Available online at www.sciencedirect.com
SCIENCE DIRECT®
Electronic Commerce Research and Applications 4 (2005) 1–13
Electronic Commerce Research and Applications
www.elsevier.com/locate/ecra

Geographic location and the diffusion of Internet technology

Chris Forman [a,*], Avi Goldfarb [b], Shane Greenstein [c]

[a] *Carnegie Mellon University, Tepper School of Business, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA*
[b] *University of Toronto, Rotman School of Management, 105 St. George Street, Toronto, Ontario M5S 3E6, Canada*
[c] *Northwestern University, Kellogg School of Management, 2001 Sheridan Road, Evanston, IL 60208, USA*

Received 20 September 2004; received in revised form 5 October 2004; accepted 11 October 2004
Available online 28 October 2004

Abstract

This study examines the sources of geographic variance in commercial Internet use. Until now, two opposing views have been argued on the relationship between Internet technology and economic agglomeration. One view, which we term *global village theory*, asserts that Internet technology helps lower communication costs and break down geographic boundaries between firms. The other view, labeled *urban density theory*, argues that the Internet follows a traditional pattern of diffusion – diffusing first through urban areas with complementary technical and knowledge resources that lower the costs of investing in new frontier technology. We provide a third view, *industry composition theory*, that asserts that demand for the Internet is increasing in location size because of the concentration of information-intensive firms in urban areas. We offer hard evidence on factors influencing the dispersion of Internet technology to businesses. We find no evidence for urban density theory in the diffusion of basic access and participation in the Internet network. We do find some evidence supporting global village theory for diffusion along this dimension. We also find that the pattern of adoption of frontier Internet technologies supports urban density theory not global village theory. Last, we show that business use of the Internet is significantly shaped by the prior geographic distribution of industry.

Keywords: Internet; Electronic commerce; Economic analysis; Adoption

1. Introduction

General Purpose Technology (GPT) theory predicts that the same technological opportunity does not result in the same commercial experience for all establishments in all locations. Such variance underlies the range of net benefits experienced across the US from the diffusion of the Internet. However, a lack of systematic data and formal theory has made it difficult to characterize variance in Internet experience among commercial users. With

* Corresponding author.
E-mail address: cforman@andrew.cmu.edu (C. Forman).

1567-4223/$ - see front matter
doi:10.1016/j.elerap.2004.10.005

a few notable exceptions, there has been little systematic empirical research on the diffusion of Internet technology to businesses in the US. [1]

This paper fills this gap with both a novel framework and novel data. We estimate probit models for Internet adoption as a function of features of a business establishment (i.e., its industry, location, and size and whether it is part of multi-establishment organization). Our analysis highlights why adoption decisions vary across locations.

There were many purposes for adopting the Internet in business. We contrast a simple purpose and a complex one. The first purpose, labeled *participation*, relates to activities such as email and web browsing. This represents a minimal and a simple use of the Internet for basic communications. The second purpose, labeled *enhancement*, relates to investment in frontier Internet technologies linked to computing facilities. These technologies are often known as "e-commerce," and involve complementary and complex changes to internal business computing processes.

Our econometric analysis tests the predictions of GPT theory. We highlight three key factors that induced variance in experience across locations and that support three theories that describe how these factors influenced Internet dispersion.

- *Overcoming agglomeration disadvantages.* Some researchers hypothesize that the Internet helped businesses bridge distances between geographically disparate economic actors. While all business establishments benefited from this increase in capabilities, geographically isolated businesses benefited comparatively more from their ability to transfer data and transact electronically in ways that were previously difficult or prohibitively expensive. We label this *global village theory*.
- *The user cost of the Internet.* Some researchers hypothesize that adoption, adaptation, and operations for the Internet were comparatively more expensive outside of urban areas. Internet infrastructure was more expensive, so too were the basic services – such as access fees, maintenance and development. We call this *urban density theory.* We further posit that complex applications of Internet technology are more sensitive to variation in geographic supply constraints than simple applications.
- *The geographic dispersion of IT-intense business.* Establishments locate in places for a wide variety of reasons. Many of these decisions were made prior to the diffusion of the Internet. Some of these establishments belonged to industries that already made heavy use of information-intensive technologies, while others did not. Because many of these firms clustered together in the same places, urban locations were coincident with high demand for the applications associated with new technology. We call this *industry composition theory.* It produces heterogeneous responses to newly available technology that are unrelated to location, and is essential for understanding whether new Internet technology becomes a substitute or complement to urban agglomeration.

This study examines Internet adoption at 86,879 establishments with over 100 employees. This data comes from a survey updated to the end of 2000. Harte Hanks Market Intelligence, a commercial market research firm that tracks use of Internet technology in business, undertook the survey. We use the County Business Patterns data from the Census and routine statistical methods to generalize our estimates to the entire population of medium to large establishments in the US. Approximately two-thirds of the US work force is employed in such establishments.

The results are consistent with GPT theory. First, participation is near saturation in a majority of industries and locations, which is evidence that benefits exceed costs for most establishments and is consistent with global village theory. Once industry composition is controlled for, we find further weak support for global village theory; in other words, there is little evidence that urban density theory shapes participation.

Second, the estimates for enhancement differ from participation in ways consistent with GPT theory. Controlling for industry composition,

[1] To be sure, geographers and government agencies have studied the diffusion of Internet technology at length, but largely among households. For statistical research about business use of the Internet, see [3,4], and [2], as well as the [11,13,30,31], and [23]. None have explored geographic variation in use within the US.

establishments in large and medium-sized metropolitan statistical areas (MSAs) are 1% more likely to use the Internet for enhancement, a significant increase given that only 12.6% of firms undertake enhancement. In other words, the urban location of an establishment does contribute to observed outcomes.

Finally, we show why enhancement and urban areas were complements over a time period in which Internet technology was diffusing. Our study is a short-run analysis that holds establishment locations fixed. While it is too soon to observe the long-run movement of establishments in reaction to this diffusion, our findings inform research about the relationship between information technology and the geographic location of firms (e.g. [17,22,28]).

We also offer a very different explanation for geographic variance in commercial Internet use from the prevailing literature on the digital divide. Mostly focusing on what we call participation, some authors argue that geographic usage patterns provide evidence of the existence of a digital divide (i.e., that the economic consequences from the Internet exacerbates regional inequalities, concentrating benefits in only a few locations). Our data show that this characterization is misleading for business use of the Internet. Business use of the Internet for enhancement is shaped in part by the prior geographic distribution of industry. Since many industries are not concentrated in a few locations, the Internet's advanced uses did not concentrate in a small number of areas. Instead, it dispersed to the vast majority of areas in the US.

2. Literature review

The present study builds upon our earlier work [14–16]. That research established that the variance in geographical patterns of technology adoption by business differed substantially from the variance uncovered in any existing research on households or infrastructure deployment. The present study differs in its focus on decomposing the factors underlying aggregate variance and in its explicit testing of hypotheses relating to these underlying factors.

Our investigation contributes novel findings to the debate about whether IT acts as a substitute or complement to the agglomeration of economic activity (e.g., [17]). Unlike some studies in this vein, we do not consider the determinants of long-run equilibrium, that is, where firms relocate after technology markets develop (e.g., [5,22]). Rather, we examine the short-run reaction of establishments to the diffusion of a technology.

Our empirical question of how location affects Internet practice is similar to [21]. Whereas [21] uses domain name registrations in the context of a periphery/central city model to find that users in cities of medium size and above have registration patterns consistent with those areas benefiting disproportionately from the Internet, we examine businesses' adoption of certain processes in cities of varying size and location.

Our study is also similar in spirit to [28]. They examine household behavior for evidence that Internet content is either a substitute or a complement to content found locally. When we examine substitutability or complementarity our focus is on adoption by business establishments, as opposed to household Internet use. Moreover, our decomposition of the geographic variation of Internet use into industry- and location-specific factors enables us to test hypothesizes unexamined by prior work.

In its focus on identifying the causes of the geographic variance in Internet adoption, the present study is broadly similar to recent studies completed by researchers at the Center for Research on Information Technology and Organizations (e.g., [32,23]). Their research focuses on characterizing the determinants of cross-country variation in e-commerce adoption. In contrast, we pursue a more focused approach of examining the causes of variation in Internet adoption within the US. This enables us to identify the effects of location size while holding constant factors such as aggregate wealth, governmental policies, and cultural differences that vary in cross-country studies.

Our analysis and findings contrast strongly with the prevailing analysis inspired by literature on the digital divide. For example, [33,34,18], among many others, examine the Internet in terms of "global city theory," where a few key locations act as economic hubs for other areas. They argue that the economic effects from the diffusion of the Internet concentrate in a small number of cities. [2] In contrast, we emphasize the sharp differences between the diffusion processes shaping participation and enhancement, both of which support very different explanations about the factors shaping geographic variation in use.

3. Theories of internet diffusion

The diffusion of the Internet can be viewed in the context of observations about technological convergence [1], which is the increasing use of a small number of technological functions for many different purposes. [7] develop this further in their discussion of GPTs, which they define as capabilities whose adaptation raises the marginal returns to inventive activity in a wide variety of circumstances. GPTs involve high fixed costs in invention and low marginal costs in reproduction. A GPT is adapted for any new use, and this adaptation takes time, additional expense and further invention. Following the literature, we label these as *co-invention expenses*. Studies have found that co-invention influences computing and Internet technology investments by business users [6,13].

We use GPT theory to construct a framework to understand business adoption of Internet technology. Consider an establishment's decision to adopt a new technology such as Internet access. Below we will consider both a simple and a complex application of the Internet to business. [3] Establishment i will adopt Internet technology by time t if

$$\mathrm{NB}(x_i, t) = B(x_i, t) - C(x_i, t) > 0,$$

where NB is the net benefit of adoption, B is the (gross) benefit of adoption, and C is the cost of adoption. We define x_i to be a vector of attributes distributed across the population of establishments. For example, x_i may describe variation in geographic conditions, industry, or prior investments.

We propose several theories on how location can influence the decision to adopt by time t. In these theories, x_i represents variables indicating the size of urban area and the density of population.

- *Global village theory* argues that adoption benefits are decreasing as urban density and size increases (i.e., $\mathrm{d}B/\mathrm{d}x_i < 0$, where x_i is density). Firms in areas with low population or low population density will benefit more from new Internet technology. This view has received considerable exposure in the trade press (e.g., [9,10]) but little empirical verification. [4] In its most basic form, global village theory argues that new Internet technology helps break down communication barriers between individuals and organizations, and these barriers are greatest for geographically isolated firms. That is, establishments in rural or small urban areas will derive the most benefit from Internet technology because their suppliers and customers are most likely to be located in other areas. Moreover, these areas will derive higher benefits because of a lack of substitute data communication technologies, such as fixed private lines. In other words, advanced Internet technology can substitute for many of the disadvantages associated with a location remote from the urban center of economic activity. In shorthand, Internet technology is a substitute for agglomeration.
- *Urban density theory* argues that adoption costs are decreasing as population size and density increases (i.e., $\mathrm{d}C/\mathrm{d}x_i < 0$, where x_i is density). [5] There are three major reasons for this hypothesis: (1) availability of complementary information

[2] See [19] for an extensive review of this literature.

[3] In our data analysis, we will consider adoption of two distinct layers of Internet technology. The analysis above applies to both layers of Internet investment, so for notational simplicity we consider only a generic application of technology in this section.

[4] One exception is [13], who examines a sample of service sector firms, and shows that rural firms and firms that are geographically dispersed are more likely to adopt Internet technology. Also see [26,27] for a close study of rural business Internet adoption and use.

[5] The value of some complementary resources can be viewed as affecting the gross benefits of adoption. For example, knowledge spillovers may be viewed as improving the benefits of adopting some e-commerce services. We view the value of such improvements as lowering the costs of achieving a certain level of service.

technology infrastructure; (2) labor market thickness; and (3) knowledge spillovers. These are closely related to the three major reasons given for industrial agglomeration (e.g., [35,24]). The availability of low-cost complementary inputs, such as broadband services or Internet access, will have a major impact on the cost of adopting participation or enhancement. Labor market thickness is closely related to complementary inputs. Adoption of new technologies often requires specialized technical skills not available within the firm and which must be procured either by hiring additional workers or through arms-length contracting. In other words, locating in an urban center acts as a complement to the use of advanced Internet technology. In shorthand, Internet technology is a complement to urban agglomeration.

- *Industry composition theory* asserts that demand for the Internet increases with urban density because information-intensive firms tend to inhabit more urban areas (i.e., $\mathrm{d}B/\mathrm{d}x_i > 0$, where x_i is density). As opposed to urban density theory, which has three different reasons for this hypothesis, industry composition theory relies on two premises. First, some establishments place higher value on information-intensive activities than others in ways that vary systematically across industry and establishment size. Second, establishments from the same industry tend to cluster in similar places for many reasons to take advantage of thicker labor markets and other shared local resources [24]. Concentration of Internet technology-intensive activity in one location could have little to do with the use of frontier Internet technology. If previous decisions to concentrate activity results in the clustering of some types of firms in urban areas, then it can result in a concentration of adoption of new Internet technology in urban areas. [6]

- *Simple versus complex applications of a GPT:* Microstudies of the adoption of computing technology argue that co-invention costs shape investments by business users [6]. Co-invention expenses arise when a GPT is adapted for commercial use. Co-invention requires high-quality labor inputs or third-party technological mediators that may cost more in sparsely populated areas than in urban areas.

Previous work (e.g., [13]) has shown that some simple Internet applications were viable as soon as the Internet commercialized. These simple applications were technologically mature, their technical costs of adoption were low, and they required little complementary innovation to be used successfully. In contrast, the complex applications that we label as "enhancements" required time, expense and invention to become viable with business users. To put it another way, these complex applications were at an earlier stage of diffusion. Because of their immaturity, they were harder to implement technically. More importantly, organizations were still learning how to use them. These complex technologies often involved considerable complementary innovation to be implemented successfully.

Therefore, co-invention theory forecasts that urban density theory is likely to be stronger for the diffusion of a complex application. That is, $\mathrm{d}C_\mathrm{C}/\mathrm{d}x_i < \mathrm{d}C_\mathrm{S}/\mathrm{d}x_i < 0$, where C_C is the cost for a complex application and where C_S is the cost for a simple application. In other words, the combination of urban density and co-invention theories suggest that adoption of a complex technology should be more sensitive to increases in population density than adoption of a simple one.

4. Data and method

The data we use for this study come from the Harte Hanks Market Intelligence CI Technology database (hereafter CI database). [7] The CI database contains establishment-level data on: (1)

[6] We note that variations in industry composition may also be complementary to the location-specific benefits detailed in urban density theory. Increases in the information intensity of firms may lead to entry by third-party IT service providers, increasing the marginal benefit of being located in a large and dense location. Such complementarity is very difficult to test in a cross-section. See [16] for further details.

[7] This section provides an overview of our methodology. For a more detailed discussion, see [14].

establishment characteristics, such as number of employees, industry and location; (2) use of technology hardware and software, such as computers, networking equipment, printers and other office equipment; and (3) use of Internet applications and other networking services. Harte Hanks Market Intelligence (hereafter HH) collects this information to resell as a tool for the marketing divisions at technology companies. Interview teams survey establishments throughout the calendar year; our sample contains the most current information as of December 2000.

Harte Hanks tracks over 300,000 establishments in the US. Since we focus on commercial Internet use, we exclude government establishments, military establishments and nonprofit establishments, mostly in higher education. Our sample contains all commercial establishments from the CI database that contain over 100 employees, 115,671 establishments in all; and HH provides one observation per establishment. We will use the 86,879 observations with complete data generated between June 1998 and December 2000. We adopt a strategy of utilizing as many observations as possible because we need many observations for thinly populated areas. [8] This necessitates routine adjustments of the data for the timing and type of the survey given by HH.

4.1. Sample construction and statistical method

Using two survey forms, HH surveyed establishments at different times. To adjust for differences in survey time and type, we econometrically estimate the relationship between an establishment's decision to adopt Internet technology as a function of its industry, location, timing of survey and form of survey. In the section below, we describe the econometric model using our measure of simple Internet technology (participation) as the dependent variable; the model for complex technology (enhancement) is exactly the same.

To be precise, our endogenous variable will be y_j. The variable y_j is latent. We observe whether the establishment makes a discrete choice when it, for example, chooses participation or enhancement. In either case, the observed decision takes on a value of either one or zero. We will define these endogenous variables more precisely below.

We assume that the value to an establishment j of participating in the Internet is

$$y_j = \sum_i \alpha_i d_{ij} + \sum_l \beta_l d_{lj} + \sum_t \gamma_t d_{tj} + \sum_{t>19905} \delta_t d_{tj} d_{pj} + \sum_p \phi x_{pj} + \varepsilon_j, \quad (1)$$

where d_{ij} and d_{lj} are dummy variables indicating the industry and location of the establishment, d_{tj} indicates the month in which the establishment was surveyed, and d_{pj} indicates whether the establishment responded to the long survey. [9] The variables x_j denote other location-specific and establishment-specific variables, such as population size, density, firm size, and whether the firm is a single- or multi-establishment firm, as well as interactions between these variables. If we assume the error term ε_j is i.i.d. normal, then the probability that establishment j participates is a probit. The probability of adopting enhancement can be written similarly.

We assume the adoption decisions of establishments are independent of one another. Such an assumption is questionable for multi-establishment firms in which a central decision-maker coordinates the decision to adopt for all of the establishments under his domain. As a robustness check, we estimated additional models in which we allowed the decision to adopt Internet technology to be a function of the decisions of other establishments within the same firm, and then measured whether this alters the estimate of the coefficient

[8] If we were only interested in the features of the most populated regions of the country, then we could easily rely solely on the most recent data from the latter half of 2000, about 40% of the data. However, using only this data would result in a very small number of observations for most regions with under one million in population.

[9] We interact the long survey dummy with time. This controls for endogeneity of survey. If establishments are selected for the long survey endogenously by HH, then the impact of receiving the long survey on adoption may vary over time.

on location, while instrumenting for decisions elsewhere. The results remain qualitatively the same.

5. Defining simple and complex applications

We identify two types of Internet applications used by establishments: *participation* and *enhancement*. [10] The first is simple and requires little co-invention. The second is complex and requires significant co-invention.

Participation. Participation is affiliated with basic communications, such as email use, browsing, and passive document sharing. It represents our measure of the minimal Internet investment required to do business on the Internet. We identify participation when an establishment has basic Internet access or has made any type of frontier investment. [11]

Enhancement. Enhancement, on the other hand, is affiliated with Internet technologies that either change existing internal operations or implement new services. [12] Enhancement is linked to the productive advance of firms and the economic growth of the regions in which these firms reside. It usually arrives as part of other intermediate goods, such as software, computing or networking equipment. Benefits accrue to the establishment that employs enhancement through the addition of competitive advantage, but the costs and delays of this activity vary widely.

Enhancement activity is less transparent than participation in the survey. We look for indications that an establishment must have made the type of investment commonly thought of as e-commerce. We identify enhancement from the presence of substantial investments in e-commerce or e-business applications. To provide confidence that we are measuring substantial investment, we look for commitment to two or more of the following projects: Internet-based enterprise resource planning or TCP/IP-based applications in customer service, education, extranet, publications, purchasing or technical support. [13, 14]

6. The dispersion of participation and enhancement

In this section, we seek to characterize differences in average participation and enhancement rates across industries and locations. We used the Census Bureau's 1999 County Business Patterns data to weight observations and obtain a representative sample.

In Table 1, we present average rates for participation and enhancement across types of locations in the US. Looking first at the country-wide figures, participation is widespread, and is undertaken by 88.6% of establishments. Consistent with GPT theory, levels of adoption of more complex enhancement technologies are lower (12.6%).

[10] An alternative strategy would be to treat Internet technology as fungible IT capital and employ an adoption measure that treated all Internet investment equally. Such a strategy would ignore the considerable heterogeneity in costs and benefits across purposes.

[11] To be counted as participating in the Internet, an establishment must engage in two or more of the following activities: (1) have an Internet service provider; (2) indicate it has basic access; (3) use commerce, customer service, education, extranet, homepage, publications, purchasing or technical support; (4) use the Internet for research or have an intranet or email based on TCP/IP protocols; (5) indicate there are Internet users or Internet developers on site; or (6) outsource some Internet activities. We looked for two or more activities to guard against "false positives." This was a minor issue as the vast majority of positive responses involved use of more than one of these criteria.

[12] See for example, [25,20], or [8].

[13] We tested slight variations on this threshold and did not find qualitatively different results.

[14] In brief, an establishment is counted as enhancing business processes when two or more hold: (1) the establishment uses two or more languages commonly used for web applications, such as Active-X, Java, CGI, Perl, VB Script, or XML; (2) the establishment has over five Internet developers; (3) the establishment has two or more "e-business" applications such as customer service, education, extranet, publications, purchasing, or technical support; (4) the establishment reports LAN software that performs one of several functions: e-commerce, enterprise resource planning, web development, or web server; (5) the establishment has an Internet server that is a UNIX workstation or server, mainframe, or minicomputer, or has five or more PC servers, or has Internet storage greater than 20 gigabytes; (6) the establishment answers three or more questions related to Internet server software, Internet/web software, or intranet applications. For a more precise description of some exceptional cases, see the appendix to [14].

Table 1
Average adoption by size of metropolitan statistical area

Population	Average participation	Average enhancement
Non-MSA	85.1%	10.6%
<250,000	75.5%	9.9%
250,000-1 million	84.9%	11.2%
>1 million	90.4%	14.7%
Nationwide	88.6%	12.6%

In Table 1, we further examine general tendencies by showing participation and enhancement rates across types of locations in the US. Table 1 shows sizable differences in participation and enhancement between large urban, small urban, and rural areas. On the surface, this evidence supports either urban density theory or industry composition theory. We see that large MSAs are somewhat exceptional, with an average participation rate of 90.4%. Participation rates in medium-sized MSAs and rural (non-MSA) areas are lower at 84.9% and 85.1%, respectively. In small MSAs the participation rates are even lower, 75.5% on average. [15]

The disparities in adoption rates for enhancement are even greater. Again large MSAs are somewhat exceptional. Establishments in large MSAs have adoption rates of 14.7%. In medium MSAs, it averages 11.2%. In small MSAs the rates are even lower, 9.9% on average. Average adoption rates in large MSAs are almost one-third greater than in medium MSAs. These averages suggest that the urban density theory or industry composition theory may hold. [16] To identify between them, we will need to identify the marginal effect of location while holding constant industry composition. We address this in the following section.

[15] From this point forward, MSAs with populations greater than 1 million will be referred to as *large MSAs*, those with between 250,000 and 999,999 will be *medium MSAs,* those with less than 250,000 will be *small MSAs*, and non-MSA areas will be called *rural*.

[16] For further details on the variance of average rates of Internet adoption, see [15].

7. The marginal impact of population concentration

7.1. The marginal impact of location

In this section, we continue to estimate Eq. (1), but our focus is now on estimating the parameters ϕ. We weight observations by the inverse probability that an establishment will appear in our sample. [17]

Table 2 presents the marginal effects from our probits. All probit regressions include dummy variables for three-digit North American Industry Classification System (NAICS), the month the data were collected, and whether the establishment was part of a multi-establishment firm. Employment and employment squared were also included as controls. Population was measured at the MSA level and density at the county level.

Participation. From Table 2, it is clear there is no support for the urban density theory. Controlling for industry and firm characteristics, location size and density has little impact on the decision to adopt at the participation level. If anything, the effects of location size and density support the global village theory. To be sure, the impact of geography is of limited economic and statistical significance, so the support is weak. For column 1, we use non-urban state areas (i.e., non-MSAs) for the base, and the results in Table 2 show that medium and large MSAs are 0.5–1.0% less likely to have adopted participation by the end of 2000. However, the effect is only significantly different from rural areas for medium MSAs. Moreover, this effect is only of marginal economic significance as participation rates average 88.6%.

In column 2, we identify the effects of size through a variable that captures the effects of increases in population in urban areas. Increases in population size decrease the probability of participation, though the effects are statistically insignificant. In column 4, we include dummies for

[17] This weighting corrects for potential sampling bias created by differences between the Harte Hanks data sample and the true population distribution in Census County Business Patterns (CBP) data. We use “sampling weights” as described in the User’s Guide to [29]. For further details on differences between the Harte Hanks data and CBP data, see [14].

Table 2
Marginal effects from weighted probit regressions (standard errors in parentheses)

	Participation				Enhancement			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Small MSA	0.0021				0.0035			
	(0.0064)				(0.0062)			
Medium MSA	−0.011				0.008			
	(0.0052)*				(0.0048)***			
Large MSA	−0.0058				0.0110			
	(0.0045)				(0.0039)**			
MSA population		−8.89e − 10	3.06e − 09			−1.54e − 10	3.67e − 09	
		(8.56e − 10)	(2.30e − 09)			(5.67e − 10)	(2.02e − 09)***	
MSA population squared			−4.37e − 16				−4.18e − 16	
Medium-low density			(2.55e − 16)***				(2.17e − 16)***	
Medium-high density				−0.00385				0.00485
				(0.00440)				(0.00399)
				−0.00639				0.0154
				(0.00456)				(0.00450)**
High density				−0.00400				0.0103
				(0.00508)				(0.00406)*

Notes. Standard errors are in parentheses.
(1) non-MSA is the base for these regressions.
(1) and (3) Since no meaningful population data was available for non-MSA areas, we include a "rural area" dummy variable in each of these regressions. The population and density variables were interacted with (1-RURAL). Therefore, the coefficients on the population variables do not include non-MSA areas.
(4) low density is the base for these regressions. One quarter of the observations fit into each density type.
* Significant at 95% confidence level.
** Significant at 99% confidence level.
*** Significant at 90% confidence level.

population density. This alternative specification gives very similar results. Variation in population density does not affect participation by more than 1%, and is always statistically insignificant.

Enhancement. The effects of population size and density on enhancement support the urban density theory. Column 5 in Table 2 shows that establishments in medium and large MSAs are 0.8–1.1% more likely to adopt enhancement. Column 8 shows that establishments in medium- and high-density regions are 1.5–1.0% more likely to adopt enhancement. All of these effects are statistically significant. They are economically significant in light of the average enhancement rates of 12.6%.

The contrast between participation and enhancement is informative. Studies of the adoption of computing technology argue that co-invention costs shape investments by business users [6]. Co-invention expenses arise when a GPT is adapted for commercial use. In the case of the Internet, co-invention involved the local demands of business establishments and the supply constraints of markets for Internet technology infrastructure and services. Co-invention theory forecasts that urban density theory is likely to be stronger for the diffusion of a complex application, a hypothesis that is supported in our data.

7.2. The role of industry composition

The differences between Tables 1 and 2 show that the effects of location on participation and enhancement fall if controls for establishment size, industry, and firm status are included. This suggests that although the global village and urban density theories may explain some of the variance

Table 3
Contribution of industry and location to explaining adoption decisions

	Participation		Enhancement	
	Pseudo R^2	Log likelihood	Pseudo R^2	Log likelihood
Full model	0.2339	−33093.4	0.0672	−28443.4
No MSA dummies	0.2251	−33475.0	0.0591	−28701.4
No NAICS dummies	0.1526	−36604.2	0.0347	−29434.6

Notes. Cities defined by CMSA. All regressions include establishment controls for size and multi-establishment status. The third row includes the regression with these controls plus location (MSA) dummies only, the second row has industry (NAICS) dummies only, and the first row has both industry and location dummies.

Table 4
Importance of industry and population on enhancement

	(1) Enhancement	(2) Enhancement	(3) Enhancement	(4) Enhancement
(a) Percent firms in top quartile	0.259***	0.214***	0.200***	0.212***
	(0.0291)	(0.0388)	(0.0498)	(0.0612)
(b) MSA population greater than 1 million		0.0127*		0.0189
		(0.00719)		(0.0221)
(c) MSA population between 250,000 and 1 million		0.00331		0.000163
		(0.00479)		(0.0132)
(a)*(b)			0.0560	−0.0250
			(0.0360)	(0.105)
(a)*(c)			0.0238	0.0199
			(0.0322)	(0.0876)
Constant	0.0706***	0.0739***	0.0758***	0.0742***
	(0.00496)	(0.00538)	(0.00598)	(0.00768)
R-sq	0.1995	0.2074	0.2060	0.2079

Notes. Standard errors in parentheses. Percent firms in top quartile is a measure of the percentage of firms in the top quartile of enhancement adopters.
* Significant at 90% confidence level.
** Significant at 95% confidence level.
*** Significant at 99% confidence level.

in simple and complex Internet technology adoption, industry composition is also supported.

In fact, industry composition explains much more of the variation in participation and enhancement rates than location. Once industry is controlled for, the incremental contribution of location in the probit regressions is small. This is shown in Table 3. The pseudo-R^2 of a probit for participation including location dummies only (no NAICS dummies) is 0.1526, whereas the pseudo-R^2 of a probit with industry dummies only (no MSA dummies) is 0.2251. [18] Adding location dummies to a probit that includes industry dummies improves the fit only marginally, from 0.2251 to 0.2339. Enhancement displays a similar pattern.

For industry composition theory to be true, however, leading industries must also be concentrated in large urban areas. Table 4 provides a partial test of this proposition. [19] It shows the results from a simple regression that predicts enhancement for an MSA, illustrating the effect of area size and industry presence. In this regression, we include a variable that indicates the percentage of establishments from the top quartile of enhancement adopters. Though this measure of industry selects on the endogenous variable, we note that the list of industries that are leading adopters of

[18] We continue to include establishment controls for size and multi-establishment firm status in these models.

[19] For further analysis of how industry composition shapes the geographic variance in adoption, see [16].

IT have changed little over time [12,14]. The coefficient shows that a 10% drop in the percent of firms from leading industries (from e.g., 30% to 20%) would lead to a 2% drop in the enhancement rate within an MSA. The importance of industry continues to come through even with the addition of MSA size effects and interaction terms. [20] In sum, an area is advanced because its establishments happen to come from a leading industry. To be fair, the presence of leading industries is not the only factor, but it is an important one. It alone explains 20% of the variance in enhancement.

Table 4 provides a foundation for an intuitively appealing observation – urban areas are comprised of establishments with a disproportionate tendency to be information-intensive. To be concrete, within large MSAs, 27.5% of establishments are in industries that are part of the top quartile of adopters, compared to 19.0% of establishments in small urban areas. The geographic dispersion of establishments from these industries favored large urban areas prior to the diffusion of the Internet and largely contributed to higher rates of participation and enhancement in large urban areas.

8. Conclusions

These findings contribute to the debate of whether new IT is a complement or substitute to urban agglomeration. We infer that geographically isolated establishments appeared to enjoy higher gross benefits because of initially high inter-organizational communication costs and few communications substitutes. Because the technology behind participation was well developed and required little co-invention, there was little variation in costs across different geographic areas. By 2000, participation had spread geographically throughout the country. The relationship between participation and agglomeration was, if anything, one of substitutes.

[20] We tried a number of variations on the same type of regression, with similar qualitative results. Hence, we show the simplest result here.

In contrast, enhancement technologies were complex and required substantial co-invention to be used successfully. We infer that enhancement and urban areas were complements over a time period in which Internet technology was diffusing. Establishments located in urban areas benefited from a combination of benefits, including thicker labor markets, availability of complementary inputs, and knowledge spillovers. Thus, adoption costs were lower in urban areas.

Although location played an important role in the diffusion of participation and enhancement, we showed that over the short run its role was secondary to that of industry. Among major urban areas, we observed a large number of locations with little difference in the extent to which specific areas contributed, on the margin, to the adoption of frontier technology. There appears to be a sense in which many urban areas are meaningful substitutes for each other in their contribution to the use of advanced technology.

8.1. Implication for managers

For practitioners, these results suggest that Internet technology may be useful in lowering coordination costs associated with distance. We showed that firms located in rural areas were somewhat more likely to adopt simple Internet technologies. However, reports of the dissolution of the city may be premature. We showed that establishments located in large urban areas had access to complementary resources that increased the likelihood of adopting complex Internet technologies. This suggests that managers of IT-intensive firms should continue to consider the value of such complementary resources when making firm location decisions. It remains for future work to sort out how these competing forces shape how new IT impacts firm location decisions.

Moreover, these results suggest that managers need to readjust their expectations for the speed of diffusion of complex technologies such as enhancement. They suggest that the rapid diffusion of early Internet technologies was caused by unusually high net benefits of adoption and low co-invention costs. Adopters of enhancement face some of the usual barriers to diffusion: technical costs, missing complementary technologies, learning costs, and co-invention. Demanders of Internet technologies should recognize this will likely translate into slower diffusion of e-commerce technologies among partners up and down the supply chain. For suppliers of Internet hardware and software, this will translate into lower projections for firm growth.

Acknowledgements

We thank Tim Bresnahan, Rob Kauffman, Steve Klepper, Roger Noll, Alicia Shems, Scott Stern, Manuel Trajtenberg, three anonymous referees, and participants at the International Conference on Electronic Commerce for comments. We also thank Harte Hanks Market Intelligence for supplying data. We received funding from the Kellogg School of Management and the GM Strategy Center and seed funding from the National Science Foundation and Bureau of Economic Analysis. All opinions and errors are ours alone.

References

[1] Edward Ames, Nathan Rosenberg, Technological change in the machine tool industry, 1840–1910, in: Nathan Rosenberg (Ed.), Perspectives on Technology, second ed., Cambridge Press, Cambridge, MA, 1984.

[2] Barbara K. Atrostic, John Gates, US productivity and electronic business processes in manufacturing, Working paper #01-11, Center for Economic Studies, United States Census Bureau, 2001.

[3] Barbara K. Atrostic, John Gates, Ron Jarmin, Measuring the electronic economy: current status and next steps, Working paper #00-10, Center for Economic Studies, United States Census Bureau, 2000.

[4] Barbara K. Atrostic, Sang V. Nguyen, Computer networks and US manufacturing plant productivity: new evidence from the CNUS data, Working paper #02-01, Center for Economic Studies, United States Census Bureau, 2002.

[5] M. Beardsell, V. Henderson, Spatial evolution of the computer industry in the USA, European Economic Review 43 (2) (1999) 431–456.

[6] Timothy Bresnahan, Shane Greenstein, Technical progress in computing and in the uses of computers. Brookings papers on economic activity, Microeconomics (1996) 1–78.

[7] Timothy Bresnahan, Manuel Trajtenberg, General purpose technologies: engines of growth?, Journal of Econometrics 65 (1) (1995) 83–108.

[8] Timothy Bresnahan, Erik Brynjolfsson, Lorin Hitt, Information technology, work organization, and the demand for skilled labor: firm-level evidence, Quarterly Journal of Economics 117 (February) (2002) 339–376.

[9] Frances Cairncross, The Death of Distance, Harvard University Press, Cambridge, MA, 1997.

[10] Manuel Castells, The Internet Galaxy, Reflections on the Internet, Business and Society, Oxford University Press, Oxford, UK, 2001.

[11] Census, Detailing Tabulations of Manufacturing E-Business Process Use in 2000, Survey Release from United States Department of Commerce, 2002. Available from: <http://www.census.gov/epcd/www/ebusiness.htm>.

[12] James W. Cortada, Information Technology as Business History: Issues in the History and Management of Computers, Greenwood Press, Westport, CT, 1996.

[13] Chris Forman, The Corporate Digital Divide: Determinants of Internet Adoption, Mimeo, Tepper School of Business, Carnegie Mellon University, 2002.

[14] Chris Forman, Avi Goldfarb, Shane Greenstein, Digital dispersion: an industrial and geographic census of commercial internet use, NBER working paper #9287, 2002.

[15] Chris Forman, Avi Goldfarb, Shane Greenstein, The geographic dispersion of commercial internet use, in: Lorrie Faith Cranor, Steven S. Wildman (Eds.), Rethinking Rights and Regulations: Institutional Responses to New Communications Technologies, MIT Press, Cambridge, MA, 2003.

[16] Chris Forman, Avi Goldfarb, Shane Greenstein, How did location affect adoption of the commercial internet? Global village, urban density and industry composition, NBER working paper #9979, 2003b.

[17] J. Gaspar, E. Glaeser, Information technology and the future of cities, Journal of Urban Economics 43 (1998) 136–156.

[18] Sean P. Gorman, Where are the Web Factories: the Urban Bias of E-business Location, Mimeo, Neoteric Media Inc., Arlington VA, 2002.

[19] Shane Greenstein, The economic geography of internet infrastructure in the US, in: M. Caves, S. Majumdar, I. Vogelsang (Eds.), The Handbook of Telecommunications Economics, II, Elsevier, Amsterdam, Netherlands, 2003.

[20] Tom Hubbard, The demand for monitoring technologies: the case of trucking, Quarterly Journal of Economics 116 (May) (2000) 533–560.

[21] Jed Kolko, The death of cities? The death of distance? Evidence from the geography of commercial internet usage, in: Ingo Vogelsang, Benjamin M. Compaine (Eds.), The Internet Upheaval, MIT Press, Cambridge, MA, 2000.

[22] Jed Kolko, Silicon mountains, silicon molehills, geographic concentration and convergence of internet industries in the US, Information Economics and Policy 14 (2) (2002) 211–232.

[23] Kenneth Kraemer, Jason Dedrick, Debora Dunkle, E-Commerce: a mile wide and an inch deep, Working paper,

Center for Research on Information Technology and Organizations (CRITO), University of California, Irvine, 2002.

[24] Paul Krugman, Geography and Trade, MIT Press, Cambridge, MA, 1991.

[25] Thomas Malone, Joanne Yates, Robert Benjamin, Electronic markets and electronic hierarchies, Communications of the ACM 30 (6) (1987) 484–497.

[26] G. Premkumar, Rural business and telecommunications technologies, in: Peter F. Korsching, Patricia C. Hipple, Eric A. Abbott (Eds.), Having All the Right Connections: Telecommunications and Rural Viability, Praeger, New York, 2000.

[27] G. Premkumar, Margaret Roberts, Adoption of new information technologies in rural small business, Omega – International Journal of Management Science 27 (1999) 467–484.

[28] Todd Sinai, Joel Waldfogel, Geography and the internet: is the internet a substitute or a complement for cities? Working paper, Wharton School of Business, University of Pennsylvania, Philadelphia, 2001.

[29] Stata, User's Guide, Version 7, Stata Press, College Station, 2001.

[30] Hal Varian, Robert E. Litan, Andrew Elder, Jay Shutter, The Net Impact Study, 2001. Available from: <http://www.netimpactstudy.com>.

[31] Andrew Whinston, Anitesh Barua, Jay Shutter, Brant Wilson, Measuring the Internet Economy, 2001. Available from: <http://www.internetindicators.com/jan_2001.pdf>.

[32] Kevin Zhu, Kenneth Kraemer, Sean Xu, A cross-country study of electronic business adoption using the technology-organization-environment framework, in: L. Applegate, R. Galliers (Eds.), Proceedings of the 23rd International Conference on Information Systems, 2002, pp. 337–348.

[33] Matthew Zooks, Internet metrics: using hosts and domain counts to map the internet globally, Telecommunications Policy 24 (2000) 6/7.

[34] Matthew Zooks, The web of production: the economic geography of commercial internet content production in the US, Environment and Planning 32 (2000) 411–426.

[35] Alfred Marshall, Principles of economics, 8th ed., Porcupine Press, New York, 1920.