

NBER WORKING PAPER SERIES

HOW DID LOCATION AFFECT ADOPTION
OF THE COMMERCIAL INTERNET?
GLOBAL VILLAGE, URBAN DENSITY, AND INDUSTRY COMPOSITION

Chris Forman
Avi Goldfarb
Shane Greenstein

Working Paper 9979
<http://www.nber.org/papers/w9979>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2003

We thank Tim Bresnahan, Karen Clay, Gilles Duranton, Steve Klepper, Roger Noll, Alicia Shems, Manuel Trajtenberg, seminar participants at the Center for Economic Studies and the ZEW Conference on Information and Communication Technologies, and especially Scott Stern 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. The views expressed herein are those of the authors and are not necessarily those of the National Bureau of Economic Research.

©2003 by Chris Forman, Avi Goldfarb, and Shane Greenstein. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

How did Location Affect Adoption of the Commercial Internet? Global Village, Urban Density and Industry Composition

Chris Forman, Avi Goldfarb, and Shane Greenstein

NBER Working Paper No. 9979

September 2003

JEL No. L63, L86, R0

ABSTRACT

The authors test opposing theories on how urban locations influenced the diffusion of Internet technology. They find evidence that, controlling for industry, participation in the Internet is more likely in rural areas than in urban areas. Nevertheless, talk of the dissolution of cities is premature. Frontier Internet technologies appear more often at establishments in urban areas, even with industry controls. Major urban areas also contain many establishments from information technology-intensive industries, whose presence could reinforce the concentration of frontier Internet technologies in these areas. However, information technology-intensive industries are numerous and widespread. Hence, so is the use of frontier technology.

Chris Forman
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213
cforman@andrew.cmu.edu

Avi Goldfarb
University of Toronto
105 St George Street
Toronto ON Canada M5S 3E6
avi.goldfarb@rotman.utoronto.edu

Shane Greenstein
J. L. Kellogg School of Management
Leverone Hall
2001 Sheridan Road
Evanston, IL 60208-2013
and NBER
greenstein@kellogg.northwestern.edu

1. Introduction

While information technology (IT) spending grew rapidly from 1990 to 1995, the timing and form of the Internet's commercialization caught all but a small number of technical cognoscenti by surprise. The vast majority of business establishments were then faced with a new and unexpected technology. They had to decide how to react to the availability of new capabilities.

American businesses reacted with the largest growth rates in investment in IT in the history of the United States. Stocks of information technology capital grew at a 20% annual growth rate from the end of 1995 to the end of 2000.¹ By 2000, computer hardware and software stocks had reached \$622.2 billion.² The majority of this investment was affiliated with enabling business applications. In 2000, for example, total business investment in IT goods and services was almost triple the level for personal consumption of similar goods.³

The level of these investment flows is immense and so is the variance across locations.⁴ In some locations, the Internet has been adopted across all facets of economic activity, while in other locations adoption is not widespread. Two competing theories for explaining this variance have received the most attention. First, "global village" theory predicts that firms in small cities and rural areas adopted the Internet more quickly than urban firms because the marginal returns to the use of the communications capabilities of the Internet are higher in remote locations or in locations lacking economies of density. In contrast, the "urban density" theory predicts adoption will be most extensive in big cities because such locations allow for the pooling of resources, which lowers the costs of adoption.

Despite policy interest and much speculation about the pattern of diffusion, little empirical research confirms or refutes the competing hypotheses. This paper is the first to adjudicate this debate with hard data. Specifically, our study concentrates on a single, major, well-defined, and reasonably well-recorded development—adoption of business applications of the Internet—to test between competing theories about the geographic diffusion of Internet technology. In contrast to previous research on use of the Internet by workers, we focus on adoption by an establishment. This is more appropriate for comparing the marginal contribution of location, industry, and application complexity.

Our estimates examine decision-making among the largest investors in IT in the U.S. economy: Approximately two-thirds of the United States work force is employed in the type of establishments studied. Specifically, we analyze Internet adoption at 86,879 establishments with over one hundred employees; this sample comprises roughly half of U.S. establishments of such size. The data come from

¹ This includes computer hardware, computer software, and communications hardware and instruments. See Price and McKittrick (2002) or Henry and Dalton (2002). The growth rates are even higher if communications hardware and instruments are excluded.

² These are constant (1996) dollars. See Henry and Dalton (2002).

³ For 2000, estimated personal consumption of IT goods and services was \$165 billion. For business it was \$466 billion. See Henry and Dalton (2002).

⁴ For an extensive description of this variance, see our previous paper, Forman, Goldfarb and Greenstein (2002).

a survey updated to the end of 2000. Harte Hanks Market Intelligence (hereafter HH), a commercial market research firm that tracks use of Internet technology in business, undertook the survey. We separate these data into two categories of adoption— *participation*, which refers to adoption of simple applications, and *enhancement*, which refers to adoption of complex applications requiring technical support and third-party servicing.

Our earlier work (Forman, Goldfarb and Greenstein 2002) establishes that the variance in geographical patterns of participation and enhancement differ substantially from the variance uncovered in any existing research on Internet adoption by households or infrastructure deployment. While we have already shown that there is variance in adoption rates, here we focus on decomposing the factors underlying aggregate variance and explicitly test hypotheses relating to these underlying factors. On the basis of our results we provide input on the debate concerning patterns of Internet diffusion. Our research supports three central findings:

1. Urban density theory is not supported when the Internet is adopted for *participation* purposes. In other words, adoption does not increase with the size and density of a city. In fact, there is some evidence that the opposite holds. Adoption of participation is closer to the more utopian *global village theory* that is associated with the promise of the Internet.
2. We reach the opposite conclusion for *enhancement*: In this case, global village theory is not supported and urban density theory is, because the probability that an establishment will adopt enhancement applications increases with the population and density of location.
3. Finally, the type of industry found in major cities plays a key role in explaining the variance of adoption rates between cities: We label this *industry composition theory*. More IT-intensive industries tend to cluster in urban areas. Urban density theory and industry composition theory interact in a complementary way for enhancement applications. This interaction could exacerbate agglomeration in use. However, because IT-intensive industries are numerous and geographically dispersed, frontier technology use is also widespread. We find little evidence of agglomeration among participation applications.

Our study is pertinent to previous studies, yet it provides a new dimension to and novel data on the way adoption processes and business location interact. In addition to bringing hard data to bear on the geographic variance of Internet use, 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., Gaspar and Glaeser 1998). 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., Beardsell and Henderson 1999; Kolko 2002). Rather, we examine the short-run reaction of establishments to something new. This situation avoids the largest difficulty in previous empirical research, the simultaneous determination of industrial location, adoption decisions and supplier activity.

Our empirical question of how location affects Internet practice is similar to Kolko (2000). He 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 the adoption by businesses of certain processes in cities of varying size and location.

Our study is also similar in spirit to Sinai and Waldfogel (2001). 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 hypotheses unexamined by prior work.

We ultimately accept hypotheses based on the formulation that the Internet is a General Purpose Technology, or GPT (Bresnahan and Trajtenberg 1995). Although others have hypothesized that the Internet is a GPT (e.g., Harris 1998), we are the first to link this formulation to a statistical analysis of adoption behavior by commercial businesses. The GPT framework underscores our distinction between participation and enhancement, both of which rely, to different degrees, on local market-based support. In our characterization of how local pooled resources influence the returns to adopting enhancement applications, our study is similar to adoption studies that examine how knowledge spillovers (e.g., Goolsbee and Klenow 2002) and variations in education and skill level (e.g., Chun 2003; Doms, Dunne and Troske 1997) affect demand for new technologies.

Our analysis and findings contrast strongly with the prevailing analysis inspired by literature on the digital divide. Even though, like previous studies, we find that some regions are leaders and some are laggards in the use of Internet technology, we do not conclude that use of the Internet is concentrated in a small number of places. Moreover, 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.

In Sections 2 and 3, we describe the theoretical framework, data, and empirical method. In Section 4, we show that there is regional variation in Internet adoption rates. In Sections 5 and 6, we separate out the relative importance of global village theory, urban density theory, and industry composition theory in driving this variation. In Section 7, we conclude with observations about how our findings shape the analysis of the economic returns from the diffusion of Internet technology.

2. Framework and Testable Hypotheses

Why are technologies, such as the Internet, adopted at different rates by different firms? We offer a framework that formalizes theories about how the benefits and costs of adoption vary across establishments in different locations.

We consider an establishment's decision to adopt a new technology such as Internet access or interfirm networking capabilities. For notational simplicity, in this section we consider only a generic application of technology. Later in our discussion we will consider adoption of both a specific participation and enhancement application.⁵ Establishment i will adopt Internet technology by time t if $NB(x_i, z_i, t) \equiv B(x_i, z_i, t) - C(x_i, z_i, t) > 0$,

where NB is the net benefit of adoption, B is the gross benefit of adoption, and C is the cost of adoption. We let x_i describe geographic conditions, such as population size and density, while z_i describes industry characteristics that may affect a firm's decision to adopt Internet technology. Both features are fixed over time for establishments.

Our empirical work examines one cross-section at time t . Since adoption of the Internet is rarely reversed, we are comfortable with the data-driven requirement that we suppress the time dimension. Under the standard "probit model" of diffusion (e.g., David 1969), adoption costs decline over time for all potential adopters, so the timing of adoption coincides with intensity of demand. We employ the standard analysis for a cross-section at any point in time: The difference between adoption and non-adoption reveals the threshold between those with high and low valuations from use.⁶

We now discuss different theories about $NB(x_i, z_i)$, each of which gives competing predictions about the role of location in the adoption of the Internet by commercial establishments. Table 1 lists each of these hypotheses and their predictions.

2.1 Global village versus urban density

Global village theory argues that gross adoption benefits (1) increase as population size and density decrease (i.e., $dB/dx_i < 0$, where x_i is population size or density) and (2) increase more slowly than costs ($dC/dx_i < -dB/dx_i$ for all x_i). Together, net benefits from adoption decrease as population size and density increase ($dNB/dx_i < 0$). While the theory is implausible for sufficiently isolated establishments (e.g., in the Mohave desert), its proponents generally focus on the predictions for existing establishments in most urban and low-density settings. Global village theory predicts that

⁵ Our data analysis will consider adoption of two distinct layers of Internet technology. It is easily generalizable to more than two. In fact, Forman (2002) finds that participation and enhancement applications tend to be adopted in clusters.

⁶ Generally, see Rogers (1995). We allow the cost term C to include the opportunity cost of not adopting at some other time $s > t$, thus the net benefit condition above is both necessary and sufficient for the establishment to adopt by t . Another standard formulation would examine an establishment's decision to adopt at time t , and the equation above would be supplemented by an "arbitrage condition" (Ireland and Stoneman 1986) that it is less beneficial to adopt at any other time $s \neq t$.

adoption of the Internet will be more common among establishments in rural areas than in urban areas, all other things being equal.

Global village theory depends on two observations. First, while all business establishments benefit from an increase in capabilities, establishments in rural or small urban areas derive the most benefit from overcoming diseconomies of small local size. That is, Internet technology substitutes for the disadvantages associated with a remote location. Second, establishments in rural areas lack substitute data communication technologies for lowering communication costs, such as fixed private lines.

The most utopian versions of this hypothesis have received considerable exposure.⁷ This hypothesis also predicts that advances in information and communication technology will help decrease concentration in economic activity. Yet, even the more tempered forecasts are generally treated with skepticism by most academic research on the geography of the Internet. This hypothesis has not been directly tested and has not had much empirical verification (Forman's (2002) study is an exception).

Urban density theory stands in opposition to global village theory. Urban density theory argues that adoption costs increase as population size and density decrease (i.e., $dC/dx_i < 0$, where x_i is density) and that these costs increase faster than benefits increase ($dC/dx_i > -dB/dx_i$ for all x_i). Together, net benefits of adoption increase as population size and density increase ($dNB/dx_i > 0$). Urban density theory predicts that adoption of the Internet will be less common in rural areas than in urban areas, all other things being equal.

There are three major factors supporting this hypothesis: (1) availability of complementary information technology infrastructure, (2) labor market thickness for complementary services or specialized skills, and (3) knowledge spillovers. These are closely related to the three major reasons given for industrial agglomeration (e.g., Marshall 1920; Krugman 1991). For example, the availability of low-cost complementary inputs, such as broadband services or Internet access, will have an impact on the cost of adoption.⁸

Urban density theory has received considerable exposure, partially because the literature on the urban/rural digital divide highlights this concern. For example, both Gorman (2002) and Zooks (2000a, b), summarizing the conclusions of many others as well as their own research, express this view quite strongly (see, also, Castells 2003).⁹ They also argue that Internet technology is a complement to urban agglomeration. (For a full review of this literature, see Greenstein, 2003).

⁷ Most discussion of this hypothesis follows from Cairncross (1997), who was an early proponent of global village theory. She states, "The death of distance as a determinant of the cost of communicating will probably be the single most important force shaping society in the first half of the next century" (p. 1).

⁸ By this time period, almost all but the poorest and most remote geographic areas were serviced by dial-up Internet Service Providers (Downes and Greenstein 2002). Yet, broadband access was disproportionately an urban technology (U. S. Department of Agriculture 2001; Crandall and Alleman 2002)

⁹ This is a particularly prominent theme in the series of reports from the National Telecommunications and Information Administration (1995, 1998, 1999, 2000a, 2000b). See, also, Moss and Townsend (1997).

In comparing global village and urban density theories, we further contrast participation purposes for business use of the Internet with enhancement purposes.¹⁰ Participation is affiliated with basic communications, such as email use, browsing, and passive document sharing. It represents our measure of the minimal investment required to do business on the Internet. It is emphasized in many studies of “universal service” in new technologies. Geographic differences in participation, such as urban/rural divisions, are important drivers of policy decisions in this area.¹¹

Enhancement, on the other hand, is affiliated with IT that either changes existing internal operations or implements new services. It is related to investment in frontier technologies linked to computing facilities, which are often known as “e-commerce” or “e-business.” 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. The costs and delays of this activity vary, and enhancement investments are the most idiosyncratic and difficult.¹²

The contrast between our participation and enhancement results is informative about local adaptation costs. Bresnahan and Trajtenberg (1995) argue that the invention of a GPT like the Internet involves high fixed costs. Users typically must adapt the GPT to their particular environments. Adoption involves a combination of (1) reproduction and (2) co-invention to meet idiosyncratic circumstances. Co-invention activity is not just R&D because it involves change to organizational processes. When requirements and constraints are idiosyncratic and either technically or organizationally complex, co-invention issues are expensive to address. When technical and organizational issues are generic and common, co-invention expenses are lower and are spread out among a larger set of users.¹³

Adaptation costs are relevant to the adoption decision for enhancement and largely negligible for participation. If the effect of increases in density on adoption benefits (dB/dx_i) is similar for both participation and enhancement, the rate of improvement in net benefits, with respect to density, rises faster for enhancement than for participation. This hypothesis predicts that enhancement will be more sensitive to increases in density than will participation (i.e., $dNB/dx_i \text{ for enhancement} > dNB/dx_i \text{ for}$

¹⁰ In this choice, we follow Forman (2002), who found that investment clustered around a few key margins of behavior. For further motivation see the discussion in our companion study, Forman, Goldfarb and Greenstein (2002).

¹¹ See, e.g., Cherry, Hammond and Wildman (1999), Compaine (2001), Noll et al. (2001).

¹² Such applications often involve complementary organizational change to be used successfully. See, for example, Hitt and Brynjolfsson (1997), Hubbard (2000) and Bresnahan, Brynjolfsson and Hitt (2002).

¹³ This is consistent with other research on IT, complexity and location. For example, in research that examines worker communication patterns using French data, Charlot and Duranton (2003) find that use of Internet technology for inter-organization communication is more likely in large cities. However, their discussion emphasizes the costs of where to locate coordination activity, while ours emphasizes the costs of changing an organization’s processes as technology diffuses.

participation). This is a prediction about the comparative difference between the two types of applications, but it has no implication for whether global village or urban density theory holds for none, either, or both participation or enhancement.

2.2 Industry composition

We distinguish between different versions of industry composition theory. These theories provide potential alternative hypotheses for why the adoption of participation and enhancement increases as population density increases. In all cases the theory argues that there is heterogeneous demand for Internet technology and that the demand increases with the information intensity of the industry. Furthermore, establishments from the same industry tend to cluster in similar places to take advantage of thicker industry-specific labor markets and other shared local resources. For the purposes of this discussion, we allow z_i to denote the IT-intensity of an industry.

In one version of industry composition theory, concentration of Internet technology-intensive activity in some locations is due entirely to agglomeration of IT-intensive industries in those locations and has little to do with population density directly. Previous decisions to concentrate activity could have resulted in the clustering of some types of firms in urban areas. Concentration of adoption of new Internet technology could be an unintended by-product of these decisions. This implies two testable hypotheses: (1) the concentration of IT-intensive industries will explain geographic variation in use for both participation and enhancement and (2) location will have no marginal impact on adoption behavior beyond that explained by variation in industry composition. In other words, this version of the theory hypothesizes that $dNB/dz_i > 0$, $\text{corr}(x_i, z_i) > 0$ and $dNB/dx_i = 0$.

A second version of industry composition theory asserts that increasing population density and IT intensity may increase the likelihood of Internet adoption. Here, industry composition plays a role, but so does location. This version hypothesizes that $dNB/dz_i > 0$ and $dNB/dx_i > 0$. We explore two variants of this alternative theory. Both variants depend on increased demand for complementary resources from IT-intensive firms and supply responses from the providers of these services.

In one variant, which we name the *industry-agglomeration complements theory*, location and industry effects may be complementary for some applications. Concentration of advanced industries attracts the entry of complementary third-party services. These complementary services reduce adaptation costs, particularly for enhancement. Under this theory, new entry increases supply and lowers prices, despite increased demand from IT-intensive firms. Moreover, firms that are IT-intensive may be better able to utilize the pooled resources available in large urban areas. To summarize, $dNB/dz_i > 0$, $dNB/dx_i > 0$, and $d^2NB/dx_i dz_i > 0$. This is the theory receiving the most exposure in the literature on the digital divide.¹⁴

¹⁴ The discussions about the concentration of industrial demand tend to focus on a narrow array of industries, such as new media, dot-coms or electronic retailing. This also leads to a focus on Silicon Valley, Silicon Alley and the greater Boston area. In contrast, we focus on all industries and locations.

In our second variant, which we title *industry-agglomeration substitutes theory*, it is possible for industry composition and urban agglomeration to be substitutes. Concentration of advanced industries leads to increased demand for limited resources. If new entry is insufficient to compensate for this resource competition, prices for local services will increase. In other words, $dNB/dz_i > 0$, $dNB/dx_i > 0$, and $d^2NB/dx_i dz_i < 0$. Thus, while cities will have more IT-intensive industries, the marginal benefit of being in a city decreases as the IT intensity of the industry rises.

Prior research has indicated that entry of third party Internet services in smaller markets was sometimes low (Greenstein 2000). Accordingly, areas with low population should show the effects of industry-agglomeration substitutes theory more than areas with high population. This theory has been discussed as a theoretical possibility (e.g., Gasper and Glaeser 1998) but has seen little empirical verification to our knowledge.¹⁵

3. Data and Method

The data we use for this study come from the Harte Hanks Market Intelligence CI Technology database (hereafter CI database).¹⁶ The CI database contains establishment-level data on (1) 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 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 United States. Since we focus on commercial Internet use, we exclude government, military, and nonprofit establishments (mostly in higher education). We focus on this unit of analysis for three reasons. First, the actions of establishments will reflect local factors better than individual workers (who are mobile) or organizations (who are in multiple locations). Second, previous studies of organizational use of IT demonstrate that most co-invention expenses are incurred at a level wider than an individual. Third, and related, productivity advances take place across a wide array of interdependent processes at an establishment, even at those where the Internet is not used widely.¹⁷

¹⁵ One exception is Kolko (2002), who examines agglomeration in the location decisions of IT-intensive firms.

¹⁶ This section provides an overview of our methodology. For a more detailed discussion, see Forman, Goldfarb and Greenstein (2002).

¹⁷ See, e.g., Hitt and Brynjolfsson (1997), Bresnahan, Brynjolfsson and Hitt (2002).

Our sample contains all commercial establishments from the CI database that contain over 100 employees, 115,671 establishments in all;^{18 19} and HH provides one observation per establishment. We will use the 86,879 clean 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.²⁰ This necessitates routine adjustments of the data for the timing and type of the survey given by HH. Table A1 in the Appendix compares the HH data with the Census data. In general, the samples are close, so most adjustments are small.

3.1. Sample construction and statistical method

Our endogenous variable will be y_j , the value to establishment j of adoption. The variable y_j is latent. We observe only discrete choices: whether or not the establishment chooses participation and whether or not it chooses 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.

In our base specification we assume that the value to establishment j of adopting the Internet is

$$(1) \quad y_j = \sum_i \alpha_i d_{ij} + \sum_l \beta_l d_{lj} + \sum_t \gamma_t d_{tj} + \sum_{t>1995} \delta_t d_{tj} d_{pj} + \sum_m \phi x_{mj} + \sum_q \lambda w_{qj} + \varepsilon_j,$$

where d_{ij} and d_{lj} are dummy variables indicating the industry and location of the establishment, respectively, d_{tj} indicates the month in which the establishment was surveyed, and d_{pj} indicates whether the establishment responded to the long survey.²¹ The variables x_{mj} and w_{qj} denote other location-specific (e.g., population size and density) and establishment-specific variables (e.g., establishment size and dummies indicating single- or multi-establishment firm), respectively. In variations of the model, we may allow for interactions among these variables. If we assume the error term ε_j is i.i.d. normal, then the probability that establishment j participates can be estimated with a probit regression.

We use this model for two research purposes. Our first purpose is descriptive. We illustrate average tendencies for particular establishments in particular locations at a particular point in time. We then weight observations using Census County Business Patterns data to obtain a representative sample.

¹⁸ Because average establishment size differs across urban and rural areas and we only observe establishments greater than 100 employees, our sampling methodology may create some selection bias could affect our interpretation of the marginal impact of location. However, we believe this bias will be minor. We have run regressions with and without size controls and the results remain qualitatively similar.

¹⁹ Previous studies (Charles, Ives and Leduc 2002; Census 2002) have shown that Internet participation varies with business size and that very small establishments rarely make Internet investments for enhancement. Thus, our sampling methodology enables us to track the relevant margin in investments for enhancement, while our participation estimates may overstate participation relative to the population of all business establishments.

²⁰ 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 sample. However, using only this data would result in a very small number of observations for most regions with a population of under one million.

We do this to establish and illustrate the extent of overall variation in adoption propensity. For the average estimates in Tables 2, 3, and 4, we calculate predicted probabilities of adoption for each establishment as if it were surveyed in the second half of 2000 and were given the long survey. The location-specific and establishment-specific variables, x_{mj} and w_{qj} , are not included in this specification.

Our second (and core) purpose is to test competing hypotheses. We analyze the marginal contribution of different factors that shape adoption decisions at the establishment. We report marginal effects from a variety of different specifications, where the model listed above is our base case. The coefficients on α , β , and ϕ are weighted to give a representative sample. (We subsequently display these results in Tables 5 through 9 and Figures 1 through 4).

Two econometric assumptions support the estimates of marginal effects:

- **Exogenous location:** We examine short-run marginal effects of industry and location variables on the decision to invest in Internet technology. To identify these effects, we assume that the location of an establishment is exogenous. We argue that this assumption is supported by the (ex-ante) unexpected rapid diffusion of the Internet, as well as by a regression on a subset of the sample.

For one, this assumption is plausible. As noted in many contemporary accounts, the widespread diffusion of the Internet took most commercial establishments by surprise. Thus, firms did not make establishment-location decisions in anticipation of the Internet. In this study we observe short-run adoption decisions five years into the diffusion of the Internet, before medium and large establishments had time to relocate. That is to say, unlike prior research that has examined the relationship between information technology and economic geography, we have a “natural experiment.” The technology was new and unexpected, so establishment location is exogenous with respect to technology use.

We can test this assumption directly by comparing results between our entire sample of establishments and a special subsample of establishments that (we are certain) fixed their locations prior to 1995 when the commercial Internet became available to most businesses. Since we find that the key estimates do not differ between these two samples, we infer that the potential endogeneity of establishment locations does not alter our inferences about the influence of location on adoption of Internet technology.

- **Simultaneity bias:** Our base econometric specification assumes that the adoption decision of one establishment is independent of every other. This assumption is questionable for multi-establishment firms in which a central executive decision maker (e.g., a CIO) possibly coordinates the choice to adopt or not adopt for each establishment under his domain. Depending on a wide variety of factors, adoption decisions at establishments from the same organization could be either substitutes or complements for one another in use. While understanding that this relationship is of independent interest, it also lies

²¹ Harte Hanks used two surveys. One asked for more details on IT use than the other. We interact the long survey dummy variable with time. See Forman, Goldfarb and Greenstein (2002) for detail.

outside the scope of this study. Here, we are concerned that simultaneity influences the coefficient of interest, the estimate of location on adoption at each establishment.

We address these concerns directly by characterizing the decisions of related establishments at other locations in a reduced form, then measuring whether this alters the estimate of the coefficient on location, while instrumenting for decisions elsewhere. Our focus will be on whether our inferences about the influence of location on adoption of Internet technology are robust to introducing simultaneity into the estimation. We find that, no matter how we measure it, our key results do not change.

3.2. Identifying participation and enhancement

As a GPT, Internet technology is employed in many different uses and applications. Our sample includes at least twenty different types of Internet technology, from basic access to software for TCP/IP-based Enterprise Resource Planning. Moreover, there are considerable differences in the applications used across establishments.

Identifying participation was more straightforward than identifying enhancement. We define participation by an establishment that has basic Internet access or has made any type of frontier investment.²² The establishment survey gives plenty of information about these activities, so we identify participation with confidence.

In contrast, enhancement activity is less transparent in the survey. We look for indications that an establishment must have made investments that involved frontier technologies or substantial co-invention. Most often, these technologies involved inter-organization communication and/or substantial changes to business processes. We identify enhancement from the presence of substantial investments in e-commerce or e-business applications. The threshold for “substantial” is necessarily arbitrary within a range.²³ To be clear, the investments we consider go beyond the downstream interactions with consumers that are traditionally thought of as retail e-commerce. They often involve upstream communication with suppliers, and/or new methods for organizing production, procurement, and sales practices. 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.²⁴

²² 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.

²³ We tested slight variations on this threshold and did not find qualitatively different results.

²⁴ 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

3.3. Descriptive statistics

To obtain a representative sample, we compared the number of establishments in our database to the number of establishments in the Census. We calculated the total number of establishments with more than 50 employees in the Census Bureau's 1999 County Business Patterns data and the number of establishments in our database for each two-digit NAICS code in each location.²⁵ We then calculated the total number in each location. This provides the basis for our weighting. The weight for a given NAICS in a given location is

$$(2) \quad \frac{\text{Total \# of census establishments in location} - \text{NAICS}}{\text{Total \# of census establishments in location}} \cdot \frac{\text{Total \# of establishments in our data in location}}{\text{Total \# of establishments in our data in location} - \text{NAICS}}$$

Each location-NAICS is given its weighting from its actual frequency in the Census. In other words, if our data under samples a given two-digit NAICS at a location relative to the Census then each observation in that NAICS-location is given more importance. In Appendix Table A.1, we compare our sample to the Census data.

In Table 2, we present average rates for participation and enhancement for the United States. Participation by establishments within the sample is at 80.7% (see Unweighted Average in Table 2). The sample underrepresents adopters. Our estimate of the economy-wide distribution, using the true distribution of establishments from the Census, is 88.6% (see Weighted Average in Table 2). Enhancement has been undertaken by 11.2% of our sample and 12.6% of the true distribution.

4. The dispersion of participation and enhancement

In this section, we argue that there is considerable variation across locations in the propensity to adopt Internet technologies. Our results in Table 3 show participation and enhancement rates across Metropolitan Statistical Areas (MSAs)²⁶ of varying size in the United States. Because there has been little prior work on variation in use of the Internet by business, these descriptive findings are not widely

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 twenty 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 Forman, Goldfarb and Greenstein (2002).

²⁵ We use 50 employees because potential differences between different times for taking the survey mean that firms could grow after the Census and therefore be in the CI database. It was necessary to be inclusive for the weighting because some small rural areas had less than three firms in both the Census and the CI database; and therefore if one firm grew from the time of the Census to the time of the CI survey, the weightings would be difficult to interpret. The results are robust to weighting by firms with more than 100 employees in the Census and those with more than 25 employees. This is not surprising given the high correlation between these values.

appreciated. In a previous work (Forman, Goldfarb and Greenstein 2002), we contrasted them with other studies of the geographic variation in household use of the Internet and infrastructure deployment. We reproduce it here because, on a broad level, this table motivates the present study. Table 3 presents sizable differences in participation and enhancement adoption between large MSAs, medium MSAs, small MSAs and rural areas. On the surface, this evidence supports either urban density theory or industry composition theory. We see that large MSAs have very high participation rates, averaging 90.4%. Participation rates in medium-sized MSAs and rural areas are lower at 84.9% and 85.1%, respectively. In small MSAs the participation rates are even lower, 75.5% on average.

The disparities in enhancement adoption rates are even greater (again, see Table 3). Large MSAs have relatively high adoption rates, with an average of 14.7%. In medium MSAs, adoption 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. Once again, these averages suggest that urban density theory or industry composition theory may hold. Clearly there is considerable variation in adoption propensity by city size.

There is also variance in adoption propensity within the subset of large MSAs. In Tables 4a and 4b, we list the participation and enhancement estimates for MSAs with over one million people, in order of highest to lowest adoption rates.²⁷ As we do in all of our tables, we list the standard errors and number of observations to identify the degree of statistical confidence in the estimates.²⁸ (For comparison, Tables 4a and 4b also list the marginal effect of location on adoption, which we will discuss later.)

In Table 4a, we show that participation is high in major urban locations. Virtually all establishments in the major urban areas are participating. Of the forty-nine MSAs, thirty-five are above 90%. All but five are within a 95% confidence interval of 90%. Nevertheless there are large differences between metropolitan areas at the extremes.

In Table 4b, the top ten MSAs that adopted enhancement include a set of areas that partially overlaps with the list in Table 4a. (Five of the top ten are also in the top ten for participation.) Again, the differences between the lowest adopting areas and the highest adopting areas are substantial. The considerable variation in adoption propensity by city size and among cities of the same size illustrated in Table 3 and Tables 4a and 4b, respectively, motivated this study. We next explore some potential causes for this variation.

²⁶ 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*.

²⁷ When two or more MSAs are part of the same urban environment, the census combines them into CMSAs, or Consolidated Metropolitan Areas. For example the Dallas-Fort Worth CMSA contains both Dallas and Fort Worth. In Table 4, we present the CMSA results rather than the individual MSA results when an MSA is part of a CMSA.

²⁸ These are computed using the delta method.

5. The Marginal Impact of Location on Internet Adoption

In this section, we estimate equation (1), while focusing on testing between global village and urban density theories. We weight observations by the inverse probability that an establishment will appear in our sample. To be precise, the weight for each observation is the total number of establishments in a state/NAICS in Census County Business Patterns data divided by the number of establishments in the state/NAICS in our sample multiplied by controls for sampling the same establishment twice.

In Table 5, we show the roles of population and density in the adoption decision. Part A presents the coefficients of the probit regressions. Part B presents the marginal effects. All probit regressions include dummy variables for three-digit NAICS, the month the data were collected, survey type, survey type interacted with month, and whether or not 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. For columns 1 and 5, we use rural state areas for the base. For columns 2, 3, 6, and 7, we include a “rural area” dummy for rural areas, since no meaningful population figures exist for these areas. In columns 4 and 8 we include population density for all urban and rural areas using low-density areas as the base.

5.1 The marginal impact of location

From Table 5, it is clear there is no support for the urban density theory in participation. Controlling for industry and firm characteristics, location size and density have little impact on the decision to adopt at the participation level. If anything, the effects of location size and density support global village theory, but the impact of geography is of limited economic and statistical significance. In column 1, we show that medium and large MSAs are 0.5% to 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 do not increase the probability of participation. While not statistically significant, the coefficient suggests a possible decrease in the probability of participation. In column 3 we include a squared population term. In this formulation, the linear term remains statistically insignificant, while the squared term is significant, albeit very small. This model implies that the effects of population will turn negative once urban areas exceed 7.039 million, a threshold that is larger than all but the five largest urban areas. In column 4 we include dummies for population density. This alternative specification gives very similar results. Variation in

population density does not affect participation by more than 1%, and it is always statistically insignificant.

In contrast to participation, the effects of population size and density on enhancement support urban density theory. Column 5 in part B of Table 5 shows that establishments in medium and large MSAs adopt enhancement at a rate 0.8% to 1.1% higher than rural areas. Column 8 shows that establishments in medium and large MSAs adopt enhancement at a rate 1.0% to 1.5% more. All of these effects are statistically significant. They are economically significant in light of the average enhancement rates of 12.6%. While column 6 suggests that a linear population term has little effect on enhancement, column 7 shows that population will have a statistically and economically significant positive effect for all MSAs below 8.8 million in size (all but New York, Los Angeles, and Chicago).

Together these results support the key prediction of GPT theory about the comparative relationship between complexity of application and geographic variance in use. The probability that an establishment adopts the Internet for enhancement is more sensitive to geographic variation in density than the similar probability for participation. We interpret this as evidence that the applications more dependent on third-party support and complementary services are most costly to deploy in less dense locations.

Variance in the role of location varies by city size. In Figures 1 and 2 we graph the marginal effect of location in the baseline probit in model (1) to reinforce the results of Table 5.²⁹ We divide locations into four types: large MSAs, medium MSAs, small MSAs, and statewide rural (non-MSA) regions. We plot the kernel density estimates of the effects of location on participation and enhancement, respectively.³⁰ We use Epanachnikov kernels with “optimal” bandwidths.

In Figure 1 small MSAs and rural areas have a fatter right tail, while the density for large MSAs reaches its peak slightly below any of the three other classes of geographic areas. Although the distributions of each MSA size are roughly centered in the same place, the plot shows that, comparatively, the large MSAs have less variance in adoption of participation than the other MSAs, especially the small MSAs. In all, this figure provides further support for global village theory: Increases in local population size and density do not increase the likelihood of participation adoption. If anything, they lower it.

In Figure 2, the density estimate for large MSAs stochastically dominates those for small MSAs, medium MSAs, and rural areas. The center peak of the large MSAs' distribution is obviously at a higher value than the others. However, it is also apparent that the variance of enhancement adoption within large MSAs and rural areas is less than that within small and medium MSAs. This figure

²⁹ This probit is not depicted in any table. We identify the effects of population size and density directly through the location-specific dummy variables.

³⁰ The omitted MSA is San Jose, the top MSA in adoption of enhancement.

provides a visual depiction of the results in Table 5 that urban density theory better describes the geographic diffusion of enhancement than global village theory.

Despite these results it is important to keep in mind our previous discussion of adoption and participation diffusion patterns in large MSAs (see Table 4), which shows that the variance of adoption rates within large MSAs is great. This is not immediately observable from Figures 1 and 2 because there is even more variance in the small MSAs, medium MSAs and rural areas.

Table 6 provides summary statistics on the marginal effects of the same regressions used for Figures 1 and 2. Again, the results show that establishments in larger MSAs are less likely to adopt participation and more likely to adopt enhancement.

We conducted a number of robustness checks on our results. As noted, we were concerned that establishment location decisions might be correlated with improvements in communications technology. To control for this potential source of endogeneity, we re-estimated the model using only establishments that had been added to the HH database prior to 1995, the year in which Internet technology began to diffuse widely to businesses. Although this restricted the size of our sample substantially (to 23,436 observations), the basic results remain the same. The correlation coefficients between our baseline marginal effects and those using pre-1995 data are 0.829 for participation and 0.997 for enhancement. Qualitative results did not change.

We tried a number of other robustness checks. In order to ensure our results were not driven by omitted variables, we experimented with a variety of different specifications, using different location variables (e.g., CMSA dummies), different firm controls (e.g., revenue, private/public), and alternative measures of population size and density. To ensure our results were not a function of the particular weighting scheme used, we tried weighting the probit regressions by three-digit NAICS/states and two-digit NAICS/MSAs, as well as not weighting at all. In all cases the results remained qualitatively the same; the correlation coefficients between our baseline coefficient estimates and the alternative specifications were between 0.88 and 0.95 for participation and 0.78 and 0.90 for enhancement.³¹

5.2 Location and multi-establishment firms

Multi-establishment firms often adopt new communication technologies at some, but not all, of their establishments. Multi-establishment firms choose to locate Internet technology in the locations that are the most cost-effective; this implies that the effects of greater population size and density (urban density) will be more important to adoption decisions for establishments that are part of multi-

³¹ We also explored whether establishment differences across geographic locations drive our results. Though unobservable establishment differences could play a role, we were unable to uncover any pronounced observable establishment differences. Using weighted data, establishments in large MSAs are larger (12.8% of establishments have >500 employees versus 9.0% for small MSAs and 9.9% for rural areas) and more likely to be multi-establishment (48.7% multi-establishment versus 43.9% for small MSAs and 33.4% for non-MSAs). When using unweighted data much of the difference disappears (12.7% of establishments in large MSAs >500 employees versus 12.1% in small MSAs 13.4% in rural areas; 46.3% of establishments in large MSAs multi-establishment versus 46.8% in small MSAs and 40.8% in rural areas).

establishments firms. This will be particularly true for complex enhancement technologies, if co-invention costs are lower for firms in locations with better local support. In this section, we systematically examine how multi-establishment status changes the marginal returns to location. We find the results to be in the expected direction, but not sufficiently large to alter the evidence for or against urban density or global village theory.³²

First, we show that multi-establishment status does predict adoption. The results in Tables 7a and 7b support the hypothesis that the effects of greater population size and density are larger (more positive) if the establishment is part of a multi-establishment firm. Equivalently, the effects of smaller population size and density are greater for single-establishment firms. This is particularly true for enhancement technologies.

Column 1 of Table 7b shows that when a multi-establishment dummy is interacted with MSA dummies the non-interacted medium and large MSA dummies now both have a statistically significant marginal effect of -1.3% and -1.0% respectively. There is no statistically or economically significant effect in the interaction terms themselves. When multi-establishment is interacted with population, there is a statistically significant but small marginal impact of 0.042% per 100,000-person increase in population.³³ The effect of population itself has a more negative impact than in Table 5.³⁴ Similar results hold when using population density rather than size.

The economic significance of multi-establishment is larger for enhancement. Interactions of multi-establishment dummies with population size and density weaken the positive effects of non-interacted population variables. In many cases marginal effects become smaller or less statistically significant. In all the columns, the interaction terms suggest that the effects of population size and density are greater for multi-establishment firms.

The economic effects can be substantial. Columns 5 and 8 suggest that establishments that are part of multi-establishment firms and are located in large MSAs or densely populated counties are 2.2% to 2.3% more likely to adopt enhancement than stand-alone establishments in the same locations. These marginal effects are sizable compared to typical enhancement rates of 13% . The marginal effects of interacting multi-establishment with MSA population (in column 6) are statistically significant, but only 0.025% per 100,000 person increase in population.

The previous discussion ignores potential simultaneity bias arising from establishment-level analysis of adoption decisions in multi-establishment firms. We extended our statistical model to

³² To our knowledge, this study is the first to address how multi-establishment status affects technology adoption. Note, however, that our ability to do this comprehensively is restricted by our data. We do not observe all of an organization's establishments. We observe only those with more than 100 employees.

³³ This calculation is not shown in the table.

³⁴ These results do not reflect any collinearity between multi-establishment and urban areas. Multi-establishment dummies had a statistically and economically significant impact in the baseline regressions in Table 5; the marginal effect was between -2.6% and -2.8% . Moreover, the correlation between multi-establishment status and location in an MSA is positive but small (0.0427).

include variables capturing the behavior of other establishments within the same firm. In particular, we added variables measuring the percentage and total of other establishments within the same firm adopting the dependent variable (i.e., participation or enhancement). Because these variables are likely to be correlated with unobserved factors affecting the decision to adopt participation and enhancement, we also used nonlinear instrumental variable techniques. For instruments, we used average population and density of other establishments in the same firm. These should be correlated with adoption decisions at the firm's other establishments, but not at the establishment of interest.

We first show whether our baseline estimates (without multi-establishment interactions) are robust to the inclusion of other establishment decisions. The results in Appendix Table A.2 show the marginal effects of probit regressions that add other establishment adoption decisions to the models in Table 5. These regressions are performed with and without IV. These robustness checks make little difference to the estimated relationship between population density and Internet adoption; the results of our tests contrasting global village and urban density theory remain unchanged. Part A of Table A.2 shows that variables capturing the percentage of establishments with participation and enhancement are positive and significant in weighted probit regressions, however their significance disappears once we instrument in Part B of Table A.2. We interpret the probit regressions without IV as picking up unobserved heterogeneity, which is ultimately eliminated in the IV probit regressions. The new variables have little effect on the population marginal effects. Moreover, statistical significance of our location variables is retained in the IV version of this model.

Similarly, Appendix Table A.3 shows that the inclusion of the decisions of other establishments has almost no impact on the estimates of our multi-establishment interaction models in Table 7. The new variables have the same pattern as in Table A.2: positive and significant in probit regressions without IV, insignificant in probit regressions with IV. We conclude that simultaneity of decisions does not significantly bias our results. In other words, multi-establishment firms chose to disproportionately place their Internet investments in urban locations. Though these results have statistical significance, for purposes of understanding the relevance of global village or urban density theory the difference in behavior is negligible.

6. Industry Composition

The differences between the average adoption rates in Table 3 and the marginal effects in Table 5 show that the effects of location on participation and enhancement fall if we include controls for establishment size, industry, and firm status. In Table 3, large MSAs have almost a 15% higher participation rate and 5% higher enhancement rate than small MSAs. In Table 5, locating in a large MSA rather than a rural area reduces the probability of participation by 0.6% and increases the probability of enhancement by 1.1%.

The large differences in adoption rates between large and small urban areas in Table 3 reflect differences in industry composition across locations. Industry composition explains much more of the variation in participation and enhancement rates than location does. Once industry is controlled for, the incremental contribution of location in the probit regressions is small. This is shown in Table 8. The pseudo- R^2 of a probit for participation including only location dummies is 0.1526, whereas the pseudo- R^2 of a probit with only industry dummies is 0.2251. Adding location dummies to a probit that includes industry dummies barely improves fit, from 0.2251 to 0.2339.

Enhancement displays a similar pattern. Location dummies explain only 0.0347 of the variation in enhancement, industry dummies explain 0.0591, and the combination of industry and location dummies explains 0.0672. While there remains a great deal of unexplained variation in our results, we conclude that an establishment's industry explains more of the variation in Internet use than does geographic location.

6.1. What does industry composition explain?

All three theories that emphasize the importance of industry composition assert that leading industries concentrate in large urban areas. To test this hypothesis, we separated establishments by geographic location type (i.e., rural and small, medium, and large MSAs) and calculated the kernel density of industry marginal effects for each type of location. The underlying marginal effects are the same across all four types of locations. However, the densities of each marginal effect differ because of differences in industry composition across locations. We did this for both participation and enhancement.

In Figure 3, we show the kernel density estimates of the marginal effects of industry by geographic area for participation.³⁵ Lead-user industries tend to be concentrated in large geographic areas. The average of the marginal effects of industry in rural and small MSAs is -18.7% and -20.2% , while the average of the marginal effects for medium and large MSAs are -18.8% and -16.9% . Except in comparing rural and medium MSAs, these averages are significantly different from one another at the 99% level. Large MSAs tend to have more lead-user industries, even for participation.

Figure 4 shows that lead users of enhancement are even more skewed toward large MSAs. Rural areas and small MSAs have the highest densities along the left tail of the distribution, whereas large and medium MSAs have higher densities along the right tail. The average marginal effect of industry on enhancement adoption increases as location size increases: -8.0% in rural, -7.8% in small MSAs, -7.7% in medium MSAs, and -7.4% in large MSAs. Again, these averages are all significantly different from one another at the 1% level. The bulk of the variation in Table 3 reflects differences in

³⁵ All industry results are unweighted. The omitted industry is information and data processing (NAICS 514). We use Epanachnikov kernels with bandwidth of 0.05 for participation and 0.005 for enhancement. These are wider than "optimal" bandwidths. Optimal bandwidths fail in this case because there are thousands of observations but

industry composition between small and large MSAs, rather than other location-specific benefits of locating in large urban areas.

We conclude that large 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 MSAs. The industries in the upper quartile are traditionally information intensive, such as utilities, finance and insurance, company headquarters, professional and scientific services, electronics manufacturing, and wholesale trade.³⁶ 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.

Two conclusions emerge. First, this supports the importance of controlling for industry composition when testing between global village and urban density theory. Second, inferences about the relationship between IT and economic activity are fraught with omitted variable biases in the absence of such controls. We note that this factor is missing from all existing analysis about the geography of Internet use.

6.2. Are industry and location complements or substitutes?

The results in Table 8 and Figures 3 and 4 suggest that industry composition theory explains a major part of establishment decisions to adopt Internet technology. However, more analysis is needed to determine whether industry and location effects are complements or substitutes. To examine the industry-agglomeration substitutes and complements hypotheses, we rerun the probit regressions in Table 5 with additional variables controlling for (1) whether the establishment is in a lead-user industry and (2) interactions of this lead user dummy with MSA-size dummies. We define lead-user industries in one of two ways: (1) the top quartile of participation or enhancement adopters among three-digit NAICS industries in our study or (2) the United States Department of Commerce's (2002) top fifteen IT-using industries as reported by Daveri and Mascotto (2002). Both measures of IT intensity have strengths and weaknesses. The measure based on the top quartile selects on the basis of the dependent variable; the measure from Daveri and Mascotto's study is based on a more general measure of IT intensity than the Internet.

Consequently, these are not final tests. We present these results as descriptive evidence that may support either a complement or substitute relationship between industry and location effects. To further supplement our analysis, we later examine whether establishments in lead-user industries (defined by high marginal effects) also tend to be located in favorable locations (locations with high marginal effects). In other words, we examine whether "good" industries are located in "good" locations.

only eighty-one possible values as there are eighty-one relevant three-digit NAICS levels. Therefore the optimal bandwidth does almost no smoothing.

³⁶ For more detail, see Forman, Goldfarb and Greenstein (2002).

Using the data on lead Internet adopters, Part A of Table 9 shows that there is little evidence of a complementary relationship between industry and location in participation; if anything, they are substitutes. An establishment in a top quartile NAICS and a large MSA is 2.6% more likely to adopt participation than an otherwise equivalent top quartile establishment in a rural area. However, large MSAs are 2.8% less likely to adopt than small MSAs, though the difference is not statistically significant. The NAICS-level controls likely explain the lack of significance of the IT-intensive industry dummy (under both definitions). Perhaps because they are based on the Department of Commerce's more general measure of IT use, the industry-location interactions in Part B of Table 9 are less significant than in Part A. However, they tell exactly the same story. An establishment in a lead-user SIC (Standard Industrial Classification) and large MSA is no more likely to adopt participation than an otherwise equivalent establishment in a small MSA.

In contrast to participation, the results for enhancement in Table 9 show a strong complementary relationship between industry and location. Part A shows that an establishment in a top quartile NAICS and a large MSA is 20.0% more likely to adopt enhancement than an otherwise equivalent top quartile establishment in a rural area. The difference between small MSAs and large MSAs is even larger. Large and medium MSA establishments are equally likely to adopt enhancement. The results based on the Department of Commerce's more general measure of IT-intensity again are weaker, but again tell the same general story: An establishment in a lead-user industry and a large MSA is more likely to adopt enhancement than an otherwise equivalent establishment in a small MSA. Moreover, the positive coefficient on the interaction of lead-user industry and large MSA is the only statistically significant interaction in the regression. In summary, as was suggested by the industry-agglomeration complements theory, industry and location are complements for complex applications.

As an alternative way of examining whether industry and location effects tend to complement or substitute one another, we compare median industries by IT use across cities. This is not an explicit test for complementarities, but rather another way of showing whether establishments in IT-intensive industries tend also to be located in favorable locations.

We calculate correlations between the marginal effects of the median industry within each urban area with those of the urban areas themselves. The results are consistent with our findings in Table 9. For participation, the marginal contribution of the median industry in a location is uncorrelated with the marginal contribution of the location itself ($\rho = -0.0214$). However, this result disguises a large difference between large and medium MSAs on the one hand and small MSAs and rural areas on the other. The correlation between the marginal contribution of median industry and location is significantly positive ($\rho = 0.307$) for large and medium MSAs but significantly negative ($\rho = -0.211$) for small MSAs and rural areas.

For larger population locations, good cities do have good industries for participation; yet for smaller areas the opposite is true. For enhancement, there do seem to be complementarities between cities and locations ($\rho = 0.161$). These complementarities do not vary much by city size. Regardless of size, good industries are in good cities and good cities are dominated by good industries. Cities and IT-intensity appear to be complements for enhancement, but not for participation, especially in low population areas.

This is consistent with the industry-agglomeration complements theory, which emphasizes complementarity for complex applications. We interpret this complementarity as a likely result of spillover effects in using frontier technologies. The exception is as interesting as the more general finding. In low population areas, favorable locations likely only help adoption in firms without in-house expertise. In that case, location and industry become substitutes. This result is potentially consistent with industry-agglomeration substitutes theory, which is most relevant for small urban areas.

7. Conclusions

Has the Internet realized its promise of reducing the importance of location to economic activity? In this paper, we tested competing views by examining hard data about the short-run decisions of firms to invest in the Internet.

By 2000, participation activities such as email and web browsing had diffused almost everywhere. For these simple technologies, there is no evidence that urban density theory held: Industry composition explains the higher levels of participation adoption in urban areas. Once industry composition was controlled for, we found that the variation across locations is best explained by the global village theory. Moreover, we found some evidence of substitution between IT-intensity and location density: IT-intensive industries in rural locations were the most likely to adopt participation technologies. For the technology with low adaptation costs, we find little evidence of complementarities between industry and location.

For complex enhancement technologies, adoption behavior is best explained by urban density theory. We find evidence of complementarities between industry and location effects in the adoption of enhancement: IT-intensive firms found greater benefits than other firms from pooled resources in large cities.

The geographic variation in use was consistent with GPT theory: enhancement costs are more sensitive to variation in density than are participation costs. Adopters of participation faced low technical hurdles to implementation, while adopters of enhancement faced high ones. Establishments overcame these costs because they had experience with overcoming technical and co-invention costs, had access to rich complementary resources through local markets, or both.

Nevertheless, our research shows that the Internet did not exacerbate geographic inequalities by diffusing primarily to urban areas with complementary technical and knowledge resources. This is true even for complex applications that enhance business computing processes. More precisely, there were forces leading to concentration of complex applications in urban areas, consistent with industry composition theory with complementarities, but these forces were not concentrated in a few cities. Users of complex applications are widely dispersed to many industries and locations.

Of more relevance to economic growth, the geographic pattern of adoption for enhancement is quite understandable as an economic matter if these applications have high value and potentially high co-invention costs, which differ depending on local conditions. In smaller MSAs and rural areas, thin technical labor markets alone could drive up costs of operating facilities employing frontier Internet technology. These effects are particularly pronounced for multi-establishment firms. Because the investment is linked to competitive settings, multi-establishment organizations, if they had a choice, would implement new business processes in the more hospitable settings of major urban areas. In addition, multi-establishment organizations would hesitate to open their own complex Internet facilities in rural areas until the costs are lowered. In any case, variation in the availability of complementary resources would lead to more use of enhancement in major urban areas. Particularly for multi-establishment organizations, the outcome is not alarming at all.

Our findings provide several avenues for future research. First, our conclusions apply only to medium to large establishments. We defer to future research to examine small establishments and newly founded firms, that may face a different array of benefits, and may have diminished access to internal resources for idiosyncratic adaptation costs. It is already apparent from existing papers that participation is lower on average in small firms,³⁷ so we speculate that smaller firms will be more sensitive to geographic variation in local complementary resources than found here.

Second, our findings suggest broadly that variations in co-invention costs across technologies and locations shaped the diffusion of Internet technology. Research on the role of co-invention costs on Internet diffusion has been hampered by the binary nature of the adoption decision considered in many studies, including this one. Future work should analyze variations in firm co-invention costs, emphasizing in particular the impact of variation in labor market conditions, spillovers, and markets for technical support.

Third, our findings have implications for variation in the dollar value of investment across location and industries and the net returns from those investments. The diffusion model of adoption used in this study implies that magnitudes of investment should follow patterns similar to the patterns for the binary adoption decision. Hence, we speculate that the flow of investment dollars will correlate

³⁷ For the beginnings of such research, see Atrostic and Gates (2001) on manufacturing establishments, and Buckley and Montes (2002) and Bitler (2002) for analysis of small business computer use.

positively with the rankings of location and industries uncovered in this study. Furthermore, if this model is correct, the investment dollars affiliated with the commercialization of the Internet were widely dispersed throughout locations and industries in the United States. This speculation awaits confirmation with data about investment behavior beyond adoption.

References

- Atrostic, Barbara, K., and John Gates. 2001. "U.S. Productivity and Electronic Business Processes in Manufacturing." Working Paper #01-11. Center for Economic Studies, US Census Bureau.
- Beardsell, M.; Henderson V. 1999. "Spatial Evolution of the Computer Industry in the USA." *European Economic Review*, 43(2), pp. 431-456.
- Bitler, Marianne, P. 2002. "Computer Adoption by Small Businesses: Who Adopts?" Mimeo, Rand Corporation, Santa Monica, CA.
- Bresnahan, Timothy, and Manuel Trajtenberg. 1995. "General Purpose Technologies: 'Engines of Growth'?" *Journal of Econometrics*. 65 (1): 83-108.
- Bresnahan, Timothy, Erik Brynjolfsson, and Lorin Hitt. 2002. "Information Technology, Work Organization, and the Demand for Skilled Labor: Firm-Level Evidence." *Quarterly Journal of Economics* 117 (February): 339-376.
- Buckley, Patricia, and Sabrina Montes. 2002. *Main Street in the Digital Age: How Small and Medium-sized Business Are Using the Tools of the New Economy*. Department of Commerce. Available at <http://www.esa.doc.gov/reports.cfm>.
- Cairncross, Frances, 1997. *The Death of Distance*. Cambridge, MA: Harvard University Press.
- Castells, Manuel, 2003. *The Internet Galaxy, Reflections on the Internet, Business and Society*. Oxford, Oxford University Press.
- Census. 2002. "Detailing Tabulations of Manufacturing E-Business Process Use in 2000." Survey Release from United States Department of Commerce. Available at www.census.gov/epcd/www/ebusiness.htm.
- Charles, S., M. Ives, and A. Leduc. 2002. "Embracing E-Business: Does Size Matter?" Research Report, Statistics Canada.
- Charlot, Sylvie and Gilles Duranton. 2003. "Cities and Workplace Communication: Some French Evidence." Mimeo, London School of Economics.
- Cherry, Barbara, Alan H. Hammond, and Steve Wildman. 1999. *Making Universal Service Policy: Enhancing the Process Through Multidisciplinary Evaluation*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Chun, Hyunbae. 2003. "Information Technology and The Demand For Educated Workers: Disentangling the Impact of Adoption Versus Use." *The Review of Economics and Statistics* 85(1): 1-8.
- Compaine, Benjamin. 2001. *The Digital Divide: Facing a Crisis or Creating a Myth?* Cambridge, MA: MIT Press.
- Crandall, Robert and Alleman, James H. 2002. *Broadband: Should We Regulate High Speed Internet Access?* AEI-Brookings Joint Center for Regulatory Studies, Brookings Institution Press: Washington DC.
- Daveri, Francesco, and Andrea Mascotto. 2002. "The IT revolution across U.S. States." Mimeo, IGIER.

- David, Paul A. 1969. "A contribution to the Theory of Diffusion" Stanford Center for Research in Economic Growth, Memorandum No. 71, Stanford University.
- Doms, Mark, Timothy Dunne, and Kenneth R. Troske. 1997. "Workers, Wages, and Technology." *The Quarterly Journal of Economics* 112(1): 253-290.
- Downes, T. and Greenstein, S., 2002, "Universal Access and Local Internet Markets in the U.S." *Research Policy* 31: 1035-1052.
- Forman, Chris. 2002. The Corporate Digital Divide: Determinants of Internet Adoption. Mimeo, Carnegie Mellon University.
- Forman, Chris, Avi Goldfarb, and Shane Greenstein. 2002. "Digital Dispersion: An Industrial and Geographic Census of Commercial Internet Use." NBER Working Paper #9287.
- Gaspar, Jess, and Glaeser, Edward. 1998. "Information Technology and the Future of Cities." *Journal of Urban Economics*. 43(1): 136-156.
- Goolsbee, Austan and Peter Klenow. 2002. "Evidence on Learning and Network Externalities in the Diffusion of Home Computers." *Journal of Law and Economics* 45(2), pp. 317-344.
- Gorman, Sean P., 2002, "Where are the Web Factories: The Urban bias of E-Business Location." Mimeo, Neoteric Media Inc., Arlington VA.
- Greenstein, Shane, 2000. "Building and Developing the Virtual World: The Commercial Internet Access Market." *Journal of Industrial Economics* 48(4): 391-411.
- Greenstein, Shane, 2003. "The Economic Geography of Internet Infrastructure in the United States," in (eds.) M. Caves, S. Majumdar, and I. Vogelsang, *The Handbook of Telecommunications Economics. Volume II*. Elsevier.
- Harris, Richard. 1998. "The Internet as a GPT: Factor Market Implications," in E. Helpman (eds.), *General Purpose Technologies and Economic Growth*. Cambridge, MA: MIT Press.
- Henry, David, and Donald Dalton. 2002. "Information Technology Industries in the New Economy." Chapter III in *Digital Economy 2002*, Department of Commerce, at <http://www.esa.doc.gov/reports.cfm>.
- Hitt, Lorin and Erik Brynjolfsson. 1997. "Information Technology and Internal Firm Organization: An Exploratory Analysis." *Journal of Management Information Systems* 14 (Summer): 1-101.
- Hubbard, Tom. 2000. "The Demand for Monitoring Technologies: The Case of Trucking." *Quarterly Journal of Economics* 116 (May): 533-560.
- Ireland, N. and Stoneman, P. 1986. "Technological diffusion, expectations and welfare." *Oxford Economic Papers* 38: 283-304.
- Kolko, Jed. 2000. "The Death of Cities? The Death of Distance? Evidence from the Geography of Commercial Internet Usage," In Ingo Vogelsang and Benjamin M. Compaine (eds.), *The Internet Upheaval*. Cambridge, MA: MIT Press.

- Kolko, Jed. 2002. "Silicon Mountains, Silicon Molehills, Geographic Concentration and Convergence of Internet Industries in the US." *Information Economics and Policy* 14(2): 211-232.
- Krugman, Paul, 1991, *Geography and Trade*. Cambridge: MIT Press.
- Marshall, Alfred. 1920. *Principles of Economics*. 8th edition. New York: Porcupine Press.
- Moss, Mitchell L., and Anthony M. Townsend, 1997. "Tracking the net: using domain names to measure the growth of the Internet in US cities." *Journal of Urban Technology* 4(3): 47-60.
- National Telecommunications and Information Administration. 2000a. Falling Through the net: Toward Digital Inclusion. See <http://www.ntia.doc.gov/reports.html>.
- , 2000b. "Advanced Telecommunications in Rural America, The Challenge of Bringing Broadband Service to All Americans." See <http://www.ntia.doc.gov/reports.html>.
- , 1999. "Falling Through the Net: Defining the Digital Divide." <http://www.ntia.doc.gov/reports.html>.
- , 1998. "Falling Through the Net II: New Data on the Digital Divide." <http://www.ntia.doc.gov/reports.html>.
- , 1995. "Falling Through the Net: A Survey of the 'Have Nots' in Rural and Urban America." <http://www.ntia.doc.gov/reports.html>.
- Noll, Roger, Dina Older-Aguilar, Gregory Rosston, and Richard Ross, 2001. "The Digital Divide: Definitions, Measurement, and Policy Issues." Paper presented at the American Economics Association Meetings, January 2002.
- Price, Lee, and George McKittrick, 2002. "Setting the Stage: The New Economy Endures Despite Reduced IT Investment." in Chapter I, *Digital Economy 2002*, Department of Commerce, at <http://www.esa.doc.gov/reports.cfm>.
- Rogers, Everett, 1995. *The Diffusion of Innovations*. Fourth Edition. New York: Free Press.
- Sinai, Todd and Joel Waldfogel. 2001. "Geography and the Internet: Is the Internet a Substitute or a Complement for Cities?" Mimeo, Wharton School of Business, University of Pennsylvania, Philadelphia.
- U.S. Department of Agriculture. 2001. "Farm Computer Usage and Ownership Report, 07.30.01." Available at <http://usda.mannlib.cornell.edu/reports/nassr/other/computer/>
- United States Department of Commerce. 2002. "A Nation Online: How Americans are expanding their use of the Internet," Joint Report from the National Telecommunications and Information Administration and the Economics and Statistics Administration.
- Zooks, Matthew. 2000a. "Internet Metrics: Using Hosts and Domain Counts to Map the Internet Globally." *Telecommunications Policy*, 24, 6/7.
- Zooks, Matthew. 2000b. "The Web of Production: the Economic Geography of Commercial Internet Content Production in the United States." *Environment and Planning* 32: 411-26.

Table 1
Predictions of Alternative Hypotheses

Hypothesis	Prediction	Testable Implication
Global Village	$dNB/dx_i < 0$	Coefficient on rural > small MSA > medium MSA > large MSA
Urban Density	$dNB/dx_i > 0$	Coefficient on rural < small MSA < medium MSA < large MSA
General Purpose Technology	dNB/dx_i for enhancement > dNB/dx_i for participation	(Coefficient on large MSA – coefficient on rural/small MSA for enhancement) > (Coefficient on large MSA – coefficient on rural/small MSA for participation)
Industry Composition (IC)	$dNB/dx_i = 0$; $dNB/dz_i > 0$	Industrial composition explains adoption and eliminates the measured impact of population (density)
Industry-Agglomeration Complements	$dNB/dz_i > 0$, $dNB/dx_i > 0$, and $d^2NB/dx_i dz_i > 0$	Urban density theory holds and the effect of IT-intensity is increasing in MSA size
Industry-Agglomeration Substitutes	$dNB/dz_i > 0$, $dNB/dx_i > 0$, and $d^2NB/dx_i dz_i < 0$	Urban density theory holds and the effect of IT-intensity is decreasing in MSA size.

Notes :

$NB(x_i, z_i) \equiv B(x_i, z_i) - C(x_i, z_i)$

x_i describes geographic conditions such as population size and density.

z_i describes the IT-intensity of an industry.

Testable implications describe manifestation of each theory in probit models of decision to adopt participation and enhancement.

Table 2
National Internet Adoption Rates (in percentages)

	Weighted Average	Unweighted Average
Participation	88.6%	80.7%
Enhancement	12.6%	11.2%

Notes:

Source: Author's calculations using HH database and census data.

Definitions for Participation and Enhancement given in the text. See also Forman, Goldfarb, and Greenstein (2002) for further documentation.

Unweighted average uses only HH database sample.

Weights are defined by equation (2), as given in the text.

Table 3
Average Adoption by Size of Metropolitan Statistical Area (MSA)

Population	Average Participation	Standard Error	Average Enhancement	Standard Error	Number of Areas
Rural: Non-MSA	85.1%	0.1%	10.6%	0.2%	49
Small MSA: <250,000	75.5%	0.2%	9.9%	0.3%	143
Medium MSA: 250,000-1 million	84.9%	0.2%	11.2%	0.3%	116
Large MSA > 1 million	90.4%	0.1%	14.7%	0.2%	57

Notes:

Source: Author's calculations using HH database and census data.

Definitions for Participation and Enhancement given in the text. See also Forman, Goldfarb, and Greenstein (2002) for further documentation.

All calculations use weighted averages, where weights are defined by equation (2), as given in the text. Standard errors are computed using the delta method.

Table 4a: Participation Among Metropolitan Statistical Areas with Over One Million People

Avg. Rank	City	Avg. Rate	Std. Err. Rate	Marg Rank	Marg. Coef.	Std. Err. Coef.	Obs.	Population
1	San Francisco--Oakland--San Jose, CA	96.4%	0.4%	7	0 (base)	N/A	2135	7,039,362
2	Denver--Boulder--Greeley, CO	95.9%	0.7%	43	-0.067	0.027	940	2,581,506
3	Cleveland--Akron, OH	94.8%	0.6%	23	-0.038	0.021	1099	2,945,831
4	Seattle--Tacoma--Bremerton, WA	93.9%	0.5%	3	0.025	0.015	1012	3,554,760
5	Salt Lake City--Ogden, UT	93.5%	0.8%	6	0.007	0.019	535	1,333,914
6	San Antonio, TX	93.3%	0.8%	1	0.035	0.021	395	1,592,383
7	Providence--Fall River--Warwick, RI--MA	93.0%	1.2%	24	-0.038	0.032	290	1,188,613
8	Grand Rapids--Muskegon--Holland, MI	93.0%	0.7%	4	0.012	0.021	503	1,088,514
9	Minneapolis--St. Paul, MN--WI	92.7%	0.5%	10	-0.011	0.017	1411	2,968,806
10	Los Angeles--Riverside--Orange County, CA	92.5%	0.4%	38	-0.061	0.017	4099	16,373,645
11	Kansas City, MO--KS	92.2%	0.6%	21	-0.035	0.025	753	1,776,062
12	Austin--San Marcos, TX	92.1%	0.7%	2	0.033	0.026	344	1,249,763
13	Dallas--Fort Worth, TX	92.1%	0.5%	36	-0.058	0.019	1720	5,221,801
14	Portland--Salem, OR--WA	92.1%	0.6%	5	0.009	0.019	776	2,265,223
15	Houston--Galveston--Brazoria, TX	91.7%	0.6%	17	-0.032	0.018	1413	4,669,571
16	Phoenix--Mesa, AZ	91.6%	0.7%	13	-0.022	0.018	988	3,251,876
17	Raleigh--Durham--Chapel Hill, NC	91.6%	0.9%	9	-0.004	0.028	398	1,187,941
18	Columbus, OH	91.5%	0.9%	28	-0.048	0.025	574	1,540,157
19	Milwaukee--Racine, WI	91.5%	0.7%	14	-0.023	0.023	855	1,689,572
20	San Diego, CA	91.5%	0.7%	32	-0.053	0.023	738	2,813,833
21	Detroit--Ann Arbor--Flint, MI	91.4%	0.6%	42	-0.067	0.021	1621	5,456,428
22	Indianapolis, IN	91.3%	0.8%	22	-0.036	0.024	646	1,607,486
23	Greensboro--Winston-Salem--High Point, NC	91.1%	0.9%	18	-0.032	0.024	570	1,251,509
24	Atlanta, GA	90.9%	0.6%	40	-0.064	0.024	1426	4,112,198
25	Miami--Fort Lauderdale, FL	90.9%	0.7%	35	-0.057	0.020	1010	3,876,380
26	Charlotte--Gastonia--Rock Hill, NC--SC	90.7%	0.9%	46	-0.083	0.029	618	1,499,293
27	Boston--Worcester--Lawrence, MA--NH--ME--CT	90.6%	0.5%	12	-0.022	0.015	2231	5,819,100
28	Chicago--Gary--Kenosha, IL--IN--WI	90.5%	0.4%	27	-0.047	0.016	3431	9,157,540
29	New York--Northern NJ--Long Island, NY--NJ--CT--PA	90.5%	0.4%	30	-0.050	0.015	4775	21,199,865
30	Washington--Baltimore, DC--MD--VA--WV	90.4%	0.5%	20	-0.034	0.017	2222	7,608,070
31	Philadelphia--Wilm.--Atlantic City, PA--NJ--DE--MD	90.3%	0.5%	16	-0.031	0.017	1745	6,188,463
32	Rochester, NY	90.3%	1.0%	19	-0.033	0.028	373	1,098,201
33	Hartford, CT	90.2%	0.9%	15	-0.024	0.027	500	1,183,110
34	Oklahoma City, OK	90.2%	1.1%	8	-0.002	0.024	339	1,083,346
35	Memphis, TN--AR--MS	90.0%	1.0%	26	-0.045	0.027	437	1,135,614
36	Louisville, KY--IN	89.9%	1.0%	25	-0.044	0.027	448	1,025,598
37	Cincinnati--Hamilton, OH--KY--IN	89.7%	0.8%	41	-0.066	0.024	772	1,979,202
38	St. Louis, MO--IL	89.7%	0.7%	11	-0.020	0.020	936	2,603,607
39	Pittsburgh, PA	89.1%	0.8%	34	-0.056	0.023	727	2,358,695
40	Buffalo--Niagara Falls, NY	88.5%	1.1%	31	-0.051	0.030	393	1,170,111
41	Tampa--St. Petersburg--Clearwater, FL	88.4%	0.9%	33	-0.054	0.021	812	2,395,997
42	Jacksonville, FL	87.6%	1.3%	47	-0.094	0.032	373	1,100,491
43	Las Vegas, NV--AZ	87.2%	1.2%	48	-0.106	0.030	417	1,563,282
44	Sacramento--Yolo, CA	87.0%	1.2%	45	-0.070	0.034	427	1,796,857
45	Norfolk--Virginia Beach--Newport News, VA--NC	86.9%	1.2%	49	-0.110	0.032	374	1,569,541
46	New Orleans, LA	86.0%	1.1%	37	-0.06	0.031	386	1,337,726
47	West Palm Beach--Boca Raton, FL	85.9%	1.2%	29	-0.049	0.029	299	1,131,184
48	Orlando, FL	85.5%	1.0%	44	-0.067	0.025	622	1,644,561
49	Nashville, TN	84.6%	1.1%	39	-0.062	0.028	466	1,231,311

Table 4b: Enhancement Among Metropolitan Statistical Areas with Over One Million People

Avg. Rank	City	Avg. Rate	Std. Err. Rate	Marg Rank	Marg. Coef.	Std. Err. Coef.	Obs.	Population
1	Denver—Boulder--Greeley, CO	18.3%	1.3%	3	0.016	0.015	940	2,581,506
2	San Francisco--Oakland--San Jose, CA	17.0%	0.9%	15	0 (base)	N/A	2135	7,039,362
3	Salt Lake City--Ogden, UT	16.7%	1.7%	6	0.013	0.017	535	1,333,914
4	Minneapolis--St. Paul, MN--WI	15.9%	1.0%	10	0.003	0.012	1411	2,968,806
5	Houston--Galveston--Brazoria, TX	15.7%	1.0%	11	0.003	0.012	1413	4,669,571
6	Atlanta, GA	15.4%	1.0%	26	-0.008	0.011	1426	4,112,198
7	Oklahoma City, OK	15.4%	2.0%	2	0.020	0.021	339	1,083,346
8	Dallas--Fort Worth, TX	15.3%	0.9%	19	-0.003	0.011	1720	5,221,801
9	San Antonio, TX	15.3%	1.9%	4	0.013	0.020	395	1,592,383
10	Portland--Salem, OR--WA	15.1%	1.3%	5	0.013	0.019	776	2,265,223
11	Providence--Fall River—Warwick, RI--MA	14.9%	2.2%	7	0.010	0.024	290	1,188,613
12	Austin--San Marcos, TX	14.7%	1.9%	27	-0.009	0.016	344	1,249,763
13	Cleveland--Akron, OH	14.7%	1.2%	21	-0.004	0.014	1099	2,945,831
14	Tampa--St. Petersburg—Clearwater, FL	14.6%	1.3%	8	0.009	0.015	812	2,395,997
15	Memphis, TN--AR--MS	14.5%	1.8%	14	0.002	0.021	437	1,135,614
16	Seattle--Tacoma--Bremerton, WA	14.5%	1.2%	16	-0.002	0.012	1012	3,554,760
17	Hartford, CT	14.4%	1.6%	25	-0.008	0.016	500	1,183,110
18	San Diego, CA	14.3%	1.3%	23	-0.005	0.014	738	2,813,833
19	Cincinnati--Hamilton, OH—KY--IN	14.2%	1.3%	24	-0.005	0.014	772	1,979,202
20	Washington--Baltimore, DC—MD—VA--WV	14.2%	0.8%	22	-0.005	0.010	2222	7,608,070
21	Chicago--Gary--Kenosha, IL—IN--WI	14.1%	0.7%	17	-0.002	0.009	3431	9,157,540
22	Rochester, NY	14.1%	1.9%	18	-0.003	0.018	373	1,098,201
23	Boston--Worcester--Lawrence, MA--NH--ME--CT	13.9%	0.8%	20	-0.004	0.011	2231	5,819,100
24	Detroit--Ann Arbor--Flint, MI	13.8%	0.9%	39	-0.016	0.010	1621	5,456,428
25	Kansas City, MO--KS	13.7%	1.3%	35	-0.014	0.013	753	1,776,062
26	Raleigh--Durham--Chapel Hill, NC	13.7%	1.7%	31	-0.012	0.017	398	1,187,941
27	Pittsburgh, PA	13.6%	1.3%	13	0.003	0.015	727	2,358,695
28	Indianapolis, IN	13.6%	1.4%	41	-0.019	0.014	646	1,607,486
29	Charlotte—Gastonia--Rock Hill, NC--SC	13.6%	1.5%	29	-0.010	0.014	618	1,499,293
30	West Palm Beach--Boca Raton, FL	13.6%	2.0%	12	0.003	0.025	299	1,131,184
31	Los Angeles--Riverside—Orange County, CA	13.5%	0.6%	37	-0.015	0.008	4099	16,373,645
32	Miami—Fort Lauderdale, FL	13.5%	1.1%	33	-0.013	0.011	1010	3,876,380
33	New York--Northern NJ--Long Island, NY-NJ-CT-PA	13.5%	0.6%	36	-0.015	0.008	4775	21,199,865
34	Philadelphia-Wilm.-Atlantic City, PA--NJ--DE--MD	13.3%	0.9%	44	-0.021	0.009	1745	6,188,463
35	St. Louis, MO--IL	13.2%	1.2%	28	-0.009	0.013	936	2,603,607
36	Louisville, KY--IN	13.2%	1.6%	9	0.006	0.024	448	1,025,598
37	Columbus, OH	13.0%	1.5%	30	-0.011	0.018	574	1,540,157
38	Buffalo—Niagara Falls, NY	12.9%	1.7%	42	-0.019	0.014	393	1,170,111
39	Phoenix—Mesa, AZ	12.4%	1.1%	34	-0.014	0.012	988	3,251,876
40	Greensboro—Winston-Salem—High Point, NC	12.2%	1.4%	43	-0.020	0.015	570	1,251,509
41	Grand Rapids--Muskegon—Holland, MI	12.0%	1.5%	47	-0.031	0.012	503	1,088,514
42	New Orleans, LA	11.9%	1.7%	40	-0.018	0.016	386	1,337,726
43	Milwaukee--Racine, WI	11.7%	1.2%	38	-0.016	0.014	855	1,689,572
44	Nashville, TN	11.7%	1.5%	32	-0.012	0.016	466	1,231,311
45	Jacksonville, FL	11.3%	1.7%	48	-0.034	0.014	373	1,100,491
46	Sacramento--Yolo, CA	11.8%	1.6%	1	0.041	0.050	427	1,796,857
47	Norfolk--Virginia Beach—Newport News, VA--NC	10.8%	1.7%	45	-0.021	0.017	374	1,569,541
48	Orlando, FL	10.5%	1.3%	46	-0.027	0.012	622	1,644,561
49	Las Vegas, NV--AZ	9.0%	1.4%	49	-0.043	0.012	417	1,563,282

Table 5
Effect of Population Size and Density on Adoption of Participation and Enhancement
(Standard errors in parentheses)

	Participation				Enhancement				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	
A. Coefficients from (Weighted) Probit Regressions	Small MSA	0.0095 (0.0285)				0.0198 (0.0350)			
	Medium MSA	-0.0491 (0.0227)*				0.0449 (0.0265)+			
	Large MSA	-0.0262 (0.0201)				0.0632 (0.0228)**			
	MSA Population		-3.95e-09 (3.80e-09)	1.36e-08 (1.02e-08)			-8.80e-10 (3.24e-09)	2.10e-08 (1.15e-08)+	
	MSA Population Squared			-1.94e-15 (1.13e-15)+				-2.39e-15 (1.24e-15)+	
	Medium-Low Density				-0.0170 (0.0194)			0.0275 (0.0224)	
	Medium-High Density				-0.0282 (0.0200)			0.0860 (0.0244)**	
	High Density				-0.0177 (0.0224)			0.0577 (0.0224)*	
	Log Likelihood	-33470.6	-33472.1	-33469.3	-33473.5	-28694.7	-28696.9	-28693.1	-28688.2
	Pseudo R ²	0.2252	0.2252	0.2252	0.2251	0.0593	0.0592	0.0593	0.0595
	B. Marginal Effects from (Weighted) Probit Regressions	Small MSA	0.0021 (0.0064)				0.0035 (0.0062)		
		Medium MSA	-0.011 (0.0052)*				0.008 (0.0048)+		
		Large MSA	-0.0058 (0.0045)				0.0110 (0.0039)**		
MSA Population			-8.89e-10 (8.56e-10)	3.06e-09 (2.30e-09)			-1.54e-10 (5.67e-10)	3.67e-09 (2.02e-09)+	
MSA Population Squared				-4.37e-16 (2.55e-16)+				-4.18e-16 (2.17e-16)+	
Medium-Low Density					-0.00385 (0.00440)			0.00485 (0.00399)	
Medium-High Density					-0.00639 (0.00456)			0.0154 (0.00450)**	
High Density					-0.00400 (0.00508)			0.0103 (0.00406)*	

Notes:

All regressions include dummy variables for 3-digits NAICS, month that data was collected, and whether it was a multi-establishment firm. Employment and Employment squared were also included as controls. Population was measured at the MSA level. Standard errors are in parentheses.

(1) & (5) Non-MSA is the base for these regressions.

(2), (3), (6), & (7) 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) & (8) Low density is the base for these regressions. One quarter of the observations fit into each density type.

+significant at 90% confidence level

*significant at 95% confidence level

**significant at 99% confidence level

Table 6
Average and Median Location Effects, by Type of Location

Type	N	Median Particip. Marginal Effect	Average Particip. on Marginal Effect	Std. Dev. Particip. Marginal Effect	Median Enhance. Marginal Effect	Average Enhance. Marginal Effect	Std. Dev. Enhance. Marginal Effect
Rural	49	-0.029	-0.0292	0.0486	-0.020	-0.0135	0.0274
Small MSA	130*	-0.0225	-0.0271	0.0772	-0.018	-0.00708	0.0495
Medium MSA	95	-0.046	-0.0535	0.0579	-0.012	-0.0111	0.0313
Large MSA	48	-0.0445	-0.0397	0.0324	-0.008	-0.00652	0.0150

Notes:

Author's calculation using estimates from probit models shown in Figures 1 and 2.

*N=127 for enhancement because three small MSAs perfectly predicted non-adoption.

Table 7a
Effect of Population Size and Density on Adoption of Participation and Enhancement,
Includes Multi-Establishment/Population Interactions: Coefficient Estimates
(Standard errors in parenthesis)

	Participation				Enhancement			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Multi-Est. dummy	-.1511 (0.0339)**	-0.1661 (0.0211)**	-0.1661 (0.0210)**	-0.1509 (0.0289)**	-0.0578 (0.0404)	-0.0014 (0.0210)	-0.0038 (0.0212)	-0.0463 (0.0334)
Small MSA	0.0247 (0.0430)				0.0013 (0.0492)			
Medium MSA	-0.0571 (0.0321)+				0.0317 (0.0342)			
Large MSA	-0.0463 (0.0279)+				0.0134 (0.0284)			
Small MSA × Multi-Est. dummy	-0.0267 (0.0570)				0.0532 (0.0696)			
Medium MSA × Multi-Est. dummy	0.0196 (0.0446)				0.0436 (0.0538)			
Large MSA × Multi-Est. dummy	0.0442 (0.0378)				0.1234 (0.0447)**			
MSA Population		-1.27e-08* (5.16e-09)	-3.07e-09 (9.37e-09)			-4.96e-09 (4.08e-09)	2.07e-08 (1.03e-08)*	
MSA Population Squared			-1.09e-15 (1.02e-15)				-2.92e-15 (1.13e-15)**	
MSA Population × Multi-Est. dummy		1.86e-08** (6.38e-09)	1.82e-08** (6.29e-09)			1.45e-08 (5.91e-09)*	1.44e-08 (6.01e-09)*	
Medium-Low Density				-0.0222 (0.0280)				-0.0024 (0.0294)
Medium-High Density				-0.0212 (0.0286)				0.0574 (0.0341)+
High Density				-0.0619 (0.0323)+				0.0059 (0.0289)
Medium-Low Density × Multi- Est.				0.0132 (0.0382)				0.0760 (0.0450)+
Medium-High Density × Multi- Est.				-0.0112 (0.0386)				0.0738 (0.0465)
High Density × Multi-Est.				0.0961 (0.0417)*				0.1238 (0.0435)**
Log Likelihood	-33467.9	-33463.5	-33462.5	-33465.8	-28686.8	-28695.9	-28689.2	-28681.9
Pseudo R ²	0.2253	0.2254	0.2254	0.2253	0.0596	0.0593	0.0595	0.0597

Notes:

All regressions include dummy variables for 3-digits NAICS, month that data was collected, and whether it was a multi-establishment firm. Employment and Employment squared were also included as controls. Population was measured at the MSA level.

(1) & (5) Non-MSA is the base for these regressions.

(2), (3), (6), & (7) 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) & (8) Low density is the base for these regressions. One quarter of the observations fit into each density type.

+significant at 90% confidence level

*significant at 95% confidence level

**significant at 99% confidence level

Table 7b
Effect of Population Size and Density on Adoption of Participation and Enhancement,
Includes Multi-Establishment/Population Interactions: Marginal Effects
(Standard errors in parenthesis)

	Participation				Enhancement			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Multi-Est. dummy	-0.0343 (0.0077)**	-0.0377 (0.0048)**	-0.0377 (0.0048)*	-0.0342 (0.0066)	-0.0101 (0.0070)	-0.0002 (0.0037)	-0.0007 (0.0037)	-0.0081 (0.0058)
Small MSA	0.0055 (0.0095)				0.0002 (0.0086)			
Medium MSA	-0.0131 (0.0075)+				0.0056 (0.0061)			
Large MSA	-0.0104 (0.0062)+				0.0023 (0.0050)			
Small MSA × Multi-Est. dummy	-0.0061 (0.0132)				0.0096 (0.0130)			
Medium MSA × Multi-Est. dummy	0.0043 (0.0099)				0.0078 (0.0098)			
Large MSA × Multi-Est. dummy	0.0099 (0.0083)				0.0224** (0.0084)			
MSA Population		-2.85e-09 (1.17e-09)*	-6.91e-10 (2.11e-09)			-8.68e-10 (7.15e-10)	3.63e-09 (1.81e-09)*	
MSA Population Squared			-2.44e-16 (2.30e-16)				-5.12e-16 (1.98e-16)**	
MSA Population × Multi-Est. dummy		4.19e-09 (1.44e-09)**	4.09e-09 (1.42e-09)**			2.53e-09 (1.03e-09)*	2.52e-09 (1.05e-09)*	
Medium-Low Density				-0.0050 (0.0064)				-0.0004 (0.0051)
Medium-High Density				-0.0048 (0.0065)				0.0102 (0.0062)+
High Density				-0.0141 (0.0075)+				0.0010 (0.0051)
Medium-Low Density × Multi- Est.				0.0030 (0.0085)				0.0138 (0.0085)
Medium-High Density × Multi- Est.				-0.0025 (0.0088)				0.0134 (0.0087)
High Density × Multi-Est. dummy				0.0208 (0.0087)*				0.0229 (0.0085)**

Notes:

All regressions include dummy variables for 3-digits NAICS, month that data was collected, and whether it was a multi-establishment firm. Employment and Employment squared were also included as controls. Population was measured at the MSA level.

(1) & (5) Non-MSA is the base for these regressions.

(2), (3), (6), & (7) 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) & (8) Low density is the base for these regressions. One quarter of the observations fit into each density type.

+significant at 90% confidence level

*significant at 95% confidence level

**significant at 99% confidence level

Table 8
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:

Source: Author's calculation. Pseudo-R² from full model shown in Figures 1 and 2, and subsets of coefficients controlling for industry and location effects.

Cities defined by CMSA.

Table 9
Interaction of Industry and Location Effects

	A. Leading Internet Adopters (NAICS)				B. Top IT-Using Industries (SIC)			
	Participation		Enhancement		Participation		Enhancement	
	Coefficient	Marginal Effect	Coefficient	Marginal Effect	Coefficient	Marginal Effect	Coefficient	Marginal Effect
Small MSA	-0.0127 (0.0302)	-0.0029 (0.0069)	0.0085 (0.0062)	0.0476 (0.0338)	-0.0175 (0.029)	-0.00397 (-0.0130)	-0.00153 (0.0405)	-0.000267 (0.00707)
Medium MSA	-0.0640 (0.0241)**	-0.0147 (0.0056)**	0.0013 (0.0047)	0.0073 (0.0269)	-0.0569 (0.0237)*	-0.0130 (0.00554)*	0.0279 (0.0291)	0.00494 (0.00521)
Large MSA	-0.0372 (0.0214)+	-0.0083 (0.0048)+	0.0042 (0.0040)	0.0243 (0.0228)	-0.0409 (0.0210)+	-0.00917 (0.00470)+	0.0371 (0.0247)	0.00646 (0.00430)
Top quartile NAICS3	0.8869 (0.7270)	0.1512 (0.0906)	0.0390 (0.0908)	0.2072 (0.4489)				
Small MSA* Top quartile NAICS3	0.2791 (0.0864)**	0.0539 (0.0140)**	-0.0121 (0.0169)	-0.0727 (0.1062)				
Medium MSA* Top quartile NAICS3	0.1618 (0.0692)*	0.0335 (0.0131)*	0.0397 (0.0181)*	0.2020 (0.0831)*				
Large MSA* Top quartile NAICS3	0.1205 (0.0582)**	0.0259 (0.0119)**	0.0382 (0.0155)*	0.1998 (0.0744)**				
IT intense SIC					-0.0739 (0.0702)	-0.0169 (0.0163)	-0.0672 (0.0662)	-0.0115 (0.0112)
Small MSA* IT intense SIC					0.1761 (0.0927)+	0.0361 (0.0171)+	0.0917 (0.0833)	0.0170 (0.0163)
Medium MSA* IT intense SIC					0.0683 (0.0707)	0.0149 (0.0149)	0.0773 (0.0677)	0.0141 (0.0129)
Large MSA* IT intense SIC					0.0996 (0.0620)	0.0217 (0.0130)	0.1022 (0.0595)+	0.0187 (0.0113)+
Log Likelihood	-33464.7	-33464.7	-28674.1	-28674.1	-33465.9	-33465.9	-28691.3	-28691.3
Pseudo R ²	0.2253	0.2253	0.0600	0.0600	0.2253	0.2253	0.0594	0.0594

Notes:

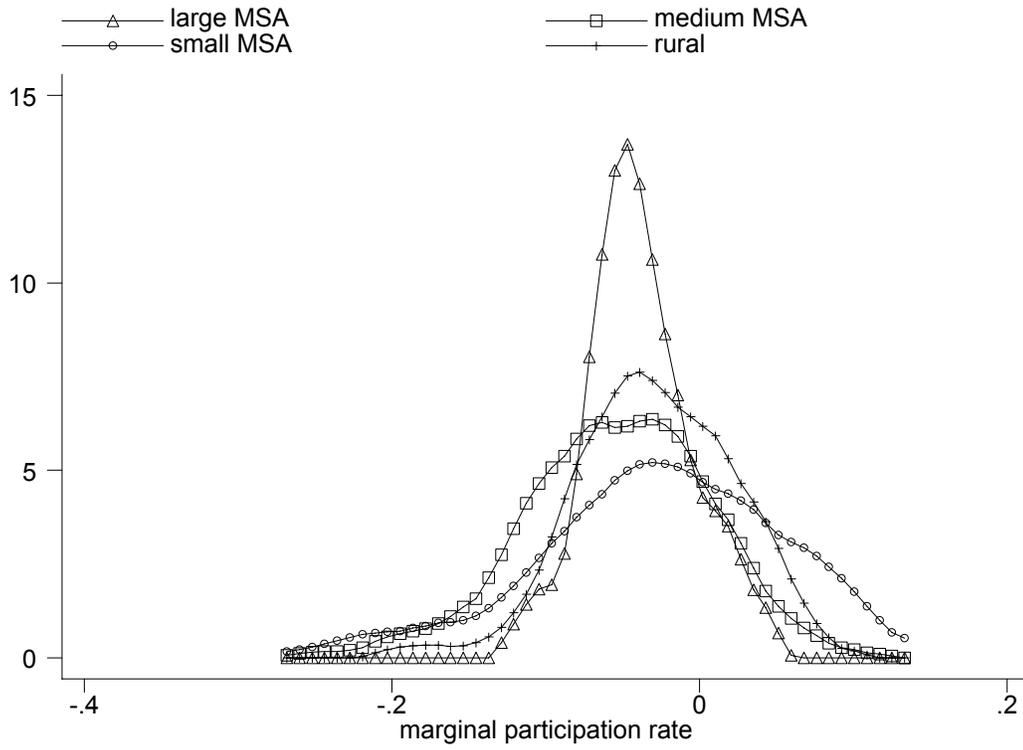
All regressions include dummy variables for 3-digits NAICS, month that data was collected, and whether it was a multi-establishment firm. Employment and Employment squared were also included as controls. Standard errors are in parentheses. Non-MSA is the base for these regressions

+significant at 90% confidence level

*significant at 95% confidence level

**significant at 99% confidence level

Figure 1
Comparison by City Size of Location Marginal Effects for Participation



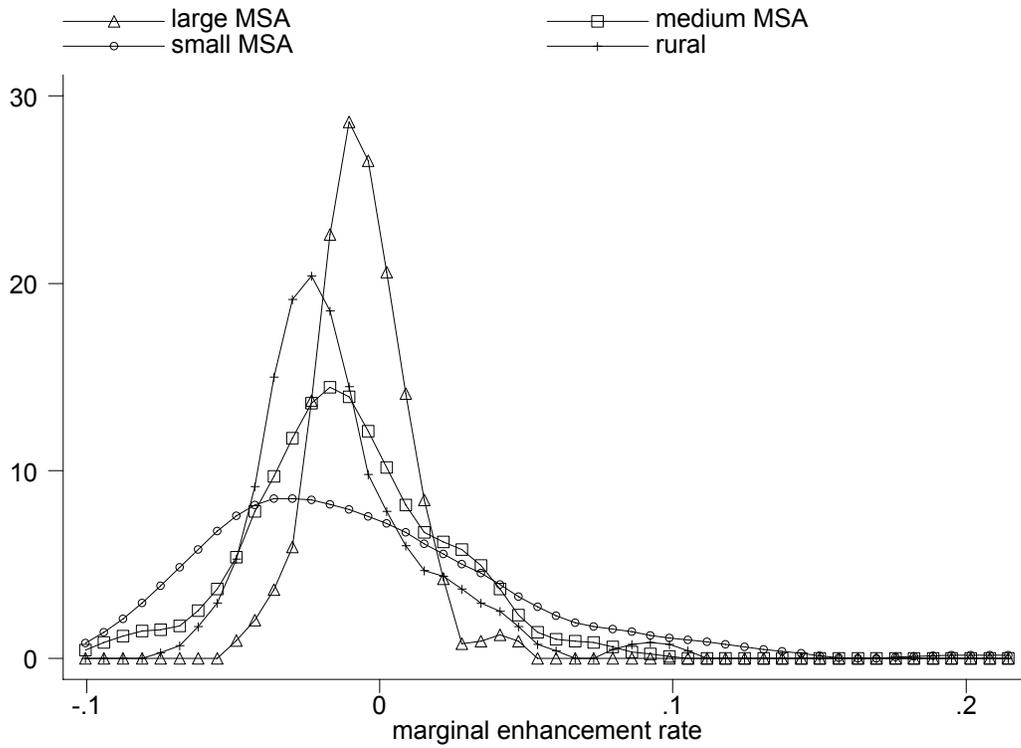
Notes:

Source: Author's calculation.

Figure shows Epanechnikov kernel density estimates of the marginal effect of location on participation, by city size.

Uses baseline probit in model (1).

Figure 2
Comparison by City Size of Location Marginal Effects for Enhancement

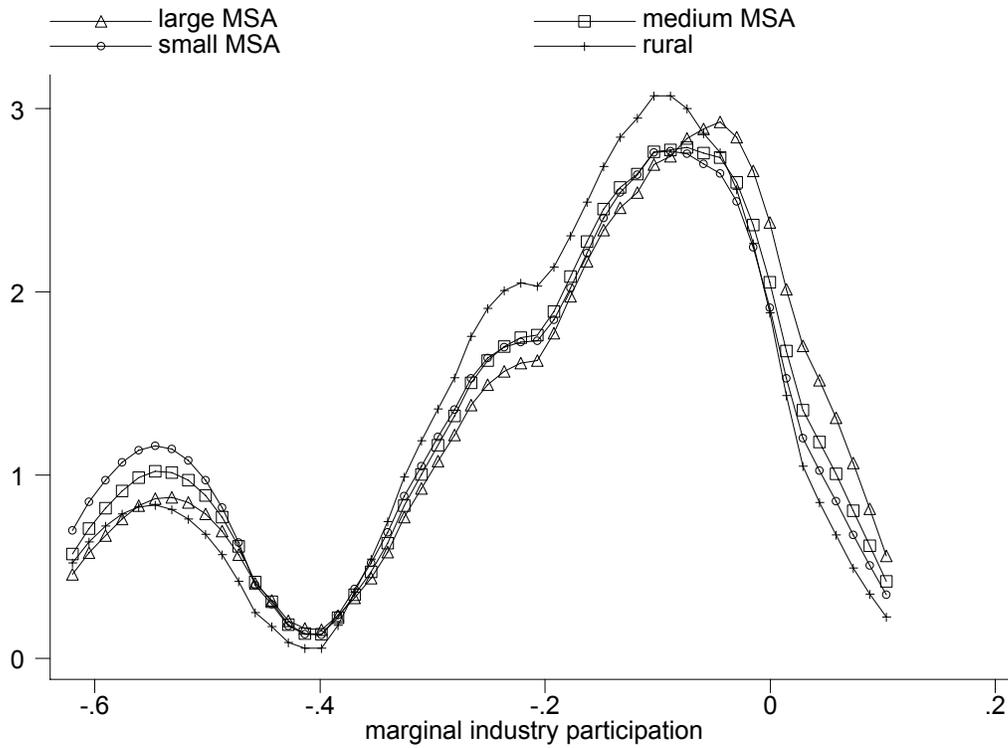


Notes:

Source: Author's calculation.

Figure shows Epanechnikov kernel density estimates of the marginal effect of location on enhancement, by city size. Uses baseline probit in model (1).

Figure 3
Industry Marginal Effects for Participation by City Size

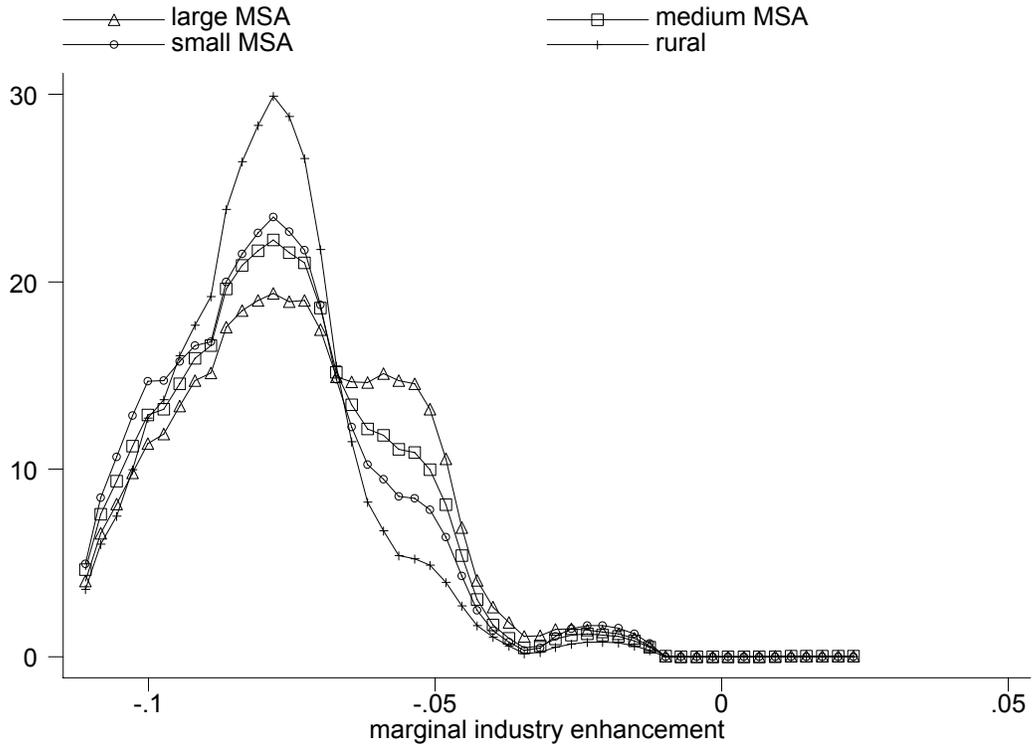


Notes:

Source: Author's calculation.

Figure shows Epanachnikov kernel density estimates of the marginal effect of industry on participation, by city size. Uses baseline probit in model (1).

Figure 4
Differences in Industry Marginal Effects for Enhancement



Notes:

Source: Author's calculation.

Figure shows Epanachnikov kernel density estimates of the marginal effect of industry on enhancement, by city size.

Uses baseline probit in model (1).

Table A1: Harte Hanks Sample Versus the Census of Business Establishments

	Sample	Census
# of Establishments with over 100 Employees	86,879	168,372
% MSA	82.5%	86.7%
% CMSA	37.2%	42.5%
% >500 Employees Given Have 100 Employees	12.7%	10.6%
% Northeast	17.7%	19.6%
% Midwest	27.9%	25.5%
% South	34.8%	34.0%
% West	19.6%	21.0%
% Agriculture, Forestry, Fishing and Hunting (NAICS = 11)	0.2%	0.1%
% Mining (NAICS = 21)	0.6%	0.5%
% Utilities (NAICS = 22)	0.8%	0.8%
% Construction (NAICS = 23)	2.9%	4.1%
% manufacturing (NAICS = 31, 32, 33)	27.9%	20.8%
% Wholesale Trade (NAICS = 42)	6.0%	4.8%
% Retail Trade (NAICS = 44, 45)	17.1%	14.7%
% Transportation and Warehousing (NAICS = 48, 49)	2.9%	3.1%
% Media, Telecommunications and Data Processing (NAICS = 51)	3.7%	3.7%
% Finance and Insurance (NAICS = 52)	4.5%	4.6%
% Real Estate and Rental and Leasing (NAICS = 53)	0.5%	1.0%
% Professional, Scientific and Technical Services (NAICS =54)	5.2%	5.0%
% Management of Companies and Enterprises (NAICS =55)	0.3%	3.2%
% Administrative and Support and Waste Management and Remediation Services (NAICS = 56)	2.7%	10.2%
% Educational Services (NAICS = 61)	0.01%	1.2%
% Health Care and Social Assistance (NAICS =62)	16.7%	12.8%
% Arts, Entertainment and Recreation (NAICS = 71)	1.6%	1.5%
% Accommodation and Food Services (NAICS = 72)	5.5%	5.1%
% Other Services (except Public Administration) (NAICS = 81)	0.9%	2.2%

Source: Author's calculation using HH and census data.

Table A.2
Population Variable Marginal Effects from Probit Regressions in Table 5,
Includes Percent Participation and Enhancement Adopters within Firm

	Model	Variable	Old	New Result
A. Weighted probits without IV	Add percentage of other establishments adopting participation to column (1)	Small MSA	0.0021	0.0021
		Medium MSA	-0.0110*	-0.0112*
		Large MSA	-0.0058	-0.0063
		Pct. shallow	N/A	0.2401**
	Add percentage of other establishments adopting enhancement to column (5)	Small MSA	0.0035	0.0038
		Medium MSA	0.0080+	0.0081+
		Large MSA	0.0110**	0.0108**
		Pct. deep	N/A	0.1026**
	Add percentage of other establishments adopting participation to column (4)	Medium-Low Density	-0.0039	-0.0039
		Medium-High Density	-0.0064	-0.0069
		High Density	-0.0040	-0.0052
		Pct. shallow	N/A	0.1637**
	Add percentage of other establishments adopting enhancement to column (8)	Medium-Low Density	0.0049	0.0049
		Medium-High Density	0.0154**	0.0152**
		High Density	0.0103*	0.0099*
		Pct. deep	N/A	0.1024**
B. Unweighted probits with IV[#]	Add percentage of other establishments adopting participation to column (1) (instrument using average population)	Small MSA	0.0032	0.0025
		Medium MSA	-0.0072+	-0.0075+
		Large MSA	-0.0045	-0.0051
		Pct. shallow	N/A	0.0193
	Add percentage of other establishments adopting enhancement to column (5) (instrument using average population)	Small MSA	0.0095*	0.0097*
		Medium MSA	0.0077*	0.0078*
		Large MSA	0.0129**	0.0128**
		Pct. shallow	N/A	0.0336
	Add percentage of other establishments adopting participation to column (4) (instrument using average density)	Medium-Low Density	-0.0019	-0.0013
		Medium-High Density	-0.0012	-0.0010
		High Density	-0.0027	-0.0028
		Pct. shallow	N/A	0.1538**
	Add percentage of other establishments adopting enhancement to column (8) (instrument using average density)	Medium-Low Density	0.0044	0.0042
		Medium-High Density	0.0167**	0.0171**
		High Density	0.0110**	0.0114**
		Pct. shallow	N/A	-0.1604

Notes:

Table compares results of probit regressions with and without variables measuring behavior of other establishments within the same firm.

[#] “Old” coefficients are different because probits are unweighted. Instruments are average population or density of locations other establishments in same firm

+significant at 90% confidence level

*significant at 95% confidence level

**significant at 99% confidence level

Table A.3
Population Variable Marginal Effects from Probit Regressions in Table 7,
Includes Percent Participation and Enhancement Adopters within Firm

	Model	Variable	Old	New Result
A. Weighted probits without IV	Add percentage of other establishments adopting participation to column (1)	Small MSA	0.0055	0.0042
		Medium MSA	-0.0131+	-0.0138+
		Large MSA	-0.0104+	-0.0111+
		Small MSA * Multi-Est. dummy	-0.0061	-0.0099
		Medium MSA * Multi-Est. dummy	0.0043	0.0054
		Large MSA * Multi-Est. dummy	0.0099	0.0098
		Multi-Est. dummy	-0.0343**	-0.1510**
		Pct. shallow	N/A	0.1636**
	Add percentage of other establishments adopting enhancement to column (5)	Small MSA	0.0002	0.0005
		Medium MSA	0.0056	0.0059
		Large MSA	0.0023	0.0029
		Small MSA * Multi-Est. dummy	0.0096	0.0096
		Medium MSA * Multi-Est. dummy	0.0078	0.0071
		Large MSA * Multi-Est. dummy	0.0224**	0.0205*
		Multi-Est. dummy	-0.0101	-0.0219**
		Pct. deep	N/A	0.1015**
	Add percentage of other establishments adopting participation to column (4)	Medium-Low Density	-0.0050	-0.0057
		Medium-High Density	-0.0048	-0.0058
		High Density	-0.0141+	-0.0145*
		Medium-Low Density * Multi-Est. dummy	0.0030	0.0042
		Medium-High Density * Multi-Est. dummy	-0.0025	-0.0016
		High Density * Multi-Est. dummy	0.0208*	0.0195*
		Multi-Est. dummy	-0.0342	-0.1505**
		Pct. shallow	N/A	0.1634**
	Add percentage of other establishments adopting enhancement to column (8)	Medium-Low Density	-0.0004	-0.0001
		Medium-High Density	0.0102+	0.0107+
		High Density	0.0010	0.0016
		Medium-Low Density * Multi-Est. dummy	0.0138	0.0131
		Medium-High Density * Multi-Est. dummy	0.0134	0.0117
		High Density * Multi-Est. dummy	0.0229**	0.0207*
		Multi-Est. dummy	-0.0081	-0.0199**
		Pct. deep	N/A	0.1014**

B. Unweighted probits with IV[#]	Add percentage other establishments adopting participation to column (1) (instrument using average population)	Small MSA	0.0050	0.0041
		Medium MSA	-0.0076	-0.0081
		Large MSA	-0.0072	-0.0077+
		Small MSA * Multi-Est. dummy	-0.0032	-0.0032
		Medium MSA * Multi-Est. dummy	0.0011	0.0011
		Large MSA * Multi-Est. dummy	0.0056	0.0049
		Multi-Est. dummy	-0.0315***	-0.0311+
		Pct. shallow	N/A	0.0150
	Add percentage other establishments adopting enhancement to column (5) (instrument using average population)	Small MSA	0.0065	0.0064
		Medium MSA	0.0029	0.0026
		Large MSA	0.0017	0.0013
		Small MSA * Multi-Est. dummy	0.0090	0.0091
		Medium MSA * Multi-Est. dummy	0.0137+	0.0140+
		Large MSA * Multi-Est. dummy	0.0288**	0.0299**
		Multi-Est. dummy	-0.0125*	-0.0037
		Pct. shallow	N/A	-0.0904
	Add percentage other establishments adopting participation to column (4) (instrument using average density)	Medium-Low Density	-0.0016	-0.0021
		Medium-High Density	0.0023	0.0013
		High Density	-0.0113*	-0.0115*
		Medium-Low Density * Multi-Est. dummy	-0.0004	0.0018
		Medium-High Density * Multi-Est. dummy	-0.0065	-0.0042
		High Density * Multi-Est. dummy	0.0175**	0.0177**
		Multi-Est. dummy	-0.0315**	-0.1316**
		Pct. shallow	N/A	0.1384**
	Add percentage other establishments adopting enhancement to column (8) (instrument using average density)	Medium-Low Density	-0.0017	-0.0024
		Medium-High Density	0.0052	0.0041
		High Density	0.0000	-0.0015
		Medium-Low Density * Multi-Est. dummy	0.0159*	0.0173**
Medium-High Density * Multi-Est. dummy		0.0279**	0.0326**	
High Density * Multi-Est. dummy		0.0279**	0.0336**	
Multi-Est. dummy		-0.0113*	0.0156	
Pct. shallow		N/A	-0.2599	

Notes:

Table compares results of probit regressions with and without variables measuring behavior of other establishments within the same firm.

“Old” coefficients are different because probits are unweighted. Instruments are average population or density of locations other establishments in same firm

+significant at 90% confidence level

*significant at 95% confidence level

**significant at 99% confidence level