizing local net benefits rather than those of the entire organization, ignoring potential complementarities arising from coordinated investment decisions across locations.

We also require assumptions regarding the nature of unobservables in our regression equation. For Equation (1) to identify the effects of Internet on cross-location pair collaboration, we must assume that unobserved factors can be decomposed into a additively separable time-invariant component and a time-varying component that is constant across location pairs (Athey and Stern 2002). This assumption will be violated if, for example, there exists unobserved time-varying factors that are correlated both with the propensity to adopt basic Internet as well as the likelihood of a cross-location pair collaboration. For example, managers who have initiated a program to encourage cross-location collaborations may adopt basic Internet to signal the importance of this program to researchers.

We do several things to explore both the validity of this assumption and to explore the robustness of our results when it is relaxed. First, we perform several sets of analyses to circumscribe how unobserved factors may influence our results. We conduct a falsification exercise where we examine whether future adoption of Internet technology at a location is correlated with the incidence of a collaborative patent. We find no significant evidence of such a correlation.[5] We further examine whether Internet adoption at only one location in the pair is associated with an increase in patent output. In particular, if Internet adoption is associated with an increase in cross-location collaborative patenting because of a decline in coordination costs, then we should observe no impact on patenting when only one location in the pair adopts Internet technology.[6] This is exactly what we find.

We also demonstrate that our results are robust to the use of instrumental variables. Our instruments proxy for variance in the costs to adoption across locations. One instrument we employ—the average (across locations in the pair) year of price cap regulation in the states in which Internet is adopted—proxies for local telecommunications costs. Another instrument—the average number of ARPANET (a wide area data communications network that was a predecessor of the Internet) nodes across locations in the pair—captures differences in local expertise in an earlier generation of networking technology that may affect the returns to Internet adoption. Further details on these instruments are discussed below.[7, 8]

[5] Although the point estimate is positive for adoption four years in the future, it is negative with a two-year window and consistently nonsignificant at any conventional level.

[6] This falsification exercise is motivated by a similar analysis conducted by Agrawal and Goldfarb (2008).

[7] We thank Avi Goldfarb and Shane Greenstein for providing these variables to us.

[8] We have also experimented with alternative transformations, such as the maximum value of price cap regulation and minimum number of ARPANET nodes across locations in the pair. The estimates are qualitatively similar to those using the baseline instruments.

As noted above, Internet adoption may also be correlated with an increase in collaborative output among researchers collocated within firm-location pairs. However, we expect the relationship to be weaker because the decline in coordination costs will be lower in the within-location case than in the cross-location case (Hypothesis 2). To measure the impact of basic Internet adoption on within-location collaborations, we estimate a variant of Equation (1) for collaborations within a single metropolitan statistical area (MSA). Our endogenous variable will be $CollaborativePatent_{ijt}$, which is a binary indicator for the incidence of a patent applied for in year t with at least two inventors in location j of a particular firm i:

$$CollaborativePatent_{ijt} = \alpha_1^{sl} X_{ijt} + \alpha_2^{sl} Z_{ijt} + \beta^{sl} Internet_{ijt} + \mu_{ij} + \tau_t + \varepsilon_{ijt}. \quad (2)$$

Here, $Internet_{ijt}$ is a binary indicator of whether basic Internet has been adopted at the location, and X_{ijt} and Z_{ijt} represent changes in firm-location and location-level controls, respectively. As noted above, we expect the marginal effect of basic Internet adoption on the incidence of patenting for collocated research teams to be smaller than for geographically dispersed inventors. In fact, if the effect on coordination costs is small, and if basic Internet adoption has little effect on the costs of shared resource use, then we may observe $\beta^{sl} = 0$.

Furthermore, to examine whether basic Internet adoption is associated with an increase in single-inventor patents, we reestimate Equation (2) using only single-authored patents (*SingleAuthoredPatents*):

$$SingleAuthoredPatents_{ijt} = \alpha_1^{slsa} X_{ijt} + \alpha_2^{slsa} Z_{ijt} + \beta^{slsa} Internet_{ijt} + \mu_{ij} + \tau_t + \varepsilon_{ijt}. \quad (3)$$

We expect the marginal effect of basic Internet adoption to be lower here than in the case of multiple inventors, because there will be no effect on coordination costs. In fact, if the adoption of basic Internet has no effect on costs of shared resource usage, then we may observe $\beta^{slsa} = 0$ or even $\beta^{slsa} < 0$.

4. Data

We use a variety of data sources to show how adoption of basic Internet influences collaborative research output within firms. In particular, we match data on IT investment from a well-known private data source on IT investments with patenting data from the USPTO. We obtain firm-level R&D data from Compustat and information for regional controls from the U.S. Census County Business Patterns. In some cases, Harte Hanks does not sample all establishments on Internet use in all years. Furthermore, for some years we do not have R&D data from Compustat. In our baseline results, we constrain our sample to a balanced panel of firm-location pairs for which we have data in all years. However, we also show all of our results using the unbalanced panel and have examined whether our results are robust to imputing missing data.

4.1. Patent Data

We use patent data from the USPTO as a measure of collaborations. Patents are dated using the year of application because of the variance in the patent application-to-grant delay over time, and because application dates are closer to the time when the innovation occurred (e.g., Griliches 1990). We map patents to firm identifiers using the patent's assignee data and the National Bureau of Economic Research (NBER) Patent Data Project's matching data set, which maps patents to a consistent set of unique firm identifiers based on the "GVKEY" code from the Compustat database (Hall et al. 2001).[9] We obtain the universe of patents with a matching GVKEY that were applied for during 1990–1998.

Our analyses will examine the geographic variance in patenting behavior across firm MSAs.[10] Using the inventor location data in U.S. patents, we map inventors to MSAs using the zip code of the inventor (obtained through the USPTO Patents BIB data product). In cases where consolidated MSAs (CMSAs) were present, we used those, because it better allowed us to capture commuting patterns.[11] In regions of the United States that are outside of MSAs, we constructed "phantom" MSAs that consisted of the region of a state outside of all of the MSAs. Our procedure will accurately map patents to the MSAs they were invented in, to the extent that inventors work in the same MSA where they reside. MSAs are constructed in part on the basis of commuting patterns and are widely used as a unit of analysis in studies of the geography of innovation (e.g., Feldman and

[9] For further details on the NBER Patent Data Project, see https://sites.google.com/site/patentdataproject/Home.

[10] This choice was made in part because of a data constraint. Although our IT data are in fact available for individual firm establishments, USPTO patent data provide only inventor locations. Thus, for multiestablishments MSAs, we are unable to identify the particular establishment at which an inventor works within an MSA.

[11] CMSAs represent regions that may contain multiple metropolitan areas, such as Baltimore, Maryland, and Washington, DC, or San Francisco, Oakland, and San Jose, California. We reran our statistical analyses using these component areas (primary MSAs (PMSAs)) instead of CMSAs, and although the results were qualitatively similar, they were somewhat weaker. We attribute these weaker results to measurement error induced by inaccurate mapping of inventors to PMSAs due to commuting patterns of inventors across PMSAs within the same CMSA.

Audretsch 1999); however, our procedure may assign some patents to the wrong MSA when one or more of its inventors commutes to or from a different MSA.

4.2. IT Data

Our data on IT investment come from the Harte Hanks Market Intelligence Computer Intelligence Technology database (hereafter, CI database). The database contains establishment- and firm-level data on characteristics such as the number of employees, personal computers per employee, and use of Internet applications. Harte Hanks collects this information to resell as a tool for the marketing divisions of technology companies. A number of researchers have used these data previously to study adoption of IT (e.g., Bresnahan and Greenstein 1996) and the productivity implications of IT investment (e.g., Bresnahan et al. 2002, Brynjolfsson and Hitt 2003, Bloom et al. 2009). Interview teams survey establishments throughout the calendar year; our sample contains the most current information as of December 1998. As has been discussed elsewhere (e.g., Forman et al. 2005), this data set represents among the best sources of information on the IT investments of private firms available.

Harte Hanks tracks over 300,000 establishments in the United States. Because we focus on industrial research, we exclude government, military, and nonprofit establishments. Our sample from the CI database contains commercial establishments with over 100 employees. Although this limits our sample to predominately large establishments, our algorithm for matching our IT data to firms using Compustat identifiers from the NBER Patent Data Project similarly requires us to focus on large firms. Furthermore, our primary research question—how the adoption of the commercial Internet affected the geography of research collaborations within firms—also circumscribes our focus to large, multiestablishment research organizations. Thus, our analysis should be viewed as a study of IT and research collaborations within large research organizations. Forman et al. (2002) conducted a detailed comparison of the 2000 Harte Hanks data to 1999 U.S. Census County Business Patterns data, we use their results to briefly motivate the extent to which our data are representative of the population of U.S. establishments.[12] They found that the Harte Hanks data contain slightly fewer than half of all establishments with over 100 employees in the United States and represent approximately one-third of all employment. They found that in terms of company size, region, industry, and urban versus rural location, the distributions are quite similar. They found that the Harte Hanks data slightly underrepresent MSAs and CMSAs; however, the regional representation is close, with a slight undersample of the Northeast and oversample of the Midwest.

Our raw data include at least 20 different specific Internet applications, from basic access to software for Internet-enabled ERP (Enterprise Resource Planning) business applications. As noted earlier, we focus on the set of applications and technologies that involve little adaptation by users to be implemented successfully; these are typically some of the technologies that diffused around the initial commercialization of the Internet such as access to the Internet and the creation of static Web pages within an organization. Our focus on this set of technologies reflects our interest in understanding how lower communication costs lowered the coordination costs of geographically dispersed, highly collaborative research.

We define an establishment as a basic Internet adopter if it indicated that it had access to the Internet (i.e., whether the establishment has an ISP), had an internal intranet based on TCP/IP (Transmission Control Protocol/Internet Protocol), or used the Internet for research purposes.[13] In particular, we do not require establishments to adopt electronic commerce or TCP/IP-enabled business application software. Our measure of Internet adoption is meant to capture whether the establishment has adopted enabling technology that will lower communication costs. We set the value of basic Internet equal to zero for all establishments in 1992 and 1994 because these years were prior to the diffusion of the commercial Internet.[14] Although our measure of basic Internet adoption shares some similarities with the measure of Internet participation used in earlier studies of Internet diffusion based on Harte Hanks data (e.g., Forman et al. 2005), there are some differences. In particular, we focus on a narrower set of applications than Forman et al. (2005) do because of our focus on an earlier time period (1996–1998 versus 2000) and changes over time in the questions asked by Harte Hanks.

As noted above, CI data are collected at the establishment level. To map our establishment-level IT data to our patent data, we match establishments to firm MSAs as we had done with the patent data. We first

[12] The 2000 sample from the Harte Hanks database they examined is identical to that used here: all nonfarm business establishments with over 100 employees.

[13] An alternative measure of basic Internet use would incorporate the use of TCP/IP-based email; however, over some periods of our data it is difficult to identify email based on Internet protocols from that based on proprietary networking protocols that were still commonly used over our sample period. To the extent that basic access is required for the use of Internet-based email, we believe our measure captures the use of such email in our sample.

[14] Although it is difficult to date the rise of the commercial Internet, as a point of reference, Netscape's browser became available in early 1995, followed by its initial public offering in December of the same year.

Table 1 Descriptive Statistics for Pairs Including Different MSAs

Variable	Mean	Standard deviation	Minimum	Maximum	Number of observations
Collaboration between inventors in pair	0.0735	0.2610	0	1	18,860
Basic Internet in both locations	0.2050	0.4037	0	1	18,860
Log of per-establishment R&D spending	3.0210	1.4698	−0.4568	7.7295	18,860
Log of establishment employees	7.6932	1.1342	5.2983	12.0369	18,860
Share of local employment in manufacturing	0.2005	0.0643	0.0366	0.5181	18,860
Local average weekly wages	543.6019	85.4429	306.4846	848.329	18,860
Log of local employment	13.8360	0.9449	10.3474	15.7005	18,860
Log of number of local patents	6.6868	1.2107	0.6931	9.1314	18,860

map the unique firm identifier used in the CI database to the GVKEY from the NBER Patent Data Project. We then assign establishments to MSAs using their zip code. For our analysis data set, we include only firm-MSA-year triplets that are from manufacturing firms (Standard Industrial Classifications 20–40) and that are in firm MSAs with at least one patent in two separate years over the period 1992–1998. These restrictions are to retain only firm organizations that perform research for our analyses (many CI database establishments perform no research function); our results are robust to alternative sample restrictions such as firm MSAs with at least one patent over 1992–1998. In cases where there are multiple establishments within an MSA,[15] we calculate a firm location as adopting basic Internet when at least one has done so.

4.3. Firm-MSA Pairs

The focus of our study is on the effects of IT investment on collaborative cross-location inventive output. We estimate the regression model in Equation (1), which allows us to examine, for each pair of firm-MSA establishments, whether the adoption of basic Internet technology in both locations is associated with an increase in the likelihood of a patent. To do this, we form the complete set of pairwise combinations of firm MSAs within a given organization. Based on coauthorship, we identify the incidence of collaborations that were performed between units in different MSAs in a given patent-application year. We further use Equations (2) and (3) to examine whether there is a relationship between basic Internet adoption and within-MSA output.

4.4. Other Controls

We combine these data with additional information from a number of sources. The additional data are used to control for time-varying factors that may be correlated with basic Internet adoption and with patent output. First, to control for variance in R&D inputs across firms, we compute the flow of R&D spending dollars using Compustat and compute the per-location R&D flow dollars by normalizing total spending by the number of firm-MSA locations in our data.[16] Second, we compute total firm-location employment as the sum of employment across establishments within the location. Unfortunately, our CI data begin with 1996, so we are unable to observe firm-location employment in 1992 and 1994. We use 1996 employment values for these to observe some time trend in employment growth; all of our results are robust to removing the employment variable. In our pair regressions, we compute the log of the average employment across the two locations.

Next, we control for a number of local factors that may influence both the likelihood of basic Internet adoption as well as innovation productivity and the propensity to patent. The data sources for these measures are at the county level and are then matched to MSAs and computed for a firm-MSA-year triplet. For our cross-location pair regressions, these data items are then averaged across triplets in a pair.[17] We use the percent of manufacturing employment in the MSA, the average weekly wage in the MSA, and the log of MSA employment. These variables are computed using U.S. Census County Business Patterns data. Using the USPTO data, we also compute the log of the total number of patents in the MSA-application

[15] This is the case for 35% of the firm MSAs in our analysis sample.

[16] An alternative procedure would be to deflate by the number of establishments. However, some establishments in our data do not engage in innovative activity. Furthermore, because our output measure is based on firm-location pairs, our procedure matches R&D input with innovative output.

[17] For our analyses of patent output within a single MSA, the average value is equal to the value of the variable for the firm-MSA-year triplet, because both triplets in the pair are equal to the same value.

Table 2 Descriptive Statistics for Within MSA Analyses

Variable	Mean	Standard deviation	Minimum	Maximum	Number of observations
Collaboration between inventors in location	0.7109	0.4534	0	1	4,276
Basic Internet in both locations	0.2935	0.4554	0	1	4,276
Log of per-establishment R&D spending	2.9930	1.4417	−0.9715	7.7295	4,276
Log of establishment employees	7.4127	1.2302	5.2983	12.2121	4,276
Share of local employment in manufacturing	0.1991	0.0823	0.0204	0.5661	4,276
Local average weekly wages	549.1758	113.4011	298.3717	860.2807	4,276
Log of local employment	13.5938	1.3047	9.8924	15.8465	4,276
Log of number of local patents	6.3291	1.6904	0	9.1411	4,276

year. For the latter two (logged) measures, we compute the log of the average across the two MSAs in our pair regressions.

Descriptive statistics for all variables are provided in Table 1 for cross-MSA collaborations and in Table 2 for within-MSA collaborations. Among MSA pairs, 0% had adopted basic Internet in 1992 and 1994, 12% had adopted by 1996, and 58% had adopted by 1998.

5. Results

We first establish a relationship between the adoption of basic Internet and the incidence of collaborative patenting at geographically dispersed research locations. We demonstrate that these results are robust to a variety of specifications and robustness checks, and to the use of instrumental variables. We then show that there is no significant effect of adoption on the number of collaborative patents invented by researchers within a location, nor on the number of single-inventor patents.

5.1. Baseline Results

We begin with some tests to demonstrate the variance in our data that identifies our core result. Table 3 reports a nonparametric difference-in-differences analysis of the percentage of cross-MSA pairs with a collaboration between 1992 and 1998 and according to their adoption (or nonadoption) of basic Internet. We study the change in collaboration patterns over these two years because they represent the beginning and end of our sample period. The results suggest a statistically significant increase in the incidence of collaborative patenting occurred for cross-location pairs adopting Internet over the period, relative to nonadopters. MSA pairs that both adopted basic Internet had an average increase in the likelihood of a collaboration that was 1.6 percentage points higher than nonadopters over this period (a difference that is statistically significant at the 5% level);[18] this compares to a pre-Internet likelihood of collaboration for future adopters of 6.4%.

In contrast, Tables 4 and 5 show that there is no significant difference between adopters and nonadopters of basic Internet in the change in the likelihood of observing a collaborative patent within single-MSA teams over the same period, nor is there any change in the likelihood of observing a patent for single inventors. In fact, lone inventors at locations that adopted basic Internet experienced a slower growth in within-location patenting than those at nonadopting locations, though the difference is significant only at the 10% level.

In Table 6 we use the regression model in Equation (1) to examine the implications of basic Internet adoption for the likelihood of observing cross-location collaborative patents (Hypothesis 1). Column (1) shows the correlation between basic Internet and the likelihood of collaboration without any controls; the correlation is significant and positive. Column (2) shows what we view as our baseline specification: it includes controls for time-varying firm-location

Table 3 Likelihood of a Pair Collaboration by Year and Whether Treated by Internet Adoption, Firm-MSA Pairs

	Before treatment (1992)	After treatment (1998)	First difference (row)
Received Internet treatment	0.0636 ($N = 3{,}301$)	0.0891 ($N = 3{,}301$)	0.0254** ($N = 3{,}301$)
Did not receive Internet treatment	0.0580 ($N = 1{,}414$)	0.0672 ($N = 1{,}414$)	0.0092 ($N = 1{,}414$)
First difference (column)	0.0056 ($N = 4{,}715$)	0.0219** ($N = 4{,}715$)	Difference-in-differences 0.0163* ($N = 4{,}715$)

Note. We base this analysis on the sample of firm-location pairs that are observed before and after the treatment between 1992 and 1998.

*Difference is significant at the 5% level; **difference is significant at the 1% level.

[18] Nonadopters include pairs where neither and only one member of the pair adopted Internet.

Table 4 Likelihood of a Collaboration Between Inventors at Location by Year and Whether Treated by Internet Adoption, Within-MSA Analyses

	Before treatment (1992)	After treatment (1998)	First difference (row)
Received Internet treatment	0.6745 ($N = 894$)	0.7092 ($N = 894$)	0.0347+ ($N = 894$)
Did not receive Internet treatment	0.6686 ($N = 175$)	0.6914 ($N = 175$)	0.0229 ($N = 175$)
First difference (column)	0.0059 ($N = 1{,}069$)	0.0177 ($N = 1{,}069$)	Difference-in-differences 0.0118 ($N = 1{,}069$)

Note. We base this analysis on the sample of firm locations that are observed before and after the treatment between 1992 and 1998.
+Difference is significant at the 10% level.

and location-specific characteristics. The coefficient on basic Internet is 0.0169; in other words, if both locations in a pair adopt basic Internet, this translates into a 1.69 percentage point increase in the likelihood of collaboration between the two locations. These results are statistically significant at the 5% level. Table 1 shows there is a mean likelihood of collaboration across location pairs of 7.35%; this suggests that basic Internet adoption is associated with a 23.0% increase in the likelihood of collaboration above the mean. As is common in linear probability models (e.g., Athey and Stern 2002), the overall R^2 (computed by excluding the fixed effects in the R^2 computation) is low. However, the R^2 values increase significantly in size once the explanatory power of the fixed effects are incorporated.

Column (3) shows that although the level of statistical significance declines slightly from 5% to 10% (p-value 0.077), the results are robust to the use of an unbalanced panel. As noted above, dropped observations in our sample often arise from missing Internet data in 1996 and 1998. Because of the small number of time periods we observe after the commercialization of the Internet and our reliance on within-panel variance for identification, we speculate that adding panels with missing data will introduce additional noise into our estimates (however, we acknowledge this assertion is impossible to test formally). We have examined the results of alternative approaches for imputing missing values in our full sample and find that our baseline results are robust to these changes.

Table 5 Likelihood of a Patent by Year and Whether Treated by Internet Adoption, Within-MSA Analyses for Single-Inventor Patents

	Before treatment (1992)	After treatment (1998)	First difference (row)
Received Internet treatment	0.4564 ($N = 894$)	0.4318 ($N = 894$)	−0.0246 ($N = 894$)
Did not receive Internet treatment	0.3600 ($N = 175$)	0.4057 ($N = 175$)	0.0457 ($N = 175$)
First difference (column)	0.0964** ($N = 1{,}069$)	0.0261 ($N = 1{,}069$)	Difference-in-differences −0.0703+ ($N = 1{,}069$)

Note. We base this analysis on the sample of firm MSAs that are observed before and after the treatment between 1992 and 1998.
**Difference is significant at the 1% level; +difference is significant at the 10% level.

We have examined the robustness of our results to a variety of different distributional assumptions, including the conditional logit, random effects probit, and fixed effects probit. In the last model, the fixed effects are estimated. To control for differences in the innovativeness and patent propensity across firm locations we include controls for the total number of patents invented at both locations (excluding the patents invented in the focal pair); we do not include these controls in our baseline specifications because they are potentially endogeneous. We also estimate our regression models excluding IT-producing industries because these may be particularly adept at using IT to facilitate research collaborations. Our results are robust to all of these changes.

5.2. Robustness Analysis

This section presents the results of a variety of tests to address omitted variable bias and potential simultaneity. We first present the results of a series of falsification exercises, and then the results of a series of instrumental variables regressions.

In columns (4) and (5) of Table 6 we show the results of a falsification test that utilizes the timing of Internet adoption. Both columns examine whether future adoption of the Internet is correlated with increases in the current likelihood of a collaborative patent. If the positive correlation between Internet and collaboration observed in column (2) reflects time-varying omitted factors, then we would expect a positive correlation between future Internet adoption and current collaboration. Column (4) of Table 6 includes additional dummies indicating whether the firm-location pair will adopt at both locations two and four years in the future. The parameter estimate of basic Internet adoption at both locations (0.0166) is very similar to that in our baseline estimate in column (2), although significance falls slightly to the 10% level (p-value 0.068). However, the parameter estimate on Internet two years and four years in the future are both statistically insignificant, and the parameter estimates on Internet adoption today and two years in the future are statistically different from one another at the 1% level (the p-value on the difference between Internet today and Internet four years from now is 0.1812). Furthermore, there is no observable time trend in the impact of Internet prior to adoption, though we acknowledge that because we have

Table 6 Baseline Results—Different CMSAs

	(1)	(2)	(3)	(4)	(5)	(6)
	Excludes time-varying pair and location controls	Baseline	Unbalanced panel	Includes indicators for four years prior to adoption	Includes indicators for four years prior to adoption (pooled)	Includes variable for when only one location adopts
Basic Internet in both locations	0.0176 (0.0069)*	0.0169 (0.0069)*	0.0119 (0.0067)+	0.0166 (0.0091)+	0.0191 (0.0087)*	0.0176 (0.0069)*
Basic Internet in only one location						−0.0040 (0.0066)
Basic Internet two years in the future				−0.0030 (0.0086)		
Basic Internet four years in the future				0.0053 (0.0075)		
Basic Internet two or four years in the future					0.0025 (0.0072)	
Log of per-establishment R&D spending		0.0300 (0.0070)**	0.0267 (0.0061)**	0.0300 (0.0070)**	0.0299 (0.0069)**	0.0299 (0.0070)**
Log of establishment employees		0.0012 (0.0149)	0.0062 (0.0144)	0.0013 (0.0149)	0.0012 (0.0149)	0.0014 (0.0149)
Share of local employment in manufacturing		−0.1345 (0.2849)	−0.1925 (0.2546)	−0.1417 (0.2852)	−0.1365 (0.2850)	−0.1363 (0.2851)
Local average weekly wages		0.0002 (0.0002)	0.0002 (0.0002)	0.0002 (0.0002)	0.0002 (0.0002)	0.0002 (0.0002)
Log of local employment		−0.0283 (0.0731)	−0.0763 (0.0646)	−0.0294 (0.0730)	−0.0292 (0.0730)	−0.0290 (0.0731)
Log of number of local patents		0.0036 (0.0191)	0.0157 (0.0167)	0.0039 (0.0191)	0.0037 (0.0191)	0.0039 (0.0192)
Observations	18,860	18,860	25,670	18,860	18,860	18,860
Number of groups	4,715	4,715	7,233	4,715	4,715	4,715
R^2 (overall)	0.00	0.01	0.00	0.01	0.01	0.01
R^2 (includes fixed effects)	0.53	0.54	0.57	0.54	0.54	0.54

Notes. The dependent variable is the incidence of a collaborative patent between inventors in both MSAs in the pair. R^2 (with fixed effects) includes fixed effects in R^2 computation. All regressions include constant term and time dummies. Robust standard errors, clustered on firm-location pairs, are in parentheses.

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

only two years of data we are unable to make strong causal assertions on the basis of these results. Column (5) constrains the parameters for two years and four years in the future to be equivalent; again the coefficient estimate on Internet adoption (0.0191) is qualitatively similar to that in column (2) (and statistically significant at the 5% level), and the coefficient estimate on Internet two or four years in the future is statistically insignificant. Furthermore, the coefficient estimates on Internet today and in the future are statistically significantly different from one another at the 5% level.

Second, following Agrawal and Goldfarb (2008), we examine whether basic Internet adoption at only one firm location is correlated with the number of collaborative patents. If Internet adoption influences research productivity primarily by lowering coordination costs, then adoption at one location should have no impact on the growth in the number of patents—adoption at both locations is necessary. However, if basic Internet influences productivity by lowering the costs of accessing shared resources, then we may observe a relationship between single-location adoption and collaborative output. Column (6) shows that basic Internet adoption at one location has no impact on the likelihood of observing a collaborative patent, and furthermore, the parameter estimate of basic Internet at both locations is significantly different from that of basic Internet at only one location at the 5% level. This result is consistent with the view that adoption of basic Internet influences collaborations by lowering coordination costs: we provide further evidence in support of this view in our tests of Hypotheses 2 and 3. In terms of robustness,

Table 7 First-Stage Instrumental Variable Results, Baseline Regressions (Different MSAs)

	(1)	(2)	(3)	(4)	(5)	(6)
	Baseline	Regulatory instrument only	ARPANET nodes only	LIML estimates	Unbalanced panel	Excludes locations in the same state
First year of ROR regulation × *1996 dummy*	−0.0062 (0.0017)**	−0.0078 (0.0017)**		−0.0062 (0.0017)**	−0.0045 (0.0014)**	−0.0063 (0.0018)**
First year of ROR regulation × *1998 dummy*	−0.0167 (0.0028)**	−0.0164 (0.0028)**		−0.0167 (0.0028)**	−0.0159 (0.0027)**	−0.0155 (0.0030)**
Number of ARPANET nodes × *1996 dummy*	0.0147 (0.0032)**		0.0143 (0.0032)**	0.0147 (0.0032)**	0.0109 (0.0026)**	0.0153 (0.0033)**
Number of ARPANET nodes × *1998 dummy*	−0.0044 (0.0041)		−0.0043 (0.0041)	−0.0044 (0.0041)	−0.0056 (0.0039)	−0.0024 (0.0042)
Log of per-establishment R&D spending	0.0237 (0.0100)*	0.0240 (0.0100)*	0.0217 (0.0100)*	0.0237 (0.0100)*	0.0171 (0.0080)*	0.0227 (0.0103)*
Log of establishment employees	0.0380 (0.0206)+	0.0398 (0.0206)+	0.0217 (0.0100)*	0.0380 (0.0206)+	0.0327 (0.0198)+	0.0321 (0.0213)
Share of local employment in manufacturing	−1.4758 (0.4970)**	−1.3499 (0.4943)**	−0.5339 (0.4998)	−1.4758 (0.4970)**	−1.2344 (0.4041)**	−1.5283 (0.0513)**
Local average weekly wages	−0.0003 (0.0003)	−0.0004 (0.0002)+	0.0001 (0.0003)	−0.0003 (0.0003)	−0.0001 (0.0002)	−0.0003 (0.0003)
Log of local employment	0.0125 (0.1079)	−0.0513 (0.1040)	−0.0274 (0.1073)	0.0125 (0.1079)	0.0332 (0.0872)	−0.0222 (0.1123)
Log of number of local patents	0.0462 (0.0264)+	0.0555 (0.0262)*	0.0409 (0.0265)	0.0462 (0.0264)+	0.0241 (0.0207)	0.0447 (0.0278)
Observations	18,860	18,860	18,860	18,860	25,594	17,824
Number of groups	4,715	4,715	4,715	4,715	7,157	4,456
F-statistic	18.26	22.67	14.42	18.26	16.82	15.38
Stock and Yogo (2005) critical values	10.27/13.96	. /11.59	. /11.59	. /3.87	10.27/13.96	10.27/13.96

Notes. First-stage dependent variable is an indicator for whether both MSAs in the pair have basic Internet. All regressions include time dummies. Stock and Yogo (2005) critical values are reported for relative bias >10% and maximal instrumental variable size >15%, respectively. Missing Stock and Yogo (2005) critical values mean they have not been computed or do not apply. Robust standard errors, clustered on firm-location pairs, are in parentheses.

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

these results suggest that if omitted variable bias is influencing our results, it must do so only when both establishments adopt basic Internet.

To further address concerns about omitted variable bias, in Tables 7 and 8 we include the results of instrumental variable estimates.[19] Both of our instruments identify cross-sectional variance in the costs to Internet adoption. Our first instrument captures differences in local regulatory policy. We identify the year in which rate of return (ROR) regulation is instituted in the two states in the firm-location pair. Greenstein and Mazzeo (2006) argue that this variable captures local variance in regulatory stringency: lower values of this variable should indicate a regulatory environment in which there is a friendlier attitude toward experimenting with competition, which should translate into lower costs for an entering competitive local exchange carrier. We expect that because of the presence of additional competition, such environments will be associated with potentially lower operating costs for Internet service providers. Thus, lower values of this variable should translate into lower Internet adoption costs for firms.

Figure 3 presents the distribution of the values of this variable across firm locations in our sample. Most regulatory policy changes occurred prior to our sample period; thus, our identification strategy rests on the timing of ROR regulation as a proxy for the local regulatory environment faced by ISPs and Internet adopters, rather than a policy shock that occurs during our sample period. The distribution of this regulatory policy change varies widely across nodes in the pair: the correlation between the values of this variable at the two different nodes in the pair is

[19] We note that one particular source of omitted variable bias that may be a concern is if managers of the firm emphasize globalization of research in the organization and use Internet adoption as a signal of their commitment to global research. We note that to the extent that our instruments are very likely to be uncorrelated with these changes in managerial focus, our use of instrumental variables should help to address this concern.

Table 8 Second-Stage Instrumental Variable Results, Baseline Regressions (Different MSAs)

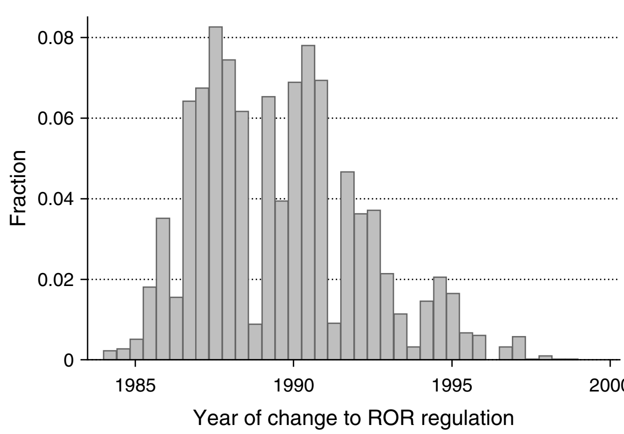
	(1)	(2)	(3)	(4)	(5)	(6)
	Baseline	Regulatory instrument only	ARPANET nodes only	LIML estimates	Unbalanced panel	Excludes locations in the same state
Basic Internet in both locations	0.1703 (0.0758)*	0.1353 (0.0979)	0.2091 (0.1230)+	0.1761 (0.0789)*	0.1816 (0.0829)*	0.2285 (0.0803)**
Log of per-establishment R&D spending	0.0266 (0.0072)**	0.0274 (0.0072)**	0.0257 (0.0077)**	0.0265 (0.0072)**	0.0239 (0.0063)**	0.0238 (0.0073)**
Log of establishment employees	−0.0041 (0.0156)	−0.0029 (0.0157)	−0.0055 (0.0160)	−0.0043 (0.0156)	0.0013 (0.0152)	−0.0087 (0.0163)
Share of local employment in manufacturing	−0.0702 (0.2999)	−0.0849 (0.2977)	−0.0540 (0.3061)	−0.0678 (0.3010)	−0.1229 (0.2685)	−0.0199 (0.3052)
Local average weekly wages	0.0002 (0.0002)	0.0002 (0.0002)	0.0002 (0.0002)	0.0002 (0.0002)	0.0002 (0.0002)	0.0002 (0.0002)
Log of local employment	−0.0140 (0.0751)	−0.0172 (0.0748)	−0.0103 (0.0764)	−0.0134 (0.0753)	−0.0691 (0.0664)	0.0159 (0.0767)
Log of number of local patents	−0.0043 (0.0198)	−0.0025 (0.0200)	−0.0063 (0.0205)	−0.0046 (0.0198)	0.0109 (0.0171)	0.0036 (0.0203)
Observations	18,860	18,860	18,860	18,860	25,594	17,824
Number of groups	4,715	4,715	4,715	4,715	7,157	4,456
Overidentification test (*p*-value)	0.3859	0.3664	0.1584	0.3868	0.1124	0.5341
Hausman test (*p*-value)	0.6737	0.9917	0.9971	0.6917	0.2688	0.3665

Notes. The dependent variable is the incidence of a collaborative patent between inventors in both MSAs in the pair. All regressions include constant term and time dummies. Robust standard errors, clustered on firm-location pairs, are in parentheses.
+Significant at 10%; *significant at 5%; **significant at 1%.

only 0.0291. Furthermore, only 5.49% of pairs in our sample consist of two MSAs that are in the same state, indicating that there is significant variance in the instrument across dyads and across nodes within dyads in our sample.

Our second instrument is the number of local connections in the MSA to the ARPANET—a wide area network that was a predecessor to the Internet.

Figure 3 Year of Change to ROR Regulation Across Firm-MSA Pairs in Estimation Sample

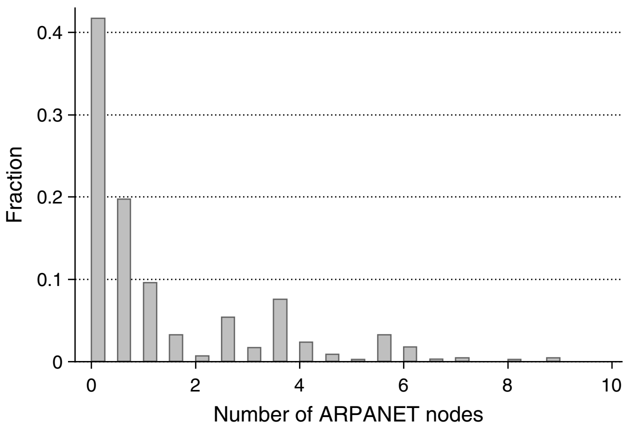


Notes. This figure presents the distribution of year of change to ROR regulation among firm-MSA pairs in estimation sample. Each data point represents the average year of change across two locations in the pair.

Increases in this variable will capture variance in local expertise with networking technologies. Forman et al. (2005, 2008) argue that firm establishments in locations with greater human capital specific to IT and networking technologies will have higher net benefits to adopting Internet technology. Furthermore, because it represents historical decisions (from the 1970s) about connectivity to U.S. Department of Defense or U.S. university networks, this variable is unlikely to be correlated with economic activity over our sample period. For these reasons, Forman et al. (2012) use this variable to instrument for local county Internet adoption in their study of the effects of Internet adoption on growth in local wages.

Figure 4 presents the distribution of the average (across locations within a pair) number of nodes across firm-location pairs in our sample. Because of the relative sparseness of the ARPANET, over 40% of observations in our sample have no proximity to an ARPANET node; however, beyond this mass point there is significant variance across pairs in our sample. Variance in this variable will capture in part urban/rural differences but will also capture something more. Even within large cities there is considerable variance in the number of nodes—for example, Washington, DC, has 11 nodes, whereas Philadelphia, Pennsylvania, has 0—and many small cities will have nodes because of the presence of

Figure 4 Number of ARPANET Nodes Across Firm-MSA Pairs in Estimation Sample



Notes. This figure presents the distribution of number of ARPANET nodes among firm-MSA pairs in estimation sample. Each data point represents the average number of nodes across two locations in the pair.

ARPANET nodes at military bases (e.g., Kirtland Air Force Base, New Mexico) or state universities (e.g., University of Illinois at Urbana–Champaign). Like the regulatory instrument, the distribution of number of ARPANET nodes varies widely across the firm locations within the dyads in our data. The correlation in values of this variable at the two different nodes in the pair is 0.0680.

As noted earlier, there are substantial differences in the penetration of basic Internet in 1996 and 1998. To allow for heterogeneous impacts on the effects of our instruments over time, we interact both with a 1996 and 1998 time dummy. Thus, we have four instruments in total: year of ROR regulation × 1996 dummy, year of ROR regulation × 1998 dummy, number of ARPANET nodes × 1996 dummy, and number of ARPANET nodes × 1998 dummy.

Table 7 presents the first-stage results of our instrumental variable regressions. Column (1) includes all four instruments, column (2) only the regulatory instruments, and column (3) the only ARPANET nodes instruments; column (4) presents limited information maximum likelihood (LIML) estimates (all others are two-stage least squares), column (5) presents the results of all four instruments with the unbalanced panel, and column (6) excludes pairs with locations in the same state from the estimation sample. As expected, increases in the time to rate of return regulation are associated with a lower likelihood of adoption; this is true both in 1996 and 1998. Increases in the number of ARPANET nodes are associated with a higher likelihood of adoption in 1996, but have little impact on adoption in 1998 when Internet technology was more mature and when local expertise was likely less important for the type of IT investment we examine. The size and direction of these parameter estimates are quite stable across specifications and insensitive to whether one or both sets of instruments are included, indicating that they capture different sources of variance in our data. The values of the F-statistics on the excluded instruments in the first-stage regression range from 14.42 to 22.67, and in all cases are significant above the 1% level. We also report the Stock and Yogo (2005) critical thresholds for weak instruments. Following Stock et al. (2002), we report critical thresholds for the test that the bias of two-stage least squares regression is no more than 10% of the inconsistency of the ordinary least squares regression, and that the size of the maximal Wald test for the first-stage instruments is large enough that a 5% hypothesis test rejects no more than 15% of the time. In all cases the F-statistic surpasses these critical values.

Table 8 presents the second-stage results. Although the direction of the estimated effect of basic Internet on research collaborations is stable across specifications, the magnitude and significance of the coefficient estimates differ—our two sets of instruments have additional power when included together. Furthermore, the coefficient estimates in Table 8 are consistently larger than those in column (2) of Table 6. We speculate that this may be because of heterogeneous effects of basic Internet on collaboration; that is, overall, the local average treatment effect for basic Internet may be largest for those pairs whose adoption is most influenced by regulatory regime and local networking skills. In other words, although the instruments are uncorrelated with the incidence of collaboration except through their influence on the likelihood of basic Internet adoption, the marginal effect of basic Internet on collaboration is largest among the group whose behavior is most strongly affected by the instruments. Despite the increase in coefficient estimates, a Hausman test retains the null that the coefficient estimates in Table 8 are not different from their counterparts without instruments.[20] All of our models are overidentified, and the p-value of the overidentification statistic for the baseline sample ranges from 0.1584 for the ARPANET-nodes-only model to 0.3868 for the LIML model. The p-value for the overidentification statistic for the unbalanced panel is 0.1124, lower than other models but still unable to reject the null at conventional levels.

In sum, our instrumental variable results provide additional evidence in support of a causal interpretation that adoption of basic Internet led to an

[20] We compare columns (1)–(4) of Table 8 to column (2) of Table 6, column (5) of Table 8 to column (3) of Table 6, and column (6) of Table 8 to a comparable sample that does not use instrumented variables (not reported).

Table 9 Same-MSA and Single-Inventor Results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Same MSA					Single inventor				
	No controls	Baseline	Unbalanced panel	For years prior to adoption	Four years prior to adoption (pooled)	No controls	Baseline	Unbalanced panel	Four years prior to adoption	Four years prior to adoption (pooled)
Basic Internet in both locations	0.0007	0.0002	−0.0018	0.0343	0.0516	−0.0122	−0.0115	0.0026	−0.0584	−0.0430
	(0.0255)	(0.0252)	(0.0237)	(0.0511)	(0.0413)	(0.0238)	(0.0237)	(0.0225)	(0.0480)	(0.0389)
Basic Internet two years in the future				0.0282					−0.0468	
				(0.0419)					(0.0385)	
Basic Internet four years in the future				0.0482					−0.0290	
				(0.0294)					(0.0267)	
Basic Internet two or four years in the future					0.0478					−0.0293
					(0.0295)					(0.0267)
Log of per-establishment R&D spending		0.1062	0.1068	0.1065	0.1061		0.0702	0.0844	0.0707	0.0703
		(0.0247)**	(0.0218)**	(0.0247)**	(0.0246)**		(0.0224)**	(0.0207)**	(0.0224)**	(0.0224)**
Log of establishment employees		0.0283	0.0387	0.0270	0.0261		−0.0253	−0.0158	−0.0231	−0.0239
		(0.0370)	(0.0385)	(0.0369)	(0.0367)		(0.0352)	(0.0368)	(0.0353)	(0.0353)
Share of local employment in manufacturing		0.0444	0.2480	0.0042	0.0055		−0.1390	−0.1042	−0.1164	−0.1152
		(10.847)	(10.148)	(10.850)	(10.853)		(0.9578)	(0.8855)	(0.9534)	(0.9544)
Local average weekly wages		−0.0008	−0.0007	−0.0008	−0.0008		−0.0002	−0.0001	−0.0002	−0.0002
		(0.0004)+	(0.0004)+	(0.0004)+	(0.0004)+		(0.0004)	(0.0004)	(0.0004)	(0.0004)
Log of local employment		0.3694	0.2492	0.3659	0.3649		0.1855	0.1008	0.1891	0.1882
		(0.2506)	(0.2320)	(0.2496)	(0.2497)		(0.2282)	(0.2166)	(0.2286)	(0.2285)
Log of number of local patents		0.1416	0.1093	0.1430	0.1443		0.1019	0.0759	0.0990	0.1002
		(0.0518)**	(0.0486)*	(0.0518)**	(0.0518)**		(0.0490)*	(0.0455)+	(0.0488)*	(0.0489)*
Observations	4,276	4,276	5,237	4,276	4,276	4,276	4,276	5,237	4,276	4,276
Number of groups	1,069	1,069	1,450	1,069	1,069	1,069	1,069	1,450	1,069	1,069
R^2 (overall)	0.00	0.02	0.02	0.02	0.02	0.00	0.03	0.03	0.03	0.03
R^2 (with fixed effects)	0.40	0.40	0.41	0.40	0.40	0.58	0.58	0.60	0.58	0.58

Notes. The dependent variable in columns (1)–(5) is the incidence of a collaborative patent between inventors in the same MSA. The dependent variable in columns (6)–(10) is the incidence of a patent from a single inventor. The R^2 (with fixed effects) includes fixed effects in the R^2 computation. All regressions include constant term and time dummies. Robust standard errors, clustered on firm locations, are in parentheses.

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

increase in the likelihood of cross-location collaboration. Although on their own they may not completely rule out a role for omitted variable bias, in combination with our controls and the variety of other robustness checks we have performed they do lend support for the view that adoption of basic Internet technology led to an increase in the likelihood of cross-location collaborations.

5.3. Results for Within-Location Patenting

In Table 9 we show the results of our model that explores the relationship between basic Internet adoption and the likelihood of within-location collaboration. The results in all columns suggest that there exists no correlation between basic Internet adoption and the likelihood of within-location patenting, either collaborative or single authored. These results are consistent with Hypotheses 2 and 3. Because there is no evidence of a correlation between basic Internet and the incidence of patenting across any of these models and that our prior is that, if anything, our coefficient estimates on basic Internet will be upward biased, to save space we do not present instrumental variable estimates of these models.[21]

In sum, Tables 3–9 show that adoption of basic Internet was associated with an increase in the incidence of collaborative, geographically dispersed research. However, there is no evidence of an increase in either collaboration within a geographic location or in output from lone inventors. This evidence—together with the results on single-location adoption in column (6) of Table 6—show that although there exists evidence that basic Internet lowered coordination costs among researchers, there is little evidence that basic

[21] However, we have estimated the analogs to the results in Table 9 using instrumental variables, and the results are consistent with the results without instruments. These results are available from the authors upon request.

Internet significantly improved researcher productivity through access to shared resources, at least in our setting and over this specific time period.

One interpretation of our results is that basic Internet has increased the productivity of larger, geographically dispersed research teams relative to other types of research collaborations. Here we address two further concerns with this assertion. First, our approach uses separate regressions to compare patent output for cross-location teams, single-location teams, and lone inventors. Our reason for this approach is that it is more flexible because it allows for heterogeneity in the effects of observables and unobservables on patent output across different types of research collaborations. However, it prevents a direct statistical test of the marginal effect of Internet across these different types of collaborations. To address this concern, we ran a regression that pools the data used to estimate Equations (1)–(3).[22] We interact our Internet variable with a dummy that indicates whether the observation corresponds to either a single-location team or lone inventor and reestimate the model with the same set of variables as before but including this new variable. We find that this variable is negative and statistically significant at the 5% level,[23] indicating Internet adoption was associated with a significantly smaller increase in the likelihood of observing a patent for single-location teams or lone inventors than for multilocation teams.

Another concern relates to differences in the likelihood of observing a patent in a given year for cross-location teams, single-location teams, and lone inventors. The mean likelihood of observing a collaborative patent for a cross-location pair is 7.35%; in contrast, the mean likelihood of observing a collaborative patent for a single-location pair-year is 71.09%, whereas the mean likelihood of observing a lone inventor patent in our data in a given year is 44.04%. Thus, our finding of an insignificant positive coefficient for Internet for these latter two groups could simply be due to the margin we consider—Internet may not shift the likelihood that we observe at least one patent but could influence the expected number of patents. To examine the salience of this alternative hypothesis, we have estimated models with the (unweighted) count of the number of patents as the dependent variable. These regressions are essentially the same as in regression models (1) through (3); however, because the dependent variable is a count we estimate the model using a conditional fixed effects Poisson regression with robust standard errors clustered by firm-location pairs (for model (1)) or firm locations (for models (2) and (3)). To conserve space, we describe a summary of these results here. We find no significant effect of Internet adoption on the expected number of patents produced by single-location teams or lone inventors. Furthermore, we find that Internet adoption is associated with a statistically significant (at the 1% level) smaller increase in patents from single-location than from cross-location pairs. Together, these results add evidence that Internet adoption has shifted the productivity of cross-location research relative to other types of collaborations.

6. Conclusion

We examine the implications of basic Internet adoption for reducing the coordination costs of industrial research teams. We match local (MSA) business IT investment data with local firm patenting activity and, using panel data fixed effects models, find robust empirical evidence that basic Internet adoption is associated with an increased likelihood of collaboration (as measured through collaborative patents) in geographically dispersed firm teams. On the contrary, we find no evidence of such a link between Internet adoption and within-location collaborative patents, nor do we find any evidence of a relationship between basic Internet and single-inventor patents. We interpret these results as evidence that basic Internet adoption lowered the coordination costs of geographically dispersed research teams; however, basic Internet adoption does not seem to be associated with increased research output as a result of easier access to electronic knowledge systems or shared resources (at least during our sample period).

Our results stand in contrast to recent work on IT and academic research that has found that IT adoption leads to a disproportionately greater increase in collaborations among researchers who are geographically close to one another (Agrawal and Goldfarb 2008). There are several potential reasons for this difference in results. First, Agrawal and Goldfarb (2008) studied BITNET, a predecessor network to the Internet. Although the latter allows for content-rich information and knowledge exchanges, one of the main benefits of the former was to share scarce computing resources. Next, whereas we look at incidence of collaborative patents, they focused on scholarly publications. The differences in costs and processes leading to these research outputs may also explain some of the differences that we observe. Finally, we look at within-firm industry collaborations, whereas Agrawal and Goldfarb (2008) examined academic collaborations across universities. Geographic proximity is

[22] In these regressions, we treat observations for single-location teams and lone inventors as "pairs" where both locations in the pair are identical.

[23] In a regression with our baseline set of controls, the coefficient estimate for *Internet* is 0.0191, and the coefficient estimate for *Internet* × *SameMSA* is −0.0283.

commonly thought to facilitate the formation of new relationships. Once relationships are formed, communication among existing partners can be facilitated through electronic channels. This mechanism has led to the argument that IT and face-to-face communication are complements to one another (e.g., Gaspar and Glaeser 1998, Charlot and Duranton 2006). However, in our setting, partnerships among researchers are likely set by the managers within the firm, so the benefits of geographic proximity to identifying research partners is less important than in an academic setting.

Our results have implications for the literature on knowledge diffusion within firms. Whereas evidence of the well-known stickiness of knowledge has been observed even across units within the same firm (e.g., Teece 1977, Szulanski 1996), collaborative ties have been found to be a very efficient way to transfer knowledge across branches, institutions, or industry boundaries (e.g., Singh 2005, Fleming et al. 2007). By providing robust evidence that IT investments can enable distant industrial R&D collaborations, and hence facilitate cross-unit integration through a decrease in coordination costs, the present study suggests that IT investments have the potential to reduce the well-known localization of knowledge flows.

There is an abundant body of research on the productivity of IT investments and more recently some work on the implications of IT investments for the growth in intangible assets like trademarks and patents (e.g., Gao and Hitt 2004, Kleis et al. 2012). However, because this latter work has focused on IT capital spending using firm-level data, it has been unable to unpack how IT investments lead to growth in intangibles. Our paper provides evidence that IT investments influenced coordination costs, but little evidence of improving productivity by lowering costs of access to shared resources or distant knowledge. This result has important implications for the design of research organizations within firms. In this way, we add to recent work in the IT productivity literature (e.g., Bloom et al. 2012) on the implications of different types of IT investment for business value and organizational design

Although our study only relies on U.S. data and on local capabilities, it has important implications for the study of the globalization of research. In designing their international R&D organization, firms are often thought to choose between a centralized organization that provides higher control but prevents access to local knowledge spillovers and a geographically dispersed and decentralized structure that enables tapping into local knowledge resources but induces higher coordination costs and more difficult knowledge sharing across firm units (e.g., von Zedtwitz and Gassmann 2002). By suggesting that Internet adoption can reduce coordination costs across distant R&D workers, our results suggest that IT investments may substantially alter this organizational trade-off and render decentralized R&D models more attractive, hence encouraging a higher geographic dispersion of R&D activities within firms.

From a managerial perspective, our results suggest that IT can be used to integrate geographically dispersed operations, either obtained through acquisition or deliberately dispersed because of a need to access local knowledge resources or markets. More broadly, they have implications for the long-run design of research organizations within firms. Our results suggest that firms that wish to disperse their research organizations to either capitalize on lower costs or on local capabilities can do so with the knowledge that these dispersed researchers can be linked through their IT investments.

Although our data are some of the best available, they are limited to one sample over one time period, therefore restricting the potential generalization of our conclusions. Future work may seek to understand how IT investments influence research collaborations in cross-country data. Extension to the cross-country context could have particularly interesting implications, because coordination costs will be higher while, simultaneously, the division of labor among researchers may be quite different (e.g., Zhao 2006). In addition, our study paves the way for further research on the effect of more advanced kinds of IT investments, such as those that facilitate social networking.

Furthermore, as noted above, our results raise several questions about the implications of IT investments for knowledge flows within organizations. Future work should examine whether new collaboration patterns enabled by IT have mediated new knowledge flows within organizations. More broadly, future research should examine to what extent IT investments have reduced or increased the importance of traditional channels of knowledge transfer, such as spatial, social, and employment relationships. We hope that our paper will help stimulate future work in these important areas.

Acknowledgments

The authors thank Ajay Agrawal, Marco Ceccagnoli, Avi Goldfarb, Shane Greenstein, Lee Fleming (the department editor), the associate editor, and three anonymous referees. They also thank participants at the 2011 National Bureau of Economic Research Summer Institute, the ZEW-KMRC Workshop on Social Software, the EPFL-ETHZ Conference on Knowledge in Organizations, the 8th ZEW Conference on the Economics of IT, the 2010 International Conference in Information Systems, the Information and Communication Technology (ICT) and Economic Growth Conference (Munich), and the 2011 Industry Studies Association Conference, and seminar participants at ECARES, Georgia State University, the University of Pennsylvania, the University of Washington, the University of Texas at Austin, and the University of Minnesota for helpful comments and suggestions. The authors thank Michael Cha for invaluable research assistance. They gratefully acknowledge funding support from the General Motors Strategy Center at the Kellogg School of Management and the Software Industry Center at Carnegie Mellon University. Chris Forman acknowledges funding from a Sloan Industry Studies Fellowship from the Alfred P. Sloan Foundation. Nicolas van Zeebroeck acknowledges funding from Fonds National de la Recherche Scientifique. They also thank Harte Hanks Market Intelligence for supplying data. All opinions and errors are the authors' own.

References

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

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

Argyres, N. S., B. Silverman. 2004. R&D, organization structure, and the development of corporate technological knowledge. *Strategic Management J.* **25**(8–9) 929–958.

Athey, S., S. Stern. 2002. The impact of information technology on emergency health care outcomes. *RAND J. Econom.* **33**(3) 399–432.

Augereau, A., S. Greenstein. 2001. The need for speed in emerging communications markets: Upgrades to advanced technology at Internet service providers. *Internat. J. Indust. Organ.* **19**(7) 1085–1102.

Augereau, A., S. Greenstein, M. Rysman. 2006. Coordination versus differentiation in a standards war: 56K modems. *RAND J. Econom.* **37**(4) 887–909.

Becker, G. S., K. M. Murphy. 1992. The division of labor, coordination costs, and knowledge. *Quart. J. Econom.* **107**(4) 1137–1160.

Bloom, N., R. Sadun, J. Van Reenen. 2012. Americans do IT better: US multinationals and the productivity miracle. *Amer. Econom. Rev.* **102**(1) 167–201.

Bloom, N., L. Garicano, R. Sadun, J. Van Reenen. 2009. The distinct effects of information technology and communication technology on firm organization. NBER Working Paper 14975, National Bureau of Economic Research, Cambridge, MA.

Bresnahan, T., S. Greenstein. 1996. Technical progress and co-invention in computing and in the uses of computers. *Brookings Papers Econom. Activity, Microeconomics* **1996** 1–83.

Bresnahan, T. F., E. Brynjolfsson, L. M. Hitt. 2002. Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence. *Quart. J. Econom.* **117**(1) 339–376.

Brynjolfsson, E., L. Hitt. 2003. Computing productivity: Firm-level evidence. *Rev. Econom. Statist.* **85**(4) 793–808.

Cairncross, F. 1997. *The Death of Distance*. Harvard University Press, Cambridge, MA.

Charlot, S., G. Duranton. 2006. Cities and workplace communication: Some quantitative French evidence. *Urban Stud.* **43**(8) 1365–1394.

Cohen, W., R. R. Nelson, J. P. Walsh. 2000. Protecting their intellectual assets: Appropriability conditions and why U.S. manufacturing firms patent (or not). NBER Working Paper 7552, National Bureau of Economic Research, Cambridge, MA.

Ding, W., S. Levin, P. Stephan, A. Winkler. 2010. The impact of information technology on academic scientists' productivity and collaboration patterns. *Management Sci.* **56**(9) 1439–1461.

Downes, T., S. Greenstein. 2007. Understanding why universal service obligations may be unnecessary: The private development of local Internet access markets. *J. Urban Econom.* **62**(1) 2–26.

Feldman, M. P., D. B. Audretsch. 1999. Innovation in cities: Science-based diversity, specialization, and localized competition. *Eur. Econom. Rev.* **43**(2) 409–429.

Fleming, L., D. Chen, S. Mingo. 2007. Collaborative brokerage, generative creativity, and creative success. *Admin. Sci. Quart.* **52**(3) 443–475.

Forman, C. 2005. The corporate digital divide: Determinants of Internet adoption. *Management Sci.* **51**(4) 641–654.

Forman, C., A. Goldfarb, S. Greenstein. 2002. Digital dispersion: An industrial and geographic census of commercial Internet use. NBER Working Paper 9287, National Bureau of Economic Research, Cambridge, MA.

Forman, C., A. Goldfarb, S. Greenstein. 2005. How did location affect the adoption of the commercial Internet? Global village vs. urban Density. *J. Urban Econom.* **58**(3) 389–420.

Forman, C., A. Goldfarb, S. Greenstein. 2008. Understanding the inputs into innovation: Do cities substitute for internal Firm resources? *J. Econom. Management Strategy* **17**(2) 295–316.

Forman, C., A. Goldfarb, S. Greenstein. 2012. The Internet and local wages: A puzzle. *Amer. Econom. Rev.* **102**(1) 556–575.

Friedman, T. 2005. *The World Is Flat: A Brief History of the Twenty-First Century*. (Farrar, Straus, and Giroux, New York).

Frost, T. S., J. Birkinshaw, P. Ensign. 2002. Centers of excellence in multinational corporations. *Strategic Management J.* **23**(11) 997–1018.

Furman, J. L., M. K. Kyle, I. Cockburn, R. M. Henderson. 2005. Public and private spillovers, location, and the productivity of pharmaceutical research. *Annales D'Economie et de Statistique* (79/80) 1–24.

Gao, G., L. Hitt. 2004. IT and product variety: Evidence from panel data. *Proc. 25th Internat. Conf. Inform. Systems (ICIS)*, Association for Information Systems, Atlanta.

Gaspar, J., E. Glaeser. 1998. Information technology and the future of cities. *J. Urban Econom.* **43**(1) 136–156.

Greenstein, S., M. Mazzeo. 2006. The role of differentiation strategy in local telecommunication entry and market evolution: 1999–2002 *J. Indust. Econom.* **54**(3) 323–350.

Griliches, Z. 1990. Patent statistics as economic indicators: A survey. *J. Econom. Literature* **28**(4) 1661–1707

Hall, B. H., R. H. Ziedonis. 2001. The patent paradox revisited: An empirical study in the U.S. semiconductor industry, 1979–1995. *RAND J. Econom.* **32**(1) 101–128.

Hall, B. H., A. Jaffe, M. Trajtenberg. 2001. The NBER patent citation data file: Lessons, insights, and methodological tools. NBER Working Paper 8498, National Bureau of Economic Research, Cambridge, MA.

Hall, B. H., A. Jaffe, M. Trajtenberg. 2005. Market value and patent citations. *RAND J. Econom.* **36**(1) 16–38.

Hamilton, B. H., J. A. Nickerson, H. Owan. 2003. Team incentives and worker heterogeneity: An empirical analysis of the impact of teams on productivity and participation. *J. Political Economy* **111**(3) 465–497.

Holmstrom, B. 1982. Moral hazard in teams. *Bell J. Econom.* **13**(2) 324–340.

Jaffe, A., M. Trajtenberg. 2002. *Patents, Citations, and Innovations: A Window on the Knowledge Economy*. MIT Press, Cambridge, MA.

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

Karau, S., K. Williams. 1993. Social loafing: A meta-analytic review and theoretical integration. *J. Personality Soc. Psych.* **65**(4) 681–706.

King, G., L. Zeng. 2001. Logistic regression in rare events data. *Political Anal.* **9**(2) 137–163.

Kleis, L., P. Chwelos, R. V. Ramirez, I. Cockburn. 2012. Information technology and intangible output: The impact of IT investment on innovation productivity. *Inform. Systems Res.* **23**(1) 42–59.

Kogut, B., U. Zander. 1993. Knowledge of the firm and the evolutionary theory of the multinational corporation. *J. Internat. Bus. Stud.* **24**(4) 625–645.

Latané, B., K. Williams, S. Harkins. 1979. Many hands make light the work: The causes and consequences of social loafing. *J. Personality Soc. Psych.* **37**(6) 822–832.

Lee, H., B. Choi. 2003. Knowledge management enablers, processes, and organizational performance: An integrative view and empirical examination. *J. Management Inform. Systems* **20**(1) 179–228.

Leiponen, A., C. E. Helfat. 2011. Location, decentralization, and knowledge sources for innovation. *Organ. Sci.* **22**(3) 641–658.

Macher, J., D. Mowery. 2008. *Innovation in Global Industries: U.S. Firms Competing in a New World*. National Academies Press, Washington, DC.

McElheran, K. 2011. Delegation in multi-establishment firms: Adaptation vs. coordination in IT purchasing authority. Working Paper 11-101, Technology and Operations Management Unit, Harvard Business School, Boston.

Rice, R. E. 1994. Relating electronic mail use and network structure to R&D work networks and performance. *J. Management Inform. Systems* **11**(1) 9–29.

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

Sambamurthy, V., R. W. Zmud. 1999. Arrangements for information technology governance: A theory of multiple contingencies. *MIS Quart.* **23**(2) 261–290.

Sauer, R. D. 1988. Estimates of the returns to quality and coauthorship in economic academia. *J. Political Econom.* **96**(4) 855–866.

Singh, J. 2005. Collaborative networks as determinants of knowledge diffusion patterns. *Management Sci.* **51**(5) 756–770.

Singh, J. 2008. Distributed R&D, cross-regional knowledge integration, and quality of innovative output. *Res. Policy* **37**(1) 77–96.

Sorenson, O., J. W. Rivkin, L. Fleming. 2006. Complexity, networks, and knowledge flow. *Res. Policy* **35**(7) 994–1017.

Stock, J. H., M. Yogo. 2005. Testing for weak instruments in linear IV regression. D. W. K. Andrews, ed. *Identification and Inference for Econometric Models*. Cambridge University Press, Cambridge, UK, 80–108.

Stock, J. H., J. H. Wright, M. Yogo. 2002. A survey of weak instruments and weak identification in generalized method of moments. *J. Bus. Econom. Statist.* **20**(4) 518–529.

Szulanski, G. 1996. Exploring internal stickiness: Impediments to the transfer of best practice within the firm. *Strategic Management J.* **17**(Special Issue) 27–43.

Teece, D. J. 1977. Technology transfer by multinational firms: The resource cost of transferring technological know-how. *Econom. J.* **87**(346) 242–261.

van Alstyne, M., E. Brynjolfsson. 2005. Electronic communities: Global village or cyber balkans? *Management Sci.* **51**(6) 851–868.

von Zedtwitz, M., O. Gassmann. 2002. Market versus technology drive in R&D internationalization: Four different patterns of managing research and development. *Res. Policy* **31**(4) 569–588.

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

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

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

Zhao, M. 2006. Conducting R&D in countries with weak intellectual property rights protection. *Management Sci.* **52**(8) 1185–1199.

Ziedonis, R. H. 2004. Don't fence me in: Fragmented markets for technology and the patent acquisition strategies of firms. *Management Sci.* **50**(6) 804–820.