The Corporate Digital Divide:

Determinants of Internet Adoption

Christopher Forman
Kellogg School of Management
Northwestern University
c-forman@kellogg.nwu.edu

May 2002

Abstract

This paper shows how organizational, technical, and environmental factors affected firm decisions to adopt Internet technologies during the early years of the commercialization of the Internet. I show that the Internet was used primarily as a basic communications technology during the early years of its diffusion. Organizations that had made prior investments in client/server networks had a higher likelihood of Internet adoption, however investments in proprietary or platform-specific client/server technologies raised the cost of switching from legacy systems. Urban firms and those that were geographically concentrated were less likely to adopt. The results suggest that low adaptation costs characterized the rapid diffusion of early Internet technologies.

This paper is based on chapter 4 of my Northwestern University dissertation. I am very grateful to Shane Greenstein for his guidance and support. I thank David Besanko, Anne Gron, Ranjay Gulati, Mike Mazzeo, Scott Stern, and seminar participants at Berkeley, Carnegie Mellon, Georgia Tech, Harvard, Illinois, Maryland, Northwestern, Rochester, Toronto, Texas, the Federal Reserve Board of Governors, the Federal Reserve Bank of San Francisco, and the 2001 Workshop on Information Systems and Economics for many helpful comments and suggestions. I would also like to thank the General Motors Strategy Center and the Kellogg School of Management for generous financial support. All errors are my own.

1. Introduction

This study examines organization decisions to adopt the Internet using a large sample concentrated in the finance and services sectors. I find that 78 percent of organizations had adopted basic Internet access and 29 percent had adopted e-commerce in 1998. The large percentages of non-adopters are surprising, given all of the promise and press associated with the Internet. Why did so many firms choose not to adopt this new technology? This paper takes a first step toward answering this question. I study organization decisions to adopt basic access and a group of more advanced e-commerce applications that enable communication between a firm, its suppliers, and its customers. The analysis shows how organizational, technical, and environmental factors affected firm decisions on whether or not to adopt the Internet by 1998.

To predict whether organizations adopt the Internet, I develop a set of theories detailing the causes of variation in the costs and benefits of adoption. I estimate a discrete choice model of organizational decisions to adopt access and e-commerce, and use the results to identify whether the empirical evidence supports each hypothesis. I analyze a sample of over 6,000 organizations surveyed by Harte Hanks Market Intelligence, concentrated primarily in the FIRE (finance, insurance, and real estate) and services sectors.

I show that during the early years of Internet commercialization, firms used the Internet primarily as a basic communications technology. Most organizations were using Internet technologies for basic research on the World Wide Web, for e-mail, or as a substitute for existing data communications systems; Internet commerce was the exception rather than the rule. Those who adopted had the greatest need for a technology that lowered communication costs.

I analyze the competing effects of prior investments in client/server technologies on the net benefits of adoption. Many Internet technologies are an extension of the client/server (C/S) computing platform, implying organizations that had previously invested in C/S will enjoy greater net benefits of adoption. Prior investment in C/S may signal high marginal benefit to decentralized computing. It may also signal the organization made complementary investments in organizational design. As a result of these factors, I find that investments in C/S increase the probability of basic access adoption.

Prior investments in C/S have an unambiguously positive impact on the probability of access adoption. However, the effects of marginal investments in C/S technologies that are proprietary or platform-specific are not unambiguous. Such investments can enable adoption because they suggest an increase in technical sophistication and familiarity with C/S technologies within the organization. They can also impede adoption if users have made complementary investments in the installed base that are incompatible with Internet protocols. On net, I find that the combined impact of these two effects is mixed, and depended on the complexity of the installed base of C/S technologies. For many organizations with complex C/S networks, the lock-in effects outweighed the effects of technical sophistication.

I study how the geographic location of employees affected the benefits of Internet adoption. I argue that geographically concentrated organizations benefited less from reductions in variable communications costs brought about by technologies like basic Internet access. The analysis shows that organizations with multiple establishments were more likely to adopt. Conditional on firm size and the existence of multiple establishments, organizations that were more geographically concentrated had a significantly lower likelihood of adoption.

I show that rurally located organizations were more likely to adopt access and e-commerce. Surprisingly, this result occurred despite evidence of lower supplies of Internet services to rural areas (Downes and Greenstein 1999). I provide evidence that, although rurally located organizations may have faced lower supplies of Internet services, they also had higher benefits to adoption. This is because there were few available communications substitutes for rural firms. This work suggests that policies to subsidize Internet access in rural areas would have little effect on the adoption of these technologies.

1.1 Prior Research

There is a large theoretical and empirical literature on firm usage of the Internet, and on e-commerce in particular. Kauffman and Walden (2001) provide an overview of the literature.[1] There is a large theoretical literature on the Internet and electronic commerce that examines pricing and product strategies in on-line markets (e.g., Bakos 1997; Bakos and Brynjolfsson 1999; Dewan, Jing, and Seidmann 2000). Several empirical papers have examined on-line pricing strategies (e.g., Clemons, Hann, and Hitt 1999; Brynjolfsson and Smith 2000; Scott Morton, Zettelmeyer, and Risso 2000; Brown and Goolsbee 2001). A burgeoning case study literature and smaller empirical literature has examined adoption (e.g., Tan and Teo 1998; Carlton and Chevalier 2000; Gertner and Spillman 2000; Chircu and Kauffman 2000) and valuation (e.g., Subramani and Walden 2001; Whinston et. al. 2001; Varian et. al. 2002) of Internet investments. In general, most prior work has focused on adoption and implementation of e-commerce technologies. These prior studies do not adequately address how, for many firms, the most valuable use of the Internet was as a basic communications technology. This paper is also unique because I study Internet adoption across a much broader sample of firms and industries than has previously been possible.

[1] There is also an extensive theoretical and empirical literature on the adoption of electronic data interchange (EDI). See, for example, Chwelos, Benbasat, and Dexter (2001).

This study views the set of communications protocols and technologies that defined the Internet as a general purpose technology (GPT).[2] Some organizations were well positioned to take advantage of this new GPT; others were not. The rate of diffusion of GPT's is often driven by the user costs of adapting new technologies to organizational needs, termed co-invention by Bresnahan and Trajtenberg (1995). Bresnahan and Greenstein (1997) show that high co-invention costs were responsible for the slow diffusion of C/S networking technologies. In this study I will argue that the reverse occurred in the early stages of Internet diffusion; low co-invention costs were responsible for rapid diffusion.

This paper shows how recent investments can sometimes hinder the adoption of new innovations. Moreover, it shows this hindrance can be greatest for technically sophisticated organizations. Theories that emphasize the age of an organization's technology portfolio as an important determinant of adoption (e.g., March and Sproull 1990; Swanson 1994) fail to consider that new systems are often incompatible with frontier technologies. Competence with newer systems can lock users in just as competencies with older ones. Moreover, the classic lead user theory of diffusion (e.g., Rogers 1995), which emphasizes the role of technical sophistication on adoption, is also insufficient. This theory fails to recognize that switching costs will be greatest when incompatible technologies are embedded in complex systems.

Because of stringent data requirements, little empirical work has shown how information technology investments can raise the switching costs of adopting new innovations.[3] Kauffman, McAndrews, and Wang (2000) show that large private ATM networks decreased the speed with

---

[2] Bresnahan and Trajtenberg (1995) define GPT's as key technologies that are characterized by their "pervasiveness, inherent potential for technical improvements, and innovational complementarities."

[3] Several recent papers have examined switching costs within the context of consumer behavior in electronic markets (e.g., Johnson et. al. 2000; Moe and Fader 2000; Brynjolfsson and Smith 2000; Chen and Hitt

which banks adopted interbank ATM networks in the mid 1980s. However, the source of switching costs in their paper was network externalities. Bresnahan and Greenstein (1997) show that slow adoption of C/S can be partially explained by lock in to the proprietary software of host-based hardware vendors. In contrast to Bresnahan and Greenstein, the experiment in this paper does not rely upon the close buyer-vendor interaction that prevailed under mainframe platforms; C/S vendors could not exert account control in the way that IBM could in the 1960s and 1970s. Moreover, the emphasis in Bresnahan and Greenstein (1997) was to show how long-run investments in mainframe software slowed adoption of a radically different platform. I show that short-run investments can slow the adoption of related information technologies.

Relatively little prior work has examined the relationship between the spatial distribution of the firm and the likelihood of adopting new innovations. Some studies have shown how the geographic distribution of establishments influences the adoption of new organizational forms. Chandler (1962) and Williamson (1975) argue that coordination and control costs are higher for firms that operate in many geographic markets, and that the multidivisional form (MDF) is a solution to this problem. Palmer et. al. (1987) and Palmer, Jennings, and Zhou (1993) examine choice of organizational form within a large sample of U.S. industrial firms, and find that geographically dispersed firms are more likely to adopt the multidivisional form. In this paper, I find that organizations adopt new communications technologies for similar reasons—to reduce the communication and coordination costs of disparate economic units.

2001). However, little work has examined how switching costs affect adoption, probably because of the difficulty of disentangling the effects of technical sophistication and lock-in.

There is a vast literature on the diffusion of new technologies.[4] This paper fits within the diffusion literature that examines the roles of heterogeneous firm incentives and environmental conditions on the adoption of new innovations. As in those papers, I distinguish between the roles played by market conditions (e.g., Hannan and McDowell 1984), geographic factors (e.g., Griliches 1957), and hypotheses related to internal firm features.

2. Early Patterns of Internet Adoption

Because of its non-commercial origins, many Internet technologies were already quite mature by the time the Internet was commercialized. Technologies for access and intranet had been perfected by years of academic and governmental use; they required little adaptation by organizations. In contrast, the complementary technologies needed to run commercial transactions successfully over the Internet were still under development. Organizations wishing to conduct commerce were forced to adapt existing Internet technologies for use in these new applications. In this section, I examine the diffusion patterns of three classes of Internet technologies: basic access, intranet, and e-commerce. The sample comes from the Harte Hanks CI Technology database (hereafter CI database), a survey of establishment technology infrastructure conducted by consultants Harte Hanks Market Intelligence.[5] The sample includes all establishments over 100 employees surveyed by Harte Hanks in SIC codes 60-67, 73, 87, and 27 over the period 1996-1998.[6]

2.1 Technology Definitions

[4] Rogers (1995) is the classic reference for work on the diffusion of innovations. Fichman (1992) provides a review of diffusion research within the IS literature.

[5] The CI is an abbreviation for Computer Intelligence, as Computer Intelligence Infocorp originally maintained this database.

[6] These are the industrial classifications for printing and publishing (27); finance, insurance, and real estate (60-67); business services (73); and engineering, accounting, and other management research and consulting firms (87). Further details on the sampling methodology are provided in Section 5 and in the Data Appendix.

Access

Employees obtain content from and send messages over the Internet by obtaining access. Access is the most basic type of Internet service; other Internet applications require access as a prerequisite. At its most basic level, access involves nothing more than users dialing up to the local Internet Service Provider (ISP), however it can also involve users connecting to the Internet through a T-1 or T-3 line.[7] Engineers had already refined the technology for access by the start of my sample, and because most early applications of access involved the retrieval of information from static web pages, adopters of basic access incurred few adaptation costs.

Intranet

An intranet is an internal network based on TCP/IP protocols.[8] Because intranets employ the same TCP/IP protocols as the Internet, organizations using an intranet can maintain a single gateway to the Internet. Organizations can use intranets to spread the cost of a T-1 or T-3 line over many users. Intranets improve security by reducing the number of openings an intruder can use to break into the organization's network. Organizations use intranets not only to connect to the Internet, but also for internal communications. For example, human resources departments use intranets to publish internal web pages describing employee benefits.

The technical and organizational costs of adoption are higher for intranet than for access. Technical costs will be higher because the network must be compatible with TCP/IP. The technical costs of migrating from a host-based (i.e., mainframe) platform to an intranet will be particularly high. Organizational costs will also be higher. For example, network administrators

---

[7] T-1 and T-3 lines are dedicated connections supporting data rates or 1.544 and 43 megabits per second. Firms commonly leased all or part of these lines from local phone companies to obtain Internet access.

[8] TCP/IP is the suite of major communication protocols that are used in the Internet.

that made specialized investments in the maintenance and management of proprietary network systems may resist adopting a technology that renders these investments obsolete.

E-commerce

The last technology is a collection of Internet applications that enable communication between the organization, its customers, and suppliers. I group these applications together because they perform similar functions and have similar adoption costs.[9]

These applications dynamically process idiosyncratic requests from customers and suppliers. They require communication between Internet applications and pre-existing firm databases, and usually require the firm to provide security and/or privacy protection to external users. Organizations adopting e-commerce must manage an entirely new distribution channel—a distribution channel that requires organizational change to firm sales and distribution techniques. Thus, e-commerce has the highest technical and organizational costs of adoption.

2.2 Technology Diffusion

In this section, I present patterns of Internet diffusion. The unit of observation in the CI database is the establishment. However, establishment-level analysis is inappropriate because an establishment's adoption decision depends on observable and unobservable characteristics of other establishments within the same organization. Because of these potential problems, I aggregate establishments up to the organization level. I perform all analyses using the organization as the unit of observation, and define an organization as the set of all establishments in the sample from the same firm.[10]

---

[9] Establishments are said to have adopted e-commerce if they use any of the following applications: business-to-business e-commerce, business-to-consumer e-commerce, customer service, education, extranet, publishing, purchasing, and technical support.

[10] A fuller analysis of the problems with establishment-level analysis, as well as a discussion of the methodology for constructing an organization-level data set, is included in the Data Appendix.

Table 1 shows penetration rates of the three technologies over 1996-1998.[11] Even in 1998, several years into the diffusion of the Internet, over 22 percent of organizations still had not adopted basic access. Given the low costs of adopting basic access, this is very surprising.

Internet technologies diffused at very different rates. Although there were many non-adopters, access diffused rapidly, achieving 33.3 percent penetration among organizations as early as 1996 and achieving 77.9 percent penetration by 1998. In contrast, the diffusion of e-commerce proceeded much slower. Only 7.1 percent of organizations had adopted by 1996, and 28.9 percent by 1998. The diffusion rates for intranet were somewhere in between. These results support the hypothesis that the fixed costs of adoption for access were far lower than those for e-commerce.

Organizations differ in their costs and benefits to adopting innovations, and respond by adopting different combinations of technologies. Table 2 shows how organizations adopted the Internet. Most organizations adopted access and intranet together, rather than adopting access alone. This pattern occurred in every year, and suggests that access and intranet have similar determinants of adoption. However, although most adopters chose to adopt intranet in conjunction with access, adoption of e-commerce lagged far behind that of other applications.

Organizations generally used access and intranet either for research, transmitting text-based information, or as a substitute for other data communications. Of those organizations adopting basic access, 71.5 percent used the Internet to do basic research on the World Wide Web and 56.6 percent used TCP/IP-based e-mail. In 1998, Harte Hanks surveyed the types of

---

[11] The size of the Harte Hanks sample increased significantly over the sample period, from 5389 observations in 1996 to 8388 observations in 1998. Throughout the study, I drop organizations that failed to report any software in the CI database. To remove concerns about changing sample composition, Tables 1 and 2 include only those observations that were in the sample in 1996. Analysis using all observations, or analysis using only organizations in the sample for all three years, gives qualitatively similar results.

applications used in internal intranets. 31.3 percent of responses were "e-mail" and 14.2 percent were "communications." In many intranet applications, employees posted information on internal web sites (19.0 percent). 15.6 percent of applications involved basic storage or retrieval of data. Only 8.7 percent of applications involved back-end processing of data and 4.5 percent involved providing information to customers.[12]

The data show substantial heterogeneity both in the rate and manner with which organizations adopted Internet technologies. By 1998, many organizations had not yet adopted any Internet technologies. Those who had adopted often adopted access and intranet together. These results suggest there exists substantial heterogeneity in the organizational and technical costs and benefits of adoption across firms. What they fail to identify, however, are the particular factors driving heterogeneity in adoption behavior. To answer this question, I will need to carefully identify the factors driving organization behavior. In the next section I present several theories to explain the variation in adoption behavior across organizations. I then develop an econometric framework to identify which are consistent with the empirical evidence.

3. Theories of Internet Adoption

Lead User Theory

The first hypothesis comes from the classic diffusion literature (e.g., Rogers 1995). It says that innovative organizations—or those that are traditionally closest to the technical frontier—will be early adopters of the Internet.

There are two interrelated hypotheses for why innovative organizations will be early adopters.[13] First, technically sophisticated organizations may have the internal skills necessary to

[12] 2.9 percent of applications involved network administration and 3.7 percent could not be classified.

[13] These hypotheses could alternatively be separated into two theories on adoption. However, because they are closely related and I am unable to identify between them in the data, I have grouped them together.

adapt TCP/IP-based technologies to user needs.[14] Second, innovative organizations will be more likely than laggards to have made the transition from host-based to C/S computing. Innovative organizations will have higher net benefits to adopting because most Internet technologies are a natural extension of the C/S paradigm.[15] Prior adoption of C/S can lower the technical costs of adoption,[16] and can lower the costs of organization adaptations if the organization has made complementary organizational investments.[17]

Competing Effects of Installed Base

The lead user theory said that prior investments in C/S increase the probability of early adoption. However, although the average effect of C/S investments is positive, marginal investments in platform-specific technologies may impede Internet adoption. The competing effects of installed base theory emphasizes how platform-specific investments raise the switching costs of migrating to a TCP/IP platform, lowering the benefits to Internet adoption. Legacy investments can impede adoption if users have developed competencies or made complementary investments in the installed base that are incompatible with new technologies. In such cases, tangible and intangible investments in the installed base may raise the costs of switching to the new technology, effectively locking in users.[18]

Lock-in can influence Internet adoption in several ways. First, investments in C/S software that is customized to current systems may be difficult to transfer to new platforms. Second, investments in proprietary vendor technologies such as Novell's NetWare may be incompatible with the Internet's protocols.

---

[14] Attewell (1992) discusses the importance of organizational know-how to the adoption of IS innovations.

[15] See Bernard (1998) and Orfali, Harkley, and Edwards (1999) on the costs of converting from proprietary C/S to an intranet.

[16] Swanson (1994) argues how aged applications systems can increase the technical costs of adoption.

The competing effects of installed base theory argues that prior investments that are current—but away from the technical frontier—can leave organizations ill-positioned to take advantage of frontier GPTs. This proposition has received comparatively little attention among empiricists studying technology diffusion.[19]

Size

The size effect theory argues that large firms will be early adopters.[20] Several hypotheses from the literature on firm size and innovativeness may explain this link. First, large firms may be early adopters because they are better able to bear the risks of a potential failure from adoption. Second, if there are substantial fixed costs to adopting e-commerce or intranet, the returns to adoption will increase if the organization can spread fixed costs over greater sales revenue. Third, large firms may benefit more from a technology that lowers communication costs among large numbers of employees.

Geographic Location

The geographic location theory says that the decreases in internal communications costs will be less valuable for organizations that are geographically concentrated. In addition, decreases in the costs of communicating with external parties will be less valuable for organizations in urban areas.

Organizations that adopt the Internet at multiple locations and have geographically dispersed establishments can send data communications over the Internet backbone. When

---

[17] Hitt and Brynjolfsson (1997) and Bresnahan, Brynjolfsson, and Hitt (2002) examine empirically the potential complementarities between IT investment and organizational design.

[18] Klemperer (1995) offers a survey of switching costs and lock-in.

[19] As noted above, Bresnahan and Greenstein (1997) and Kauffman, McAndrews, and Wang (2000) study the role of switching costs on technology adoption in varying contexts.

[20] Swanson (1994) discusses the relationship between organization size and the adoption of IS innovations. Bertschek and Fryges (2001) examine the link between firm size and probability of B-to-B e-commerce adoption among a sample of German firms.

organizations are widely dispersed, this technology, known as virtual private network, has relatively few alternatives. Traditional local area networks cannot reach over large distances, and private line services are expensive. In contrast, organizations that are geographically concentrated have many readily available low-cost substitutes to the Internet such as LANs or private lines services. The effects of geographic concentration will have the greatest impact on the adoption of access and intranet technologies that enable coordination between users within the organization.[21]

Rural organizations will have higher benefits to adopting access and e-commerce. There are two reasons for this. First, there are few readily available substitutes to the Internet for rurally located firms. Private line services are often unavailable or expensive. Second, rural organizations are, on average, farther from customers and suppliers. Rural organizations need to communicate electronically with customers and suppliers, and have a higher net benefit to adopting than urban ones.

External Environment

Recent models of diffusion have argued that industry or geographical spillovers may influence adoption through network externalities (e.g., Farrell and Saloner 1985; Katz and Shapiro 1996) or through learning effects (e.g., Irwin and Klenow 1994). Recently, Goolsbee and Klenow (2000) have examined the role of network externalities and learning effects on consumers' home PC adoption decisions. Another strand of the literature has estimated structural models that derive the effects of industry-wide diffusion on the costs and benefits to adoption (e.g., Karshenas and Stoneman 1993).

[21] Cortese (1996) describes some early uses of corporate intranets, including some examples of how geographically dispersed firms can benefit particularly from this technology.

Empirically identifying the role of spillovers or competitor actions on adoption is inherently difficult. For example, an observed positive correlation between the decision to adopt and the behavior of competitors could simply be the result of unobservable factors reflecting the relative benefits of adoption across industries. In the analysis that follows, I control for the relative benefits of adoption across industries by including SIC dummies and variables measuring competitor behavior. I also include variables measuring the adoption behavior of local agents. However, I will be unable to explicitly identify the role of competitor actions or spillovers on adoption behavior.

4. Empirical Model

I use discrete choice (e.g., McFadden 1981) to model an organization's joint decision to adopt access and e-commerce.[22] Each organization *i* associates some utility with a choice *j*, U_{ij}, where *j*=1 denotes a decision not to adopt any form of the Internet, *j*=2 denotes a decision to adopt access only, and *j*=3 denotes a decision to adopt access plus e-commerce. Utility takes the form of a random utility model, $U_{ij} = u_{ij} + e_{ij}$. An organization's utility for a choice has two components: (1) a deterministic component u_{ij} that is a function of organization characteristics and choice-specific attributes and (2) a random error term e_{ij} that captures the effects of unmeasured variables.

Researchers commonly use the multinomial logit model in discrete choice analysis. However, the multinomial logit is unattractive because of the well-known independence of irrelevant alternatives (IIA) property. This property imposes independence on the error terms of individual alternatives, and places severe restrictions on the substitution patterns in the model.

To allow for correlation in the unobservables, I employ a nested logit model. I assume the tree shown in figure 1 describes the adoption decision for an organization and assume that utility is additively separable into components that vary with the decision to adopt access and the decision to adopt e-commerce:

$$U_{ij} = \alpha_k' X_i + \beta_l' Z_i + \varepsilon_{ikl} .$$

α_k and β_l represent parameters that I estimate, while X_i and Z_i represent variables influencing the decisions to adopt access and e-commerce.[23] Following the nested logit literature (e.g., McFadden 1981), I assume the error term ε_{ikl} follows a generalized extreme value distribution. Under this distribution, errors within a nest are positively correlated, however errors across nests are uncorrelated.

The joint probability of a choice j is $P_{ij} = P_{ik} P_{il|k}$, where P_{ij} represents the joint probability of an access/e-commerce decision, P_{ik} is the marginal probability of an access choice, and $P_{il|k}$ is the probability of an e-commerce choice conditional on an access decision.

The generalized extreme value distribution implies that given a choice of access, the conditional probability of making an e-commerce decision l^* will be

$$P_{il^*|k} = \frac{\exp(\beta_{l^*}' Z_i)}{\sum_{l \in C_k} \exp(\beta_{l^*}' Z_i)} ,$$

where C_k denotes the set of choices available at node k in the tree.[24]

At the next level up, the probability of an access decision k^* will be

---

[22] The baseline model excludes organization decisions to adopt intranet because access and intranet adoption are driven by common factors. This is suggested in table 2 and shown in table 8.

[23] Because the explanatory variables do not vary by choice, the model parameters must vary by choice to obtain identification.

$$P_{ik} = \frac{a'_{k^*} X_i + l I_{ik^*}}{\sum_{k \in C} a'_k X_i + l I_{ik}},$$

where $I_{ik} = \log \left[\sum_{l \in C_k} \exp(b'_l Z_i) \right]$ is the inclusive value, the expected aggregate value of choice k.

The coefficient on the inclusive value, l, measures the dissimilarity of alternatives available to the buyer given different choices k. McFadden (1978) has shown that the nested logit model is consistent with expected utility maximization if and only if the inclusive value lies within the unit interval. I estimate the model using full information maximum likelihood.

The model above can be generalized directly to three or more levels.[25] As a robustness check of some of the hypotheses, I will also estimate a three-level nested logit model that incorporates the joint decisions to adopt access, intranet, and e-commerce. Figure 2 shows the assumed nesting of decisions in this three-level model.

One weakness of the nested logit model is that the coefficient estimates are sensitive to the assumed nesting of choices. In the baseline model, I have strong priors that figure 1 describes the appropriate nesting; an organization can adopt e-commerce only if it has adopted access. However, another potential nesting for the three choice model is the one in figure 3. As a robustness check, I estimated a model consistent with figure 3 and found the results to be qualitatively similar.[26]

The nested logit model imposes a very particular covariance structure on the error terms. One common alternative to the nested logit is the multinomial probit model. However, the

---

[24] In practice, because I do not observe organizations adopting e-commerce without access, I constrain the marginal probability of e-commerce adoption conditional on no access to be equal to one.

[25] Maddala (1983) describes nested logit models that are three levels and higher.

[26] The inclusive value parameters of both models all lie within the unit interval, so I am unable to reject one of these models based on this criteria.

multinomial probit presents identification problems for this study because the majority of my explanatory variables do not vary by choice. As shown in Keane (1992), identification in the multinomial probit model is tenuous in the absence of choice-specific variables. Movements in the coefficients mimic changes in the covariance parameters.[27] As a result, restricting the covariance parameters has little effect on the fit of the model. I estimated a version of the multinomial probit model with and without the restriction that the error terms were i.i.d. N(0,1), and a likelihood ratio test was unable to reject this constraint.[28] Because of these concerns, I present only the nested logit estimates.

5. Data

As noted above, this study uses data from the Harte Hanks CI Technology Database. Other empirical researchers have used earlier versions of the CI database,[29] however this paper uses a newer version that includes information on establishment usage of TCP/IP-based technologies. The database is unique because it contains detailed information on the hardware and software in use at individual firm establishments. Section 5.1 describes the sample, and sections 5.2 and 5.3 describe the dependent and independent variables.

5.1 Sample

I obtained data over the period 1996-1998 from the CI database. 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

---

[27] It is important to note that this model is *formally* identified, however I am unable to identify *statistically* between movements in the coefficients and parameters in the covariance matrix.

[28] Keane (1992) suggests adding exclusion restrictions, or variables that vary by choice, to obtain identification. The model is in fact estimated with two sets of exclusion restrictions, however these have insufficient explanatory power to alleviate the identification problem.

[29] See, for example, Hitt and Brynjolfsson (1997), Ito (2000), and Bresnahan, Brynjolfsson, and Hitt (2002).

computers, networking equipment, printers, and other office equipment; and (3) use of Internet applications and other networking services. Harte Hanks surveys establishments throughout the calendar year; my sample of annual data contains the most current information as of December of each year.

To keep the analysis of manageable size, I obtained data from the CI database on SIC codes 60-67, 73, 87, and 27. These SIC codes correspond to the industrial groupings on Finance, Insurance, and Real Estate (60-67); Business Services (73); Engineering, Accounting, Research, Management, and Related Services (87); and Printing and Publishing (27). I selected these industries because they are heavy users of information technology. The sample contains data from the CI database on all establishments of over 100 employees in these industries. All establishments are from the U.S.

I use 1998 adoption data to estimate the models described in section 4. In the baseline model, I use prior year data on firm characteristics as explanatory variables.[30] As a result, each observation requires two consecutive years of data. Because establishments enter and exit the database, this method of data construction required that I drop some organizations from the analysis sample. Moreover, I dropped some observations due to missing data. The establishment-level data originally obtained from the CI database contained 18,725 *establishments* in 1998. The final analysis sample contains 6156 *organizations.*

5.2 Variables measuring Internet adoption

In this paper, I focus on an organization's first adoption of technology. An organization is coded as adopting a technology if the technology was acquired by at least one establishment within that organization by 1998. The CI database includes several measures of establishment

[30] I also estimate models using 1995 explanatory variables as robustness checks.

Internet use. The Data Appendix describes how I use these measures to construct the adoption variables.

5.3 Exogenous Variables

Table 3 contains the names of the variables, a short description, and their descriptive statistics. Descriptions of the variables are organized by hypothesis below.[31]

Lead User

The variable *PCPEREMP* represents the number of PCs per employee in the organization. Organizations with a high PC-to-employee ratio are likely to have more decentralized computing structures, potentially implying lower organizational costs of adoption. Moreover, because PCs represent the most common way of accessing the Internet, a high value for *PCPEREMP* likely represents lower technical costs of adoption.

CLIENT indicates the percentage of establishments within the organization that have installed Internet-ready clients.[32] *CLIENT* indicates a minimal level of technical sophistication necessary for Internet adoption. Low values of *CLIENT* should imply a lower probability of adopting access or intranet.

NOAPP is a dummy variable that is one when an organization reports zero software applications.[33] Because many explanatory variables describe software use at the firm, I include this variable as a control.

---

[31] Note that organizational measures of the variables below are derived by aggregating the related establishment measures. The method of aggregation will sometimes depend on the variable. The description column in Table 3 will provide an explanation of the method when it is not clear from the context.

[32] For the purposes of this paper, Internet-ready clients include those with the following operating systems: UNIX, Windows, Macintosh OS, and several smaller others. Unfortunately, I was unable to identify between sites with Windows 95 or above versus those with Windows 3.1 or below, however I was able to identify and exclude sites with DOS installed.

[33] Establishments surveyed for the CI database do not supply a complete listing of software to Harte Hanks; rather, they report the most important software in use. Thus, a value of 1 for *NOAPP* may indicate nonresponse or a potential lack of technical sophistication at the site.

The variable *PCTLAN* indicates the percentage of customized (non-PC) applications that are accessed through an organization's LAN. I code organizations without customized applications as zero. Organizations with a high value for *PCTLAN* are ones that have adopted C/S heavily. These organizations should have lower technical and organizational costs of adoption.

SYSCOM indicates that an organizations uses system LAN software applications.[34] *NETWARE* is a dummy that indicates an organization uses some version of Novell's NetWare, a LAN operating system best known as a file server. The variable *INTRANETWARE* indicates that an organization has invested in Novell's IntranetWare operating system, a recent version of NetWare that supports the Internet's TCP/IP protocols. The effects of *INTRANETWARE* are added to those of *NETWARE*, so that the net effect of investments in *INTRANETWARE* on access is $b_{NETWARE} + b_{INTRANETWARE}$.[35] *MSNT* is a dummy that indicates use of NT, a recent network operating system developed by Microsoft. Since respondents to the Harte Hanks survey list only those applications they feel are important, a value of one for any for these variables represents a well-developed C/S network at the firm.

The variable *PCTMAIN* represents the percentage of customized (non-PC) applications that are accessed over a mainframe or minicomputer. Investments in mainframe software should decrease the net benefits to adoption.

Competing Effects of Installed Base

SYSCOM and *NETWARE* indicate prior investments in C/S that may lower the technical and organizational costs of adoption. However, they also indicate investments in platform-

---

[34] System software in this data set represents primarily software for the management, maintenance, and backup of LANs.

specific technologies that may lower the net benefits to switching to a new platform. To identify the effects of *SYSCOM* and *NETWARE* on switching costs, I interact each with *PCTLAN*. For example, in an organization with *NETWARE* and *SYSCOM*, the net effect of *PCTLAN* on access adoption is

$$(b_{PCTLAN} + b_{NETWARE \times PCTLAN} + b_{SYSCOM \times PCTLAN})PCTLAN \text{ .}$$

A negative coefficient on these interaction terms indicates that the positive impact of *PCTLAN* on Internet adoption is weakened if some C/S investments were directed toward technologies that are incompatible with Internet protocols. This negative impact will increase as *SYSCOM* and *NETWARE* are integrated into larger and more complicated networks (i.e., as *PCTLAN* increases).

I also interact *INTRANETWARE* with *PCTLAN*. I include this term to show that $NETWARE \times PCTLAN$ captures switching costs, rather than unobservable factors correlated with *NETWARE* use that decrease the returns from *PCTLAN*. For an organization with *INTRANETWARE* but not *SYSCOM*, the net effect of PCTLAN on access adoption is

$$(b_{PCTLAN} + b_{NETWARE \times PCTLAN} + b_{INTRANETWARE \times PCTLAN})PCTLAN \text{ .}$$

Because *INTRANETWARE* supports Internet protocols, the switching costs of adopting TCP/IP from an IntranetWare network should be negligible. Therefore, if $b_{NETWARE \times PCTLAN} < 0$ indicates the effects of switching costs, then we expect $b_{INTRANETWARE \times PCTLAN} > 0$. The lock-in effects of *NETWARE* will be reduced if some of an organization's *NETWARE* investments were not specific to its native IPX protocols.[36]

---

[35]This example ignores the effects of the interaction terms described below. Also, *NETWARE* and *INTRANETWARE* will also impact the decision to adopt access through the inclusive value term.

[36] IPX is the networking protocol used by Novell's NetWare.

I also interact *MSNT* with *PCTLAN*. Microsoft NT was a new LAN operating system that had only recently diffused widely. Thus, the vintage of NT software is younger, and should be more compatible with TCP/IP. Thus, if $NETWARE \times PCTLAN$ captures switching costs, one should expect $b_{MSNT \times PCTLAN}$ to be less negative than $b_{NETWARE \times PCTLAN}$.

Size

To capture the effects of organization size on the probability of adoption, I include the natural logarithm of the total number of employees in the organization. Because I expect the marginal effects of increases in the number of employees to vary with number of employees, I allow the coefficient on this variable to vary across ranges of firm size. Formally, $b_{EMPLE} EMPLE$ captures the effects of organization size on access adoption, where $EMPLE$ is the log of the number of organization employees and b_{EMPLE} is defined as

$$b_{EMPLE} = \begin{cases} b^1_{EMPLE}, & 100 \le EMPLE < 200 \\ b^2_{EMPLE}, & 200 \le EMPLE < 500 \\ b^3_{EMPLE}, & 500 < EMPLE \end{cases}.$$

I model the effects of *EMPLE* on e-commerce adoption similarly.

The variable *TOTMIPS* indicates the sum of *MIPS* across the organization's mainframes, minicomputers, and servers.[37] It measures the scale of the computing infrastructure within the organization.

Geographic Concentration

The variable *MULTEST* indicates the organization has more than one establishment. I interact *MULTEST* with *EMPCON* to obtain a measure of the concentration of employees within an organization. I calculate *EMPCON* by summing the squared shares of employees in each

---

[37]This variable does not include the computing power of the organization's PCs.

establishment across all establishments in the organization.[38] A high value for *EMPCON* indicates that employees are concentrated within a small number of establishments, and should lower the probability of adoption.

I interact *MUTLTEST* and *PCT100* to identify how greater geographic dispersion influences adoption. To calculate *PCT100*, I use longitude and latitude to find the distance between each establishment in an organization.[39] I then calculate the percentage of pairwise establishment combinations that are within 100 miles in distance. A high value for *PCT100* indicates greater concentration of organization employees, and should lower the probability of adoption.

The variable *PCT500* is the direct analog to *PCT100*, computed using a 500 mile radius. *AVGDIST* indicates the log of the average distance between establishments within the organization. I include *PCT500* and *AVGDIST* as robustness checks.

The variable *URBAN* indicates the percentage of establishments located in a major metropolitan statistical area (MSA). It identifies whether the benefits to Internet adoption are lower for organizations located in urban areas.

External environment

---

[38]Formally, let e_{ij} be the number of employees in establishment *i* and organization *j*, and $E_j = \sum_{i=1}^{N_j} e_{ij}$, where N_j is the number of establishments in organization *j*. Then, s_{ij}, the share of employees in organization *j* that are in establishment *i* will be $s_{ij} = e_{ij} / E_j$. $EMPHERF = \sum_{i=1}^{N} s_{ij}^2$.

[39] Wallsten (1999) used this approach to calculate distances between firms to measure the effects of agglomeration and spillovers on the probability of obtaining a grant from the federal government's Small Business Innovation Research (SBIR) program.

The variables *PCIACC*, *PCIINT*, and *PCICOM* indicate the percentage of competitors in the organization's three-digit SIC code that have adopted access, intranet, and e-commerce.[40] These variables are included to capture the role of competitor actions on firm behavior. They also capture differences in the relative value of Internet technologies across industries.

PCNTYACC, *PCNTYINT*, and *PCNTYCOM* indicate the percentage of establishments in the organization's county that have adopted access, intranet, and e-commerce. For multi-county organizations, I (arithmetically) average across counties. My sample contains some counties that contain zero or one observations. In these cases, I normalize the variables to equal zero. *DNCTYZERO* is a control that indicates Harte Hanks sampled no other establishments in the same county.

To control for industry effects, I include variables indicating the percentage of an organization's establishments in SIC codes 60-67, 73, and 87. SIC 27 is the omitted category. In alternative specifications, I used three-digit SIC code dummies. In some specifications, I interact dummy variables indicating the top twenty U.S. metropolitan areas with employment. I do this to capture heterogeneity in large/small firm differences across major metropolitan areas.

6. Results

This section presents the results of the empirical models. The results are organized by hypothesis. I also present extensions to the baseline model as robustness checks for each of these theses.

6.1 Effects of Internal Organizational Factors

---

[40] This variable is averaged across industries in the case of multi-industry organizations.

Table 4 presents the baseline results. To ease interpretation of the coefficients, Table 5 shows the marginal effects of a change in each dependent variable on the probability of access and e-commerce adoption.[41]

• **Lead User:** Table 4 provides evidence that technical sophistication and prior investment in C/S technologies play a role in Internet adoption, particularly in the adoption of access. The coefficients on *PCPEREMP*, *CLIENT*, and *NOAPP* are significant and consistent with the lead user theory. Organizations with low values of these basic measures of technical sophistication are unlikely to adopt even basic access. The expected probability of access adoption for a typical organization in which *PCPEREMP* and *CLIENT* are each one standard deviation below their means and for which *NOAPP*=1 is only 39.1 percent.[42]

Table 4 shows that prior investments in C/S technologies, as measured by *PCTLAN*, significantly increase the probability of access adoption but have less impact on the probability of e-commerce adoption. *PCTLAN* has a statistically significant effect on access but a statistically insignificant impact on e-commerce. An increase in *PCTLAN* from 0 to 1 increases the probability of access adoption by 18.4 percent, but increases the probability of e-commerce by only 10.5 percent. Moreover, due to the nested structure of the discrete choice model, some of this 10.5 percent increase is caused by *PCTLAN*'s effect on access adoption. Conditional on access adoption, an increase in *PCTLAN* from 0 to 1 raises the probability of e-commerce

---

[41] For variables with a range between zero and one (dummy variables or percentages), the table shows how a change from zero to one affects the probability of adoption. For continuous variables that have a range outside of the unit interval (e.g., employment or total MIPS), the table shows the effects of a one standard deviation increase in the variable. The marginal effects in the table are computed by averaging the effects across all organizations in the sample.

[42] This typical organization has 200 employees and mean values for *SIC60-SIC87*, *URBAN*, *PCTMAIN*, and *TOTMIPS*. It also has mean values for the variables measuring external environment (e.g., *PCNTYACC*). This organization, because it reports no applications, also has values of zero for *PCTLAN*, *SYSCOM*, *NETWARE*, *SYSCOM*× *PCTLAN*, and *NETWARE*× *PCTLAN*.

adoption by only 6.6 percent. For the remainder of this study, I will present the marginal effects on e-commerce only in this conditional probability form.

Investments in *NETWARE* and *SYSCOM* have a statistically significant effect on access, increasing the probability of access adoption by 8.1 and 5.0 percent. However, *NETWARE* and *SYSCOM* have no significant effect on e-commerce, with marginal effects of 4.0 and 2.6 percent.

Why do *PCTLAN*, *NETWARE*, and *SYSCOM* affect access but not e-commerce? Access generally requires complementary investments in an organization's network to be effective. In contrast, third-party providers often host e-commerce applications. E-commerce applications generally do not require the pervasive use of C/S that is captured by *PCTLAN*, *SYSCOM*, and *NETWARE*.[43]

The coefficients on *PCTMAIN* have the expected negative effect on the probability of access and e-commerce adoption, however are economically and statistically insignificant.[44]

• **Competing Effects of Installed Base:** The results show that recent investments in platform-specific technologies can lower the net benefits to access adoption. The coefficients on *NETWARE* × *PCTLAN* and *SYSCOM* × *PCTLAN* have a negative impact on both adoption decisions and a statistically and economically significant impact on the decisions to adopt access. Changing *NETWARE*×*PCTLAN* from zero to one decreases the probability of access adoption by 18.2 percent and decreases the conditional probability of e-commerce adoption by 5.5 percent. The marginal effects for *NETWARE* show that for very complex networks—or ones for

---

[43] The fact that I am analyzing the diffusion of early Internet applications is important. More recent e-business implementations, as described by, for example, Choi and Whinston (2000) or Kalakota and Robinson (2001), emphasize the integration of e-commerce applications with the internal information systems infrastructure.

[44] This result is not due to collinearity between *PCTLAN* and *PCTMAIN*. The omitted categories of software in the adoption equation are applications on minicomputers, small servers, and workstations that are not accessed by users over the LAN. I also estimated the adoption equation without *PCTLAN* and any of its interaction terms. *PCTMAIN* remained insignificant.

which *PCTLAN* is large—the negative lock-in effects will more than offset the benefits of technical sophistication. The results for *SYSCOM* are qualitatively similar.

On net, do *NETWARE* and *SYSCOM* have a positive or negative impact on access adoption? In other words, which is more important, technical sophistication or lock-in? The answer will depend on the complexity of the organization's C/S network. If the installed base of LAN applications is relatively small, then the positive impact of technical sophistication will outweigh the negative effects of additional switching costs. Investment in *NETWARE* and *SYSCOM* will then increase the probability of access adoption. Conversely, if the installed base of LAN applications is very large, then the effects of large switching costs predominate, lowering the likelihood of access adoption.

In most cases, the positive effects of technical sophistication will outweigh the negative effects of lock-in. Simulations of the effects of *NETWARE* on access adoption show that *NETWARE* increases the probability of access adoption for 65.5 percent of organizations. For *SYSCOM*, the corresponding probability is 65.2 percent.

• **Size:** Tables 4 and 5 shows that size has little effect on access adoption. This result is different from that argued by Swanson (1994) and is also different from some classic articles in diffusion research (e.g., Moch and Morse 1977; Kimberly and Evanisko 1981; Hannan and McDowell 1984). Increases in computing power—as measured by *TOTMIPS*— have no effect on access adoption. Total employment also has little effect on access adoption, except for the largest organizations. Among organizations over 500 employees, a one standard deviation increase in employment (equivalent to roughly 259 employees) increases the likelihood of adoption by only 3.3 percent.

Bresnahan and Greenstein (1997) also find no clear size effect in their study of client/server adoption. Distributed computing technologies, whether traditional client/server or TCP/IP-based, are attractive to organizations of all sizes.

Size does play some role in decisions to adopt e-commerce, however. Among organizations greater than 500 employees, a one standard deviation increase in employment increases the probability of e-commerce adoption by 7.7 percent. The coefficient on *TOTMIPS* is small but is significant, suggesting that very large computing stocks will have an economically significant impact on e-commerce adoption. A one standard deviation increase in *TOTMIPS* will increase the likelihood of e-commerce adoption by 4.4 percent. These size results probably reflect the fact that organizations adopting e-commerce faced sizable fixed costs during the early years of the commercialization of the Internet.[45] Such fixed costs could only be recovered by organizations with sufficient scale.

• **Robustness checks:** Tables 6 and 7 include the results of alternative models. I use these alternative models as robustness checks of the lead user and competing effects of installed base theories.

One potential concern with the baseline model is that the term $NETWARE \times PCTLAN$ captures unobservable factors related to NetWare use that lower the marginal benefit of prior LAN investment, but which are unrelated to the competing effects of installed base theory. To investigate this possibility, I re-estimate the model using different vintages of NetWare and different types of LAN operating systems.

---

[45] This result depends on the early sample period of this study. Organizations adopting e-commerce faced many security and privacy problems that had not yet been fully solved. Internet Service Providers had also not yet begun to offer e-commerce hosting services.

INTRANETWARE has a statistically insignificant effect on both access and e-commerce adoption, and has little effect on the coefficient estimate of *NETWARE* in either equation. This suggests that, by itself, use of *INTRANETWARE* adds little additional information. However, the interaction term $INTRANETWARE \times PCTLAN$ has a statistically significant positive coefficient estimate of 1.8557. Investment in *INTRANETWARE* removes the negative effects of *NETWARE*. While organizations investing in NetWare alone were more likely to adopt access than those without NetWare for 65.5 percent of organizations in our sample, investment in *NETWARE* combined with *INTRANETWARE* increased the likelihood of access adoption for all organizations.[46]

I used NetWare as a proxy for the competing effects of installed base theory because of (1) the proprietary nature of its IPX protocols and (2) its large installed base of pre-Internet systems. Microsoft's NT server operating system had a much younger installed base than NetWare, and discussions with networking professionals suggest it was on average easier to integrate with Internet protocols than the older NetWare systems.

I re-estimated a version of the model that included Microsoft's NT server operating system. If the negative coefficient on $NETWARE \times PCTLAN$ captures switching costs, then $MSNT \times PCTLAN$ should have little effect on adoption. This is the pattern that I observe. The term *MSNT* has a statistically significant positive effect on the probability of access adoption, while the interaction of *MSNT* with *PCTLAN* has no effect on access adoption. *MSNT* had no lock-in effect and unambiguously increased the likelihood of access adoption.

If $NETWARE \times PCTLAN$ reflects the effects of unobservable factors other than lock-in, these unobservables would have to be brand-specific. Moreover, they would have to be specific

---

[46] IntranetWare added Internet accessibility features to NetWare, but did not require organizations to run IP

to vintages of NetWare prior to IntranetWare. Although the existence of such unobservables is possible, it is difficult to identify what they might be.

I re-estimate the model using three-digit SIC codes and three-digit SIC codes plus interaction of employment with the top 20 U.S. metropolitan areas.[47] I include these additional terms to capture the effects of unobservable industry and organization size effects that vary across urban and rural areas. The coefficient estimates remain almost identical to those in the baseline model.

One could also criticize the aggregation of establishments to the organization level. Because only one establishment needs to adopt for the organization to be coded as an adopter, the behavior of "outlier" establishments could be driving the results. If these outliers are unrepresentative of establishments in the population, then the results may be unrepresentative of the factors driving Internet adoption. As a further robustness test, I re-estimate the model using only data on single-establishment organizations. The results are in the last column in table 6. The coefficient estimates are very similar to those in the baseline model. In fact, the magnitude of the effects of our key variables—*PCTLAN, SYSCOM*, *NETWARE*, *NETWARE*×*PCTLAN*, and *SYSCOM*×*PCTLAN*—are even stronger than in the baseline model. Thus, outlier establishments do not drive the results.

Another potential criticism of the model relates to the use of prior year variables as covariates. If organizations change their investments in information technology in advance of planned adoption of Internet applications, then our covariates will be endogenous, yielding inconsistent estimates of the model parameters. The last column of table 7 includes the results of

---

through their networks. Thus, IntranetWare is not a perfect predictor of access or intranet adoption.

[47] In these models, I drop *PCIACC* and *PCIADV* because there exists no within-industry variation in adoption rates.

using 1995 covariates to explain 1998 adoption decisions.[48] Because strong anecdotal evidence suggests that the Internet took many firms by surprise in 1995, we can safely assume that 1995 covariates are exogenous with respect to the error terms. [49] Unfortunately, due to entry and exit of organizations from the sample, this required me to discard over 39 percent of the observations from my sample. In the access equation, variables generally retain their sign, however they often lose their statistical significance. *PCTLAN* retains an economically and statistically significant effect on the decision to adopt access, with a coefficient estimate of 0.6989. However, while *SYSCOM*, *NETWARE*, $NETWARE \times PCTLAN$, and $SYSCOM \times PCTLAN$ retain their signs, they are no longer statistically significant. The marginal effects of *NETWARE* and $NETWARE \times PCTLAN$ on access adoption are 2.8 percent and -4.1 percent, while the marginal effects of *SYSCOM* and $SYSCOM \times PCTLAN$ are 1.7 percent and -6.9 percent. Not surprisingly, an organization's 1995 installed base has a similar, but weaker, impact on 1998 adoption than does its 1997 installed base.

Because of the problems with using 1995 variables to explain 1998 adoption decisions, I also re-estimated the baseline model using 1995 variables to explain 1996 adoption. The factors affecting the decision to adopt Internet technologies by 1996 are likely to differ somewhat from those affecting the decision to adopt by 1998. However, there are likely to be similarities. In the access equation, the variables *CLIENT*, *NOAPP*, *PCTLAN*, *SYSCOM*, and $SYSCOM \times PCTLAN$ have identical signs to the baseline model and have economically and statistically significant effects. There are two major differences with the baseline results. *PCTMAIN* now has an

---

[48] In this model, prior year variables continued to be used for the external environment variables *DCNTYZERO*, *PCNTYACC*, *PCIACC*, *PCNTYADV*, and *PCIADV*.

[49] For instance, as has been noted in antitrust proceedings, Bill Gates gives almost no attention to Internet technologies in the first edition of his book *The Road Ahead*, which provided a roadmap for Microsoft's vision of the future.

economically and statistically significant effect on access adoption, where before it did not. In contrast, *NETWARE* and $NETWARE \times PCTLAN$ have no impact on access adoption. I have some explanation for the first result, but less explanation for the second.

The slow and ongoing transition from mainframe to C/S computing throughout the 1990s may be driving the *PCTMAIN* results. The mean of *PCTMAIN* fell from 25.0 percent in 1995 to 16.8 percent in 1997. By 1998, many organizations had completed their transition from mainframe to C/S. Those mainframe applications remaining, such as back office accounting or transaction processing, involved little coordination with the remainder of the computer network. These applications would have relatively little effect on Internet adoption.

6.2 Effects of Geographic Location

The results for geographic location are divided into the effects of geographic concentration and the role of local population density.

• **Geographic concentration:** The proxies for geographic concentration have among the strongest effects on the probability of adoption. Even controlling for organization size, multi-establishment organizations have a much higher probability of adopting both Internet technologies. The marginal effects of *MULTEST* on access and e-commerce adoption are 21.7 percent and 27.2 percent.

The variable *EMPCON* captures the effects of employee dispersion in a multi-establishment organization. $MULTEST \times EMPCON$ has a strong and significant effect on the probability of access and e-commerce adoption. For multi-establishment organizations, an increase in *EMPCON* from 0 to 1 will decrease the probability of access (-42.9 percent) and e-commerce (-24.2 percent).

The coefficient on $MULTEST \times PCT100$ captures the pure effects of geographic dispersion on adoption. $MULTEST \times PCT100$ has an economically and statistically significant effect on access adoption; an increase in $MULTEST \times PCT100$ from 0 to 1 decreases the likelihood of adopting access by 8.9 percent. However, as predicted by the geographic location theory, *PCT100* has little effect on e-commerce adoption.

The results in the baseline model are consistent with the hypothesis that geographic dispersion increases the likelihood of adopting Internet technologies that lower within-organization communication costs, but has little impact on adopting technologies that lower communications costs across organizations. To confirm this hypothesis, I expand the choice set in the discrete choice model and examine the joint decision to adopt access, intranet, and e-commerce in 1998. As described in section 4, I continue to use a nested logit model and assume a tree consistent with figure 2. The results are presented in Table 8. In the new model, *PCT100* has a strong, statistically significant effect on the decision to adopt intranet (conditional marginal effect –15.9 percent), but has no statistically significant effect on the decision to adopt access or e-commerce.[50]

• **Population density:** Table 4 shows that the coefficient for *URBAN* is negative in both the access and e-commerce equations. *URBAN* has a statistically significant effect on e-commerce adoption (marginal effect -6.7 percent) and has an effect on access that is just barely statistically insignificant at the ten percent level (10.1 percent significance; marginal effect -4.3 percent). This supports the proposition in the geographic location theory: urban organizations will be less likely to be early adopters than rural ones. One explanation for this result is that rural

[50] To save space, I have not included the coefficient estimates for the final level of the model that identifies the factors affecting the e-commerce decision. I have also estimated a version of the model in which e-commerce is

organizations have less access to packet-switching and private line services that are readily available to urban firms. The data support this hypothesis. In 1995, 37.5 percent of rurally located firms had private line services, while the comparable figure for urban firms was 45.1 percent. Although these are simple univariate statistics that may be capturing systematic differences among urban and rural firms, they are suggestive.

Overall, the results confirm the hypothesis that organizations that are geographically concentrated or located in urban areas will be less likely to adopt Internet applications.

• **Robustness checks:** Table 9 includes the results of robustness checks of the geographic location theory.

As discussed above, the proxies for geographic concentration could be capturing the unobservable effects of organization size. I control for organization size with variables describing employment, total MIPS, and through the use of multi-establishment dummies. Still, there may be some unobserved factor such as "network complexity" that is correlated with organization size and which I am unable to control for. These unobserved factors could be driving the results. To see whether the empirical evidence is consistent with this alternative theory, I examine the effects of alternative distance measures on the likelihood of adoption.

The geographic location theory argues that organizations operating in dispersed geographic areas will be more likely to adopt Internet technologies because of the high cost of communications substitutes such as private line or packet-switching services. The pricing of these communications substitutes is somewhat discontinuous; prices increase once communications must extended between (as opposed to within) metropolitan areas. Organizations may choose to adopt the Internet because Internet services are free from long

---

above intranet in the tree. I am unable to reject that model, however the results are qualitatively similar to those distance charges. I included *PCT100* in the baseline model to capture precisely this phenomenon: controlling for multiple establishments and employee dispersion, *PCT100* captured the marginal effect of establishments being located within the same metropolitan area.

If *PCT100* captures a discontinuity in the pricing of communications substitutes, other distance measures should have less power in explaining the variation in adoption decisions. However, if *PCT100* captures unobservable effects of organization size, then other distance measures should have equal power in explaining adoption. To examine this hypothesis, I re-estimate the model using the alternative distance measures *PCT500* and *AVGDIST*.[51] Both are unsuccessful in explaining the variation in access adoption decisions.

Organization location decisions could also be endogenous with the cost of communications. The availability of low-cost Internet services may cause organizations to locate establishments farther apart, implying that the geographic location theory reverses the true direction of causality. I explore this possibility by re-estimating the model using 1996 adoption decisions and 1995 explanatory variables. Organization location decisions made in 1995 or before were made prior to the widespread diffusion of the Internet; these location decisions will be exogenous with Internet adoption.[52] In this model, the effects of geographic location on access adoption are identical. Changing *PCT100* from 0 to 1 increases the probability of access

---

presented in the text.

[51] Under the alternative size hypothesis, the true explanatory variable would be (*1-PCT100)*, indicating large firms with complex networks are most likely to adopt. Variables indicating large distances between establishments such as (*1-PCT500*) or *AVGDIST* should have equivalent explanatory power.

[52] I have also estimated the model using 1998 adoption decisions and 1995 covariates. As described above in section 6.1, this model is estimated on a much smaller sample due to entry and exit of organizations from the database. In this alternative model, the signs of the coefficients are identical to those in the baseline model but statistical significance is generally lost.

adoption by 12.3 percent, compared to an 8.9 percent marginal effect in the baseline adoption equation.[53]

6.3 Effects of External Environment

Many of the SIC variables are significant, and tell an interesting story about differences in the costs and benefits of Internet adoption across industries. The coefficients on SIC 60—which includes commercial banks, credit unions, and savings institutions—were strong, negative, and significant for access and e-commerce. Why were banks so slow to adopt, particularly given the promise of on-line banking? An important component of banking information systems is transaction processing, which requires sizeable technical investments to ensure security and reliability. Early Internet technologies could ensure neither security nor reliability, implying high co-invention costs.

Within industries, the value to adoption often differed greatly across technologies. SIC 87 includes engineering firms and management consultants that were quick to adopt access. However, the firms in these industries, who typically offer customized services to a small number of clients, had relatively little need to adopt e-commerce technologies that excelled at distributing standardized products to a large number of customers.

Table 4 indicates that the variables capturing industry and county adoption rates generally have a positive and statistically significant effect on access and e-commerce adoption. However, as mentioned above, these variables may not only be capturing network effects or the effects of

[53] If rural establishments are more likely to be from multi-establishment firms in the sample, the variable *URBAN* may be capturing multi-establishment effects. Column 4 of table 9 shows the effects of *URBAN* when re-estimating the model on single-establishment organizations. *URBAN* continues to be strong and significant. To conserve space, I have not included the results of the models including additional SIC and metropolitan area controls. The results are again very similar, and are available upon request. The primary difference in these models is that the effect of *URBAN* is diminished in the models with additional metropolitan area controls. In these models, all metropolitan area controls that are significant have negative coefficients.

competition, but may also be proxying for unobserved heterogeneity in net benefits to adoption across industries and counties.

7. Assessment

• What factors separated adopters of the Internet from non-adopters? The organizations most likely to adopt the Internet were those that benefited most from the decreases in communications costs made possible by early Internet technologies. Organizations that had made heavy investments in information and communication technologies in the past were most likely to adopt. Organizations that were geographically concentrated benefited less from basic communication technologies, and often chose not to adopt. Urban organizations were less likely to adopt technologies such as access and e-commerce that permitted easier communication with suppliers and customers.

Organizations usually adopted access and intranet together, however by 1998 many organizations still had not yet adopted e-commerce. Why would organizations adopt access and intranet but not e-commerce? One reason may be that the technical and organizational costs of adopting access and intranet were lower than those of e-commerce. Most of the technical problems associated with the adoption of access and intranet had already been solved by scientists and engineers. Moreover, because early uses of access and intranet primarily involved activities such as sending email or viewing web pages, the co-invention costs of these technologies were low.

The costs of adopting e-commerce were much higher that those of access or intranet. There are several reasons for this. First, many e-commerce applications were simply more complicated than other Internet technologies. Second, many of the complementary technologies necessary to conducting commerce had yet to appear. The World Wide Web's core technologies—the TCP/IP communication protocol, the hypertext markup language (HTML), and the uniform resource locator (URL) address system—were ill-equipped to deal with security issues or to facilitate dynamic communication between database servers and web pages.[54]

These results suggest a common theme in the pattern of Internet adoption. Organizations first demanded access and intranet technologies that were relatively inexpensive to adopt. These early technologies lowered communication costs through applications such as e-mail and the World Wide Web, and were adopted first by organizations with the most to gain from reductions in communications costs. These technologies diffused rapidly. However, there remained many organizations for which these early applications had little value; these organizations responded by delaying adoption.

• **Were sophisticated users the ones most likely to adopt?** This study shows that increases in technical sophistication do not always translate into an increased likelihood of adopting new innovations. Theories that state recent investments in information technology increase the probability of adopting new innovations (e.g., March and Sproull 1990; Swanson 1994) are incomplete. Investments in certain types of LAN operating systems or network management software actually decreased the likelihood of adoption for some organizations. Moreover, the costs of platform-specific software were greatest among those organizations with the most complex network infrastructures.

I showed that switching costs could slow adoption in an environment without network externalities (Kauffman, McAndrews, and Wang 2000) and without close buyer-vendor interaction (Bresnahan and Greenstein 1997). Prior work had not shown how recent technical investments could delay adoption in such an environment. This is likely because it is difficult to

[54] For a detailed description of some of the problems of adapting Internet technologies to commerce

identify between the effects of technical sophistication and the effects of lock-in. This paper provided a unique way of disentangling these two effects.

• **How does geography influence adoption?** I focused on the relationship between geographic dispersion and adoption because of the early use of the Internet as a basic communications technology. Relatively few studies have examined how the geographic dispersion of the organization influences the adoption of new innovations. One notable exception is the literature on the diffusion of the MDF (Palmer et. al. 1987; Palmer, Jennings, and Zhou 1993). Both that literature and this paper found that increases in geographic dispersion lead to the adoption of new innovations. The reason in both sets of studies was the same; the adoption of the new innovation lead to a reduction in communication and coordination costs. This suggests that new innovations that reduce communications costs could be a complement or substitute to changes in organizational forms that are designed to reduce coordination costs.[55]

8. Conclusion

In this study I examined reasons for the "digital divide" between adopters and non-adopters of the Internet. I suggested several hypotheses that identified organization features that potentially affected the costs and benefits of adoption. I then examined whether the empirical evidence was consistent with each hypotheses using a set of proxies.

The rapid diffusion of Internet technologies that characterized the late 1990s was a diffusion of technologies like access and intranet. The early "digital divide" among firms was due primarily to variation in the relative benefits of adopting a technology that lowered communications costs. Technologies that involved commerce or dynamic interaction between

---

activities, see Orfali, Harkey, and Edwards (1999).

[55] For further discussion of the relationship between changes in organizational structure and technological change in IT, see Malone, Yates and Benjamin (1987), Gurbaxani and Whang (1991), and George and King (1991).

the organization and external parties took much longer to diffuse. Their diffusion continues today.

This study raises many new questions for future research. First, future work should more explicitly examine the roles of geographic and competitive factors on Internet adoption. In particular, what role did new web-based entrants play on organizational decisions to adopt access, intranet, and e-commerce? Further work should also examine intra-organization diffusion of Internet technologies and more carefully study the evolution of usage within organizations. Future studies should try to quantify the benefits of Internet adoption, examining the revenue-enhancing and cost-reducing effects of these technologies. Last, as noted above, this study has examined only the early stages of diffusion for a set of evolving technologies. More work should be done examining the diffusion of more complicated—and costly—Internet technologies. In all, this paper has taken a first look at understanding the forces driving investment in early Internet technologies. Much work remains to be done in understanding how these technologies have been implemented, and the impact of their usage.

References

Attewell, Paul. 1992. Technology Diffusion and Organizational Learning: The Case of Business Computing. *Organizational Science* 3 (February): 1-19.

Bakos, Yannis. 1997. Reducing Buyer Search Costs: Implications for Electronic Marketplaces. *Management Science.* 43 (December): 1676-1692.

Bakos, Yannis and Erik Brynjolfsson. 1999. Bundling Information Goods: Pricing, Profits, and Efficiency. *Management Science.* 45 (December): 1613-1630.

Bailey, Joseph. 2000. High Speed Internet Access Diffusion: An Analysis of Internet Service Providers. Working Paper, R.H. Smith School of Business, University of Maryland, College Park.

Bernard, Ryan. 1998. *The Corporate Intranet*. New York: John Wiley & Sons.

Bertschek, Irene and Helmut Fryges. 2002. The Adoption of Business-to-Business E-Commerce: Empirical Evidence for German Companies. Discussion Paper 02-05, Centre for European Economic Research, Mannheim.

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.

Bresnahan, Timothy and Shane Greenstein. 1997. Technical Progress and Co-invention in Computing and the Uses of Computers. *Brookings Papers on Economic Activity: Microeconomics*: 1-83.

Bresnahan, Timothy and Manuel Trajtenberg. 1995. General Purpose Technologies: 'Engines of growth'? *Journal of Econometrics* 65: 83-108.

Brown, Jeffrey and Austan Goolsbe. 2001. Does the Internet Make Markets More Competitive? Evidence from the Life Insurance Industry. Working Paper, Graduate School of Business Administration, University of Chicago, Chicago.

Brynjolfsson, Erik and Michael D. Smith. 2000. Frictionless Commerce? A Comparison of Internet and Conventional Retailers. *Management Science* 46 (April): 563-585.

Brynjolfsson, Erik and Michael D. Smith. 2000. The Great Equalizer? Consumer Choice Behavior at Internet Shopbots" Working Paper, Sloan School of Management, MIT, Boston.

Carlton, Dennis and Judith Chevalier. 2001. Free Riding and Sales Strategies for the Internet. Working Paper 8067, National Bureau of Economic Research, Boston.

Chandler, Alfred. 1962. *Strategy and Structure: Chapters in the History of the American Industrial Enterprise.* Cambridge: MIT Press.

Chen, Pei-Yu and Lorin Hitt. 2001. Switching Cost and Brand Loyalty in Electronic Markets: Evidence from On-Line Retail Brokers. Working Paper, The Wharton School, University of Pennsylvania, Philadelphia.

Chircu, Alina and Robert Kauffman. 2000. Limits to Value in Electronic Commerce-Related IT Investments. *Journal of Management Information Systems.* 17 (Fall): 59-80.

Choi, Soon-Yong and Andrew Whinston. 2000. *The Internet Economy: Technology and Practice*. Austin: SmartEcon Publishing.

Chwelos, Paul, Izak Benbasat, and Albert Dexter. 2001. Research Report: Empirical Test of an EDI Adoption Model. *Information Systems Research* 12 (September): 304-321.

Clemons, Eric, Il-Horn Hann, and Lorin Hitt. 1999. The Nature of Competition in Electronic Markets: An Empirical Investigation of Online Travel Agent Offerings. Working Paper, The Wharton School, University of Pennsylvania, Philadelphia.

Cortese, Amy. 1996. Here Comes the Intranet. *Business Week*. February 26.

Dewan, Rajiv, Bing Jing, and Abraham Seidmann. 2000. Adoption of Internet-Based Product Customization and Pricing Strategies. *Journal of Management Information Systems* 17 (Fall): 9-28.

Downes, Tom and Shane Greenstein. 1999. Do Commercial ISPs Provide Universal Access? In *Competition, Regulation and Convergence: Current Trends in Telecommunications Policy Research*, ed. Sharon Gillett and Ingo Vogelsang, 195-212, Mahwah: Lawrence Erlbaum Associates.

Farrell, Joseph and Garth Saloner. 1985. Standardization, Compatibility, and Innovation. *RAND Journal of Economics* 16 (Spring): 70-83.

Fichman, Robert. 1992. Information Technology Diffusion: A Review of Empirical Research. In *Proceedings of the 13th International Conference on Information Systems held in Dallas December 1992*, ed. Janice DeGross, Jack Becker, and Joyce Elam, 195-206. New York: ACM Press.

George, Joey and John King. 1991. Examining the Computing and Centralization Debate. *Communications of the ACM* 34 (July): 63-72.

Gertner, Robert and Robert Stillman. 2000. Vertical Integration and Internet Strategies in the Apparel Industry. Working Paper, Graduate School of Business, University of Chicago, Chicago.

Goolsbee, Austan and Peter Klenow. 2000. Evidence on Learning and Network Externalities in the Diffusion of Home Computers. Working Paper, Graduate School of Business, University of Chicago, Chicago.

Griliches, Zvi. 1957. Hybrid Corn: An Exploration in the Economics of Technical Change. *Econometrica* 25 (October): 501-522.

Gurbaxani, Vijay and Seungjin Whang. 1991. The Impact of Information Systems on Organizations and Markets. *Communications of the ACM* 34 (January): 59-73.

Hannan, Timothy and John McDowell. 1984. The determinants of technology adoption: the case of the banking firm. *RAND Journal of Economics* 15 (Autumn): 328-335.

Hitt, Lorin and Erik Brynjolfsson. 1997. Information Technology and Internal Firm Organization: An Exploratory Analysis *Journal of Management Information Systems* 14 (Summer): 1-101.

Irwin, Douglas and Peter Klenow. 1994. Learning By Doing Spillovers in the Semiconductor Industry. *Journal of Political Economy* 102 (December): 1200-1227.

Ito, Harumi. 2000. The Structure of Adjustment Costs in Mainframe Computer Investment. Working Paper, Economics Department, Brown University, Providence.

Johnson, Eric, Wendy Moe, Peter Fader, Steven Bellman, Jerry Lohse. 2000. On the Depth and Dynamics of Online Search Behavior. Working Paper, Wharton School of Business, University of Pennsylvania, Philadelphia.

Kalakota, Ravi and Marci Robinson. 2001. *e-Business 2.0: Roadmap for Success.* Reading: Addison-Wesley.

Karshenas, Massoud and Paul Stoneman. Rank, Stock, Order, and Epidemic Effects in the Diffusion of New Process Technologies: An Empirical Model. *RAND Journal of Economics* 24 (Winter): 503-528.

Katz, Michael and Carl Shapiro. 1986. Technology Adoption in the Presence of Network Externalities. *Journal of Political Economy* 94 (August): 822-841.

Kauffman, Robert, James McAndrews, and Yu-Ming Wang. 2000. Opening the "Black Box" of Network Externalities in Network Adoption. *Information Systems Research* 11 (March): 61-82.

Kauffman, Robert and Eric Walden. 2001. Economics and Electronic Commerce: Survey and Directions for Research. *International Journal of Electronic Commerce*. 5 (Summer): 5-116.

Keane, Michael. 1992. A Note on Identification in the Multinomial Probit Model. *Journal of Business and Economic Statistics* 10 (April): 193-200.

Kimberly, John and Michael Evanisko. 1981. Organizational Innovation: The Influence of Individual, Organizational, and Contextual Factors on Hospital Adoption of Technological and Administrative Innovations. *Academy of Management Journal* 24 (December): 689-713.

Klemperer, Paul. 1995. Competition When Consumers Have Switching Costs: An Overview with Applications to Industrial Organization, Macroeconomics, and International Trade. *Review of Economic Studies* 62 (October): 515-539.

Maddala, G.S. 1983. *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge: Cambridge University Press.

Malone, Thomas, Joanne Yates, and Robert Benjamin. 1987. Electronic Markets and Electronic Hierarchies. *Communications of the ACM* 30 (June): 484-497.

March, James and Lee Sproull. 1990. Technology, Management, and Competitive Advantage. In *Technology and Organizations*, ed. Paul Goodman and Lee Sproull, 144-173. San Francisco: Jossey-Bass.

McFadden, Daniel. 1978. Modeling the Choice of Residential Location. In *Spatial Interaction Theory and Residential Location*, ed. Anders Karlgvist et. al., 75-96. Amsterdam: North Holland.

McFadden, Daniel. 1981. Econometric Models of Probabilistic Choice. In *Structural Analysis of Discrete Data with Econometric Applications*, ed. Charles Manski and Daniel McFadden, 198-272. Cambridge: MIT Press.

Moch, Michael and Edward Morse. 1977. Size, Centralization, and Organizational Adoption of Innovations. *American Sociological Review* 42 (October): 716-725.

Moe, Wendy and Peter Fader. 2000. Capturing Evolving Visit Behavior in Clickstream Data. Working Paper, Wharton School of Business, University of Pennsylvania, Philadelphia.

Mokyr, Joel. 1992. *The Lever of Riches: Technological Creativity and Economic Progress.* Oxford: Oxford University Press.

Orfali, Robert, Dan Harkley, and Jeri Edwards. 1999. *The Client/Server Survival Guide*, New York: John Wiley and Sons.

Palmer, Donald, Roger Friedland, P. Devereaux Jennings, and Melanie Powers. 1987. The Economics and Politics of Structure: The Multidivisional Form and the Large U.S. Corporation. *Administrative Science Quarterly* 32 (March): 25-48.

Palmer, Donald, P. Devereaux Jennings, and Xeuguang Zhuo. 1993. Late Adoption of the Multidivisional Form by Large U.S. Corporations: Institutional, Political, and Economic Accounts. *Administrative Science Quarterly* 38 (March): 100-131.

Rogers, Everett M. 1995. *Diffusion of Innovations.* New York: Free Press.

Scott Morton, Fiona, Florian Zettelmeyer, and Jorge Risso. 2000. Internet Car Retailing. Working Paper 8067, National Bureau of Economic Research, Boston.

Subramani, Mani and Eric Walden. 2001. The Impact of E-Commerce Announcements on the Market Value of Firms. *Information Systems Research*. 12 (June): 135-154.

Swanson, E. Burton. 1994. Information Systems Innovation Among Organizations. *Management Science* 40 (September): 1069-1092.

Tan, Margaret and Thompson S.H. Teo. 1998. Factors Influencing the Adoption of the Internet. *International Journal of Electronic Commerce*. 2 (Spring): 5-18.

Varian, Hal, Robert Litan, Andrew Elder, and Jay Shutter. 2002. The Net Impact Study: The Projected Economic Benefits of the Internet in the United States, United Kingdom, France, and Germany [article on-line]; available from http://www.netimpactstudy.com/NetImpact_Study_Report.pdf.

Wallsten, Scott. 1999. Geographic Clustering and Spillovers: An Empirical Approach Using Geographic Information Systems and Firm-Level Data. Working Paper, Stanford Institute for Economic Policy Research, Stanford University, Palo Alto.

Whinston, Andrew, Anitesh Barua, Jay Shutter, and Brant Wilson. Measuring the Internet Economy [article on-line]; available from http://www.internetindicators.com/jan_2001.pdf.

Williamson, Oliver. 1975. *Markets and Hierarchies: Analysis and Antitrust Implications*. New York: The Free Press.

Table 1
Percentage of Organizations Adopting Internet Technologies

	Access	Intranet	E-commerce
1996	33.3	19.4	7.1
1997	65.5	39.8	20.0
1998	77.9	59.0	28.9

Sources: (1) Harte Hanks Market Intelligence. 2000. CI Technology Database [computer file]. (2) Author's calculations.

Notes: (1) Sample only includes observations in the CI Technology Database in 1996 and reported software. (2) Number of observations: 5239 in 1996; 4339 in 1997; 3704 in 1998.

Table 2
Percentage of Organizations Adopting
Internet Technology Combinations

	None	Access Only	Access & Intranet	Access & Commerce	Access, Intranet, & Commerce
1996	66.7	11.9	14.6	2.5	4.3
1997	34.5	19.6	26.2	6.6	13.0
1998	22.1	14.8	34.4	5.5	23.2

Sources: (1) Harte Hanks Market Intelligence. 2000. CI Technology Database [computer file]. (2) Author's calculations.

Notes: (1) Sample only includes observations in the CI Technology Database in 1996 and reported software. (2) Number of observations: 5239 in 1996; 4339 in 1997; 3704 in 1998.

Table 3
Description of Variables

Variable	Description	Mean	Std. Deviation	Minimum	Maximum
		Lead User			
PCPEREMP	Total organization PCs divided by total organization employees	0.686	0.534	0	6.118
CLIENT	Percentage of establishments with Internet-ready clients	0.924	0.233	0	1.000
NOAPP	Dummy indicating no applications in organization	0.101	0.301	0	1.000
PCTLAN	Percentage of customized (non-PC) applications accessed over LAN	0.351	0.256	0	1.000
SYSCOM	Dummy indicating presence of system LAN applications	0.327	0.469	0	1.000
NETWARE	Dummy indicating use of Netware LAN OS	0.616	0.486	0	1.000
INTRANETWARE	Dummy indicating use of Intranetware LAN OS	0.053	0.223	0	1.000
MSNT	Dummy indicating use of Microsoft NT LAN OS	0.374	0.484	0	1.000
PCTMAIN	Percentage of customized (non-PC) applications accessed over mainframe	0.168	0.257	0	1.000
		Competing Effects of Installed Base			
NETWARE × PCTLAN	NETWARE interacted with PCTLAN	0.263	0.267	0	1.000
SYSCOM × PCTLAN	SYSCOM interacted with PCTLAN	0.151	0.241	0	1.000
INTRANETWARE× PCTLAN	INTRANETWARE interacted with PCTLAN	0.023	0.105	0	1.000
MSNT ×PCTLAN	MSNT interacted with PCTLAN	0.155	0.236	0	1.000

Table 3 Continued

Variable	Description	Mean	Std. Deviation	Minimum	Maximum
		Size			
EMP100-200	Log of number of employees (between 100 and 200)	0.166	0.237	0	0.693
EMP201-500	Log of number of employees (between 201 and 500)	0.319	0.534	0	1.609
EMP500+	Log of number of employees (greater than 500)	0.586	1.206	0	6.914
TOTMIPS	Log of total MIPS from non-PC computing hardware	2.132	2.352	0	11.381
		Geographic Location			
MULTEST	Dummy indicating multiple establishments in the organization	0.137	0.344	0	1.000
MULTEST × EMPCON	MULTEST interacted with employee Herfindahl	0.057	0.164	0	0.936
MULTEST × PCT100	MULTEST interacted with percentage of pairwise establishment combinations within 100 miles in distance	0.035	0.160	0	1.000
MULTEST × PCT500	MULTEST interacted with percentage of pairwise establishment combinations within 100 miles in distance	0.069	0.227	0	1.000
MULTEST × AVGDIST	MULTEST interacted with average distance between pairwise establishment combinations	0.777	2.134	-9.739	7.863
URBAN	Percentage of establishments in an MSA	0.940	0.229	0	1.000
		External Environment			
DCNTYZERO	Percentage of establishments in which organization has only establishment in county	0.037	0.189	0	1.000

Table 3 Continued

Variable	Description	Mean	Std. Deviation	Minimum	Maximum
PCNTYACC	Percentage of other organizations in county that have adopted access, averaged across organization counties	0.433	0.161	0	1.000
PCIACC	Percentage of other organizations in 3-digit SIC industry that have adopted access, averaged across organization industries	0.551	0.135	0.083	1.000
PCNTYINT	Percentage of other organizations in county that have adopted intranet, averaged across organization counties	0.241	0.142	0	1.000
PCIINT	Percentage of other organizations in 3-digit SIC industry that have adopted intranet, averaged across organization industries	0.354	0.108	0	0.655
PCNTYCOM	Percentage of other organizations in county that have adopted e-commerce, averaged across organization counties	0.090	0.078	0	1.000
PCICOM	Percentage of other organizations in 3-digit SIC industry that have adopted e-commerce, averaged across organization industries	0.175	0.078	0	0.530
SIC60	Percentage of establishments in SIC60	0.081	0.269	0	1.000
SIC61	Percentage of establishments in SIC61	0.024	0.147	0	1.000
SIC62	Percentage of establishments in SIC62	0.032	0.172	0	1.000
SIC63	Percentage of establishments in SIC63	0.091	0.282	0	1.000
SIC64	Percentage of establishments in SIC64	0.033	0.173	0	1.000
SIC65	Percentage of establishments in SIC65	0.032	0.176	0	1.000
SIC67	Percentage of establishments in SIC67	0.031	0.162	0	1.000
SIC73	Percentage of establishments in SIC73	0.290	0.448	0	1.000
SIC87	Percentage of establishments in SIC87	0.183	0.383	0	1.000

Note: Number of observations is 6156.

Table 4
Baseline Estimates from Two-level Nested Logit Model

	Access	E-Commerce
	Lead User	
PCPEREMP	0.4072**	0.1925**
	(0.0865)	(0.0707)
CLIENT	1.2487**	0.3427
	(0.1644)	(0.2316)
NOAPP	-0.6717**	-0.3338
	(0.1491)	(0.2005)
PCTLAN	1.3203**	0.3409
	(0.2670)	(0.2533)
SYSCOM	0.3372*	0.1346
	(0.1970)	(0.1721)
NETWARE	0.5133**	0.2099
	(0.1334)	(0.1332)
PCTMAIN	-0.1537	-0.1707
	(0.1454)	(0.1586)
	Competing Effects of Installed Base	
NETWARE × PCTLAN	-1.1237**	-0.2937
	(0.3085)	(0.2975)
SYSCOM × PCTLAN	-0.7440*	-0.2042
	(0.3861)	(0.3485)
	Size	
EMP100-200	-0.0307	0.1374
	(0.1791)	(0.1916)
EMP201-500	0.1260	0.4700**
	(0.1154)	(0.0820)
EMP500+	0.2275**	0.4006**
	(0.1036)	(0.0496)
TOTMIPS	-0.0332	0.0948**
	(0.0287)	(0.0176)
	Geographic Location	
MULTEST	2.2427	1.2554**
	(0.6353)	(0.2442)
MULTEST × EMPCON	-2.2057**	-1.8622**
	(1.0253)	(0.4290)
MULTEST × PCT100	-0.5388*	-0.0694
	(0.2999)	(0.2312)
URBAN	-0.2916	-0.3363**
	(0.1776)	(0.1482)

Table 4 Continued

	Access	E-Commerce
	External Environment	
DCNTYZERO	0.2451	-0.0253
	(0.2417)	(0.1937)
PCNTYACC	0.4070*	. . .
	(0.2312)	. . .
PCIACC	0.5594*	. . .
	(0.3228)	. . .
PCNTYCOM	. . .	1.1005**
	. . .	(0.4020)
PCICOM	. . .	0.8014
	. . .	(0.5465)
SIC60	-0.9781**	-0.4227**
	(0.1651)	(0.1633)
SIC61	-0.3151	-0.0736
	(0.2500)	(0.2263)
SIC62	-0.6686**	-0.4618**
	(0.2184)	(0.2073)
SIC63	-0.4950**	-0.4560**
	(0.1658)	(0.1362)
SIC64	-0.3448	-0.2327
	(0.2152)	(0.2003)
SIC65	-0.5788**	-0.7184**
	(0.2252)	(0.2543)
SIC67	-0.9097**	-1.0760**
	(0.2727)	(0.2463)
SIC73	-0.3232**	-0.1406
	(0.1105)	(0.1000)
SIC87	-0.2256	-0.3898**
	(0.1397)	(0.1124)
Inclusive Value	0.2128	
	(0.6360)	
N	6156	
Log-likelihood	-5605.7506	

Source: Author's calculations.

Notes: *Indicates significance at 10% level. **Indicates significance at 5% level. Standard errors in parentheses.

Table 5
Marginal Effects in Baseline Model

	Marginal Effect of Change in RHS Variable on Probability of Adoption		
	Access	Commerce	Commerce (conditional on access)
Lead User			
PCPEREMP	0.065	0.046	0.036
CLIENT	0.236	0.105	0.062
NOAPP	-0.118	-0.077	0.044
PCTLAN	0.184	0.105	0.066
SYSCOM	0.050	0.035	0.026
NETWARE	0.081	0.054	0.040
PCTMAIN	-0.025	-0.032	-0.032
Competing Effects of Installed Base			
NETWARE × PCTLAN	-0.182	-0.090	-0.055
SYSCOM × PCTLAN	-0.124	-0.062	-0.038
Size			
EMP100-200[a]	-0.001	0.004	0.005
EMP201-500[a]	0.005	0.018	0.020
EMP500+[a]	0.033	0.073	0.077
TOTMIPS[a]	-0.010	0.032	0.044
Geographic Location			
MULTEST	0.217	0.316	0.272
MULTEST ×EMPCON	-0.429	-0.234	-0.242
MULTEST × PCT100	-0.089	-0.034	-0.013
URBAN	-0.043	-0.068	-0.067
External Environment			
PCNTYACC	0.059	0.016	. . .
PCIACC	0.084	0.023	. . .
PCNTYADV	0.012	0.189	0.231
PCIADV	0.008	0.133	0.164

Source: Author's calculations.

Notes: Table presents marginal effect of change in variable on probability of outcome in column. Marginal effects are calculated by changing variables from 0 to 1, unless otherwise noted. Effects are calculated for each organization, then averaged across organizations in the sample.

[a] Denotes marginal effects calculated by increasing variable by one standard deviation.

Table 6
Nested Logit Estimates,
Lead User and Competing Effects of Installed Base Theories

	Baseline	Including Intranetware	Including Microsoft NT	3-digit SIC Codes	3-digit SIC & Top 20 Metro	Single Est. Only
			Access			
			Lead User			
PCPEREMP	0.4072**	0.4106**	0.3458**	0.4131**	0.4042**	0.5516**
	(0.0865)	(0.0866)	(0.0851)	(0.0908)	(0.0936)	(0.1175)
CLIENT	1.2487**	1.2685**	1.2274**	1.1911**	1.2104**	1.4376**
	(0.1644)	(0.1651)	(0.1628)	(0.1662)	(0.1694)	(0.2227)
NOAPP	-0.6717**	-0.6791**	-0.6242**	-0.6006**	-0.6217**	-0.6847**
	(0.1491)	(0.1487)	(0.1483)	(0.1544)	(0.1573)	(0.1737)
PCTLAN	1.3203**	1.3371**	1.0116**	1.3083**	1.2926**	1.4988**
	(0.2670)	(0.2669)	(0.2849)	(0.2770)	(0.2823)	(0.3094)
SYSCOM	0.3372*	0.3439*	0.2445	0.3283	0.3285	0.4755**
	(0.1970)	(0.1960)	(0.1970)	(0.2033)	(0.2055)	(0.2371)
NETWARE	0.5133**	0.5354**	0.5111**	0.4619**	0.4655**	0.6676**
	(0.1334)	(0.1332)	(0.1330)	(0.1372)	(0.1395)	(0.1672)
INTRANETWARE	. . .	-0.3886	. . .	. . .	. . .	. . .
	. . .	(0.4834)	. . .	. . .	. . .	. . .
MSNT	. . .	. . .	0.4564**	. . .	. . .	. . .
	. . .	. . .	(0.1734)	. . .	. . .	. . .
PCTMAIN	-0.1537	-0.1581	-0.1144	-0.1262	-0.1320	-0.2636
	(0.1454)	(0.1453)	(0.1449)	(0.1505)	(0.1530)	(0.1699)
			Competing Effects of Installed Base			
NETWARE ×PCTLAN	-1.1237**	-1.2078**	-0.8025**	-1.0636**	-1.0657**	-1.3256**
	(0.3085)	(0.3090)	(0.3099)	(0.3179)	(0.3230)	(0.3631)
INTRANETWARE ×	. . .	1.8557*	. . .	. . .	. . .	. . .
PCTLAN	. . .	(1.1058)	. . .	. . .	. . .	. . .

Table 6 Continued

	Baseline	Including Intranetware	Including Microsoft NT	3-digit SIC Codes	3-digit SIC & Top 20 Metro	Single Est. Only
MSNT × PCTLAN	. . .	. . .	0.0759	. . .	. . .	. . .
	. . .	. . .	(0.3687)	. . .	. . .	. . .
SYSCOM × PCTLAN	-0.7440*	-0.7581**	-0.6150	-0.7555*	-0.7382*	-0.9901**
	(0.3861)	(0.3847)	(0.3875)	(0.3985)	(0.4028)	(0.4511)
Commerce						
Lead User						
PCPEREMP	0.1925**	0.1953**	0.1754**	0.1985**	0.2110**	0.2194**
	(0.0707)	(0.0711)	(0.0722)	(0.0714)	(0.0725)	(0.0781)
CLIENT	0.3427	0.3419	0.3163	0.2918	0.2948	0.4122
	(0.2316)	(0.2308)	(0.2323)	(0.2320)	(0.2328)	(0.2887)
NOAPP	-0.3338	-0.3324*	-0.2660	-0.2537	-0.2589	-0.2323
	(0.2005)	(0.2008)	(0.2031)	(0.2021)	(0.2028)	(0.2085)
PCTLAN	0.3409	0.3433	0.4472	0.3634	0.3869	0.3439
	(0.2533)	(0.2540)	(0.2816)	(0.2520)	(0.2530)	(0.2730)
SYSCOM	0.1346	0.1178	0.0562	0.1106	0.0978	0.1896
	(0.1721)	(0.1727)	(0.1761)	(0.1721)	(0.1727)	(0.1888)
NETWARE	0.2099	0.1761	0.2020	0.1957	0.2111	0.2597*
	(0.1332)	(0.1352)	(0.1342)	(0.1334)	(0.1351)	(0.1410)
INTRANETWARE	. . .	0.5732	. . .	. . .	. . .	. . .
	. . .	(0.3621)	. . .	. . .	. . .	. . .
MSNT	. . .	. . .	0.3218**	. . .	. . .	. . .
	. . .	. . .	(0.1449)	. . .	. . .	. . .
PCTMAIN	-0.1707	-0.1694	-0.1433	-0.1669	-0.1642	-0.1329
	(0.1586)	(0.1591)	(0.1601)	(0.1585)	(0.1598)	(0.1693)

Table 6 Continued

	Baseline	Including Intranetware	Including Microsoft NT	3-digit SIC Codes	3-digit SIC & Top 20 Metro	Single Est. Only
			Competing Effects of Installed Base			
NETWARE ×PCTLAN	-0.2937	-0.2258	-0.2501	-0.2569	-0.3011	-0.3744
	(0.2975)	(0.3021)	(0.3045)	(0.2964)	(0.2985)	(0.3207)
INTRANETWARE ×	. . .	-1.1681	. . .	. . .	. . .	. . .
PCTLAN	. . .	(0.7670)	. . .	. . .	. . .	. . .
MSNT × PCTLAN	. . .	. . .	-0.5010	. . .	. . .	. . .
	. . .	. . .	(0.3140)	. . .	. . .	. . .
SYSCOM × PCTLAN	-0.2042	-0.1761	-0.0629	-0.1928	-0.1504	-0.2445
	(0.3485)	(0.3497)	(0.3571)	(0.3472)	(0.3489)	(0.3851)
N	6156	6156	6156	6150	6150	5311
Log-likelihood	-5605.7506	-5601.2068	-5586.4154	-5545.0510	-5497.2906	-4965.6004

Source*: Author's calculations.*

Notes: *Indicates significance at 10% level. **Indicates significance at 5% level.
Standard errors in parentheses. 3-digit SIC code estimates calculated excluding SIC codes 601 (Central Reserve Depository Institutions and 654 (Title Abstract Offices).

Table 7
Nested Logit Estimates,
1995 Covariates

	1995 Adoption	1998 Adoption
	Access	
	Lead User	
PCPEREMP	0.1366	0.1019
	(0.1045)	(0.0875)
CLIENT	0.6589**	0.6576**
	(0.1419)	(0.1457)
NOAPP	-0.7288**	-0.3250
	(0.2126)	(0.2089)
PCTLAN	0.4451*	0.6989**
	(0.2561)	(0.3311)
SYSCOM	0.4758**	0.1421
	(0.2147)	(0.2807)
NETWARE	-0.0055	0.1909
	(0.1500)	(0.1778)
PCTMAIN	-0.4094**	-0.1585
	(0.2047)	(0.2040)
	Competing Effects of Installed Base	
NETWARE ×PCTLAN	0.0605	-0.2819
	(0.3165)	(0.3938)
SYSCOM × PCTLAN	-0.8639**	-0.4806
	(0.4078)	(0.5090)
	Commerce	
	Lead User	
PCPEREMP	0.3727**	0.1405**
	(0.0926)	(0.0621)
CLIENT	0.3755	0.2377
	(0.3704)	(0.1824)
NOAPP	-0.4924	-0.1433
	(0.4732)	(0.2468)
PCTLAN	-0.3916	-0.3292
	(0.5346)	(0.3177)
SYSCOM	0.5928	-0.3478
	(0.4057)	(0.2489)
NETWARE	-0.2685	-0.1451
	(0.3019)	(0.1767)
PCTMAIN	-1.3074**	-0.4051**
	(0.3221)	(0.1785)

Table 7 Continued

	1995 Adoption	1998 Adoption
	Commerce	
	Competing Effects of Installed Base	
NETWARE × PCTLAN	0.6581	0.2465
	(0.6443)	(0.3836)
SYSCOM × PCTLAN	-1.4631**	-0.3478
	(0.7360)	(0.2489)
N	5239	3744
Log-likelihood	-3912.3788	-3532.8257

Source: Author's calculations.

Notes: *Indicates significance at 10% level. **Indicates significance at 5% level. Standard errors in parentheses. Second column indicates estimates of 1996 adoption decisions using 1995 covariates; third column indicates estimates of 1998 adoption decisions using 1995 covariates.

Table 8
Three-Level Nested Logit Estimates

	Access	Intranet
	Lead User	
PCPEREMP	0.4123**	-0.0234
	(0.0822)	(0.1210)
CLIENT	1.2119**	0.3449
	(0.2154)	(0.3059)
NOAPP	-0.6977**	0.2015
	(0.1557)	(0.2050)
PCTLAN	1.2430**	0.6532*
	(0.3615)	(0.3562)
SYSCOM	0.2924	0.3481
	(0.2471)	(0.2592)
NETWARE	0.4527**	0.5108*
	(0.2289)	(0.3013)
PCTMAIN	-0.1764	0.1918
	(0.1519)	(0.2917)
	Competing Effects of Installed Base	
NETWARE × PCTLAN	-0.9972**	-1.0579**
	(0.4744)	(0.4041)
SYSCOM × PCTLAN	0.2924	-0.2862
	(0.2471)	(0.4376)
	Geographic Location	
MULTEST	2.0744**	1.2620
	(0.9480)	(1.0855)
MULTEST × EMPCON	-2.0229	-1.3980
	(1.3239)	(1.7656)
MULTEST × PCT100	-0.4090	-0.9930**
	(0.4712)	(0.4396)
URBAN	-0.3047*	0.0903
	(0.1727)	(0.2029)
	Size	
EMP100-200	-0.0532	0.1966
	(0.1971)	(0.2052)
EMP201-500	0.1069	0.1948
	(0.1460)	(0.1656)
EMP500+	0.2023	0.2285*
	(0.1524)	(0.1388)
TOTMIPS	-0.0355	0.0278
	(0.0324)	(0.0264)

Table 8 Continued

	Access	Intranet
	External Environment	
DCNTYZERO	0.1805	0.5011*
	(0.2925)	(0.2940)
PCNTYACC	0.3969*	. . .
	(0.2366)	. . .
PCIACC	0.5667*	. . .
	(0.3198)	. . .
PCNTYINT	. . .	0.1789
	. . .	(0.2707)
PCIINT	. . .	-0.0408
	. . .	(0.5795)
SIC60	-0.9261**	-0.4731**
	(0.2454)	(0.1792)
SIC61	-0.3537	0.3103
	(0.2843)	(0.7266)
SIC62	-0.6759**	0.0452
	(0.2152)	(0.2545)
SIC63	-0.4986**	0.0293
	(0.1569)	(0.2769)
SIC64	-0.3465	0.0114
	(0.2172)	(0.3986)
SIC65	-0.5655**	-0.1500
	(0.2379)	(0.2322)
SIC67	-0.9450**	0.2398
	(0.2204)	(0.3279)
SIC73	-0.3465	0.1958
	(0.1163)	(0.1881)
SIC87	-0.2724**	0.3757*
	(0.1388)	(0.2143)
Inclusive Value, Level 1	0.1858	
	(0.4819)	
Inclusive Value, Level 2	1.0017	
	(1.3903)	
N	6156	
Log-likelihood	-8271.3096	

Source: Author's calculations.

Notes: *Indicates significance at 10% level. **Indicates significance at 5% level. Standard errors in parentheses.

Table 9 Nested Logit Estimates,
Geographic Location Theory

	Baseline, PCT100	Baseline, PCT500	Baseline, AVGDIST	Single Establishment	1996 Adoption
		Access			
MULTEST	2.2427**	2.1376**	1.6331**	. . .	2.7522**
	(0.6353)	(0.6429)	(0.7416)	. . .	(0.3500)
MULTEST × EMPCON	-2.2057**	-2.2592**	-2.1752**	. . .	-2.8975**
	(1.0253)	(1.0220)	(1.0422)	. . .	(0.5669)
MULTEST × PCT100	-0.5388*	. . .	. . .	. . .	-0.7431**
	(0.2999)	. . .	. . .	. . .	(0.2312)
MULTEST × PCT500	. . .	-0.1204	. . .	. . .	. . .
	. . .	(0.2975)	. . .	. . .	. . .
MULTEST × AVGDIST	. . .	. . .	0.0791	. . .	. . .
	. . .	. . .	(0.0563)	. . .	. . .
URBAN	-0.2916	-0.2862	-0.2925*	-0.4330**	-0.2631
	(0.1776)	(0.1779)	(0.1774)	(0.2035)	(0.1671)
		Commerce			
MULTEST	1.2554**	1.2277**	1.5560**	. . .	1.2117**
	(0.2442)	(0.2555)	(0.3765)	. . .	(0.3641)
MULTEST × EMPCON	-1.8622**	-1.8943**	-1.9857**	. . .	-1.7740**
	(0.4290)	(0.4268)	(0.4353)	. . .	(0.6248)
MULTEST × PCT100	-0.0694	. . .	. . .	. . .	-0.3521
	(0.2312)	. . .	. . .	. . .	(0.4335)
MULTEST × PCT500	. . .	0.0518	. . .	. . .	. . .
	. . .	(0.2043)	. . .	. . .	. . .
MULTEST × AVGDIST	. . .	. . .	-0.0472	. . .	. . .
	. . .	. . .	(0.0440)	. . .	. . .
URBAN	-0.3363**	-0.3377**	-0.3349**	-0.2913*	-0.3460
	(0.1482)	(0.1482)	(0.1484)	(0.1522)	(0.3183)
N	6156	6156	6156	5311	5239
Log-likelihood	-5605.7506	-5607.3315	-5605.9859	-4965.6004	-3912.3788

Figure 1
Two-level Nested Logit Tree
For Internet Adoption Choice

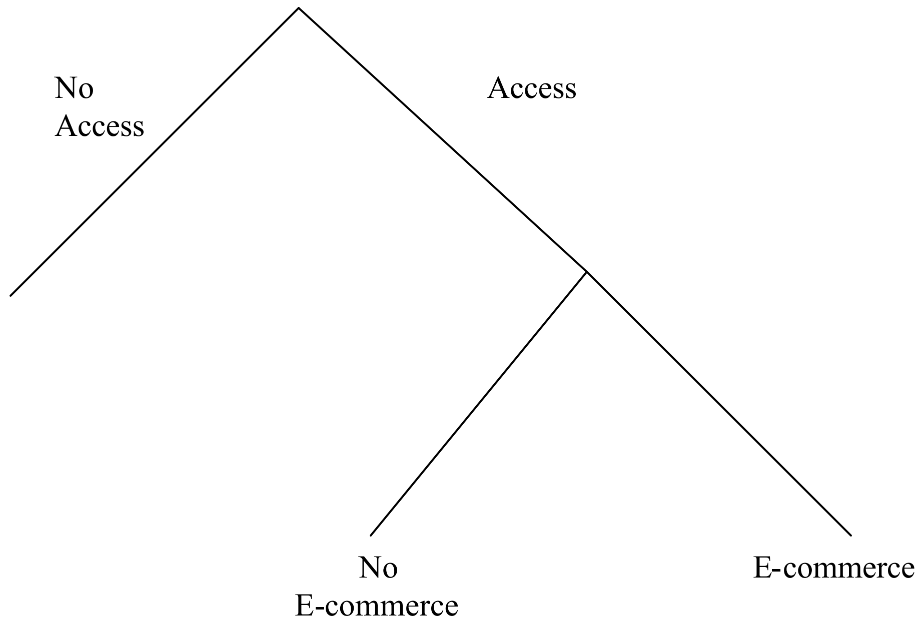


Figure 2
Three-level Nested Logit Tree
For Internet Adoption Choice

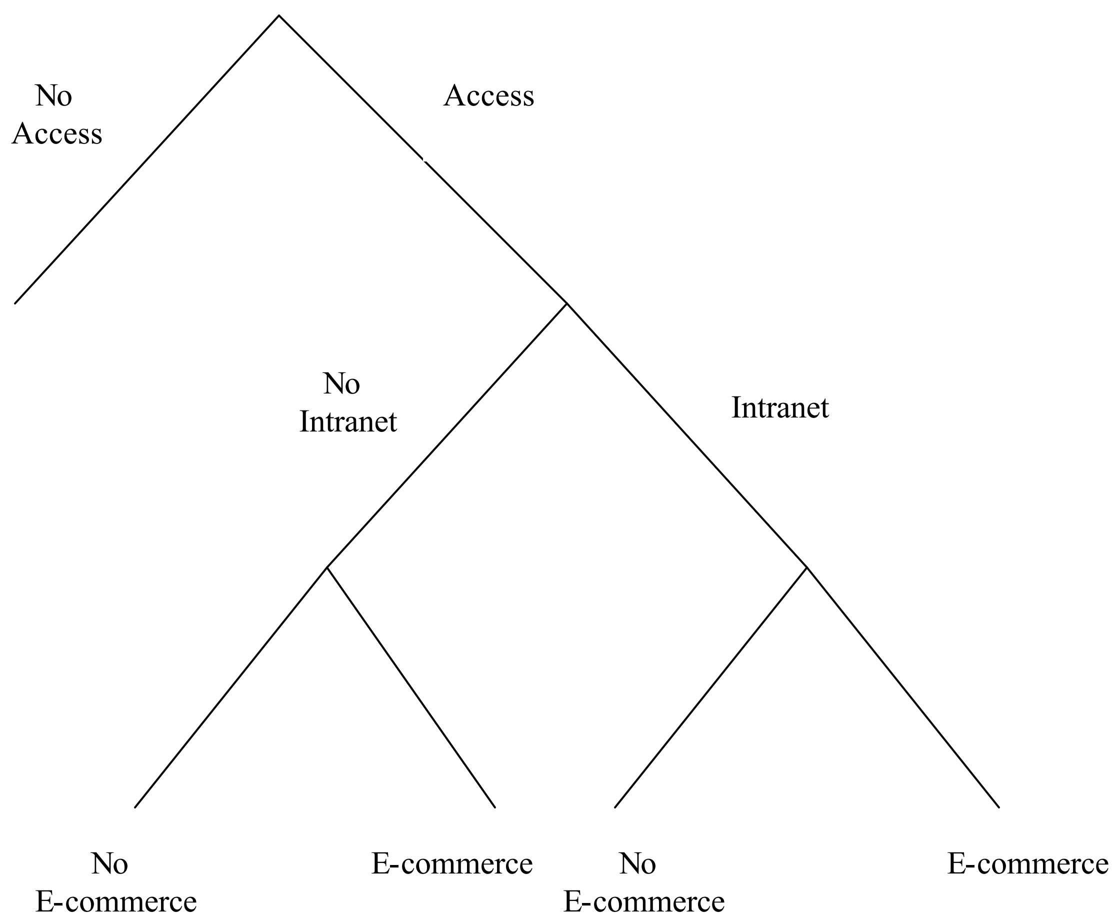
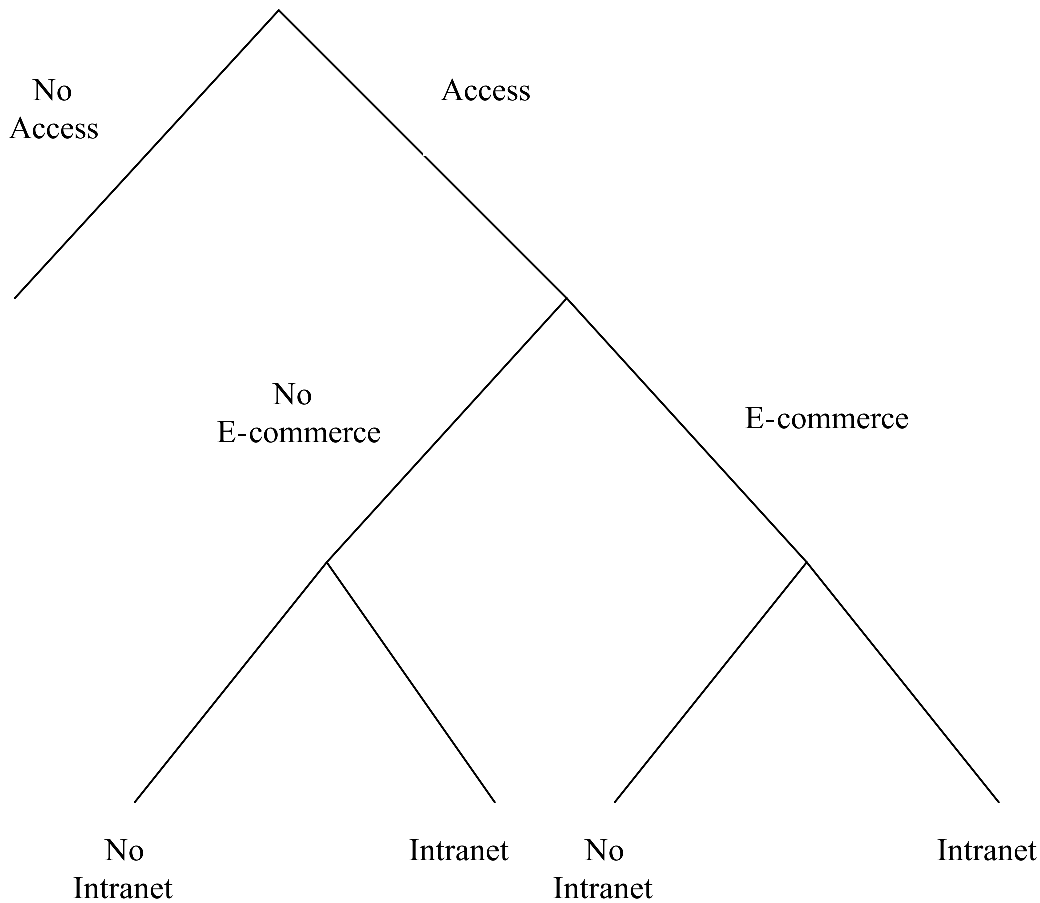


Figure 3
Alternate Three-level Nested Logit Tree
For Internet Adoption Choice



Data Appendix

Unit of Observation

The unit of observation in the CI database is an establishment/year. Roughly speaking, an establishment refers to a particular branch or location of a firm. It is similar to the concept of establishment used by government organizations such as the Bureau of Labor Statistics. Thus, the database will often have data on multiple establishments for a given firm.

The establishment is an inappropriate unit of observation because the technology adoption decision for an establishment likely depends on observable and unobservable attributes of other establishments within the same organization. Table A.1 and Figure A.1 illustrate the potential dangers associated with establishment-level analysis. Table A.1 presents the distribution of number of establishments per organization. The vast majority of organizations have only one or two establishments, although a significant percentage of establishments (over 10%) are part of multi-establishment organizations. To show the manner in which multi-establishment organizations adopt, Figure A.1 shows the distribution of number of establishments adopting Internet technologies among multi-establishment organizations that have adopted some form of access. The distribution of total number of establishments in multi-establishment organizations among this sub-sample is included for comparison purposes. The figure shows that many multi-establishment organizations adopt the Internet at only one location. The mode number of establishments adopting access, intranet, and e-commerce is one for all three technologies. In particular, for e-commerce, over 60% of multi-establishment organizations adopt e-commerce at only one location. Although this graph is of course partially capturing within-firm diffusion, the high percentage of single-establishment adopters suggests that there is some danger that adoption decisions are being made at the firm rather than establishment level, and that the establishment may be an inappropriate unit of observation. To avoid these problems, I conduct all analyses of the paper at the organization level.

I define an organization in a year as the aggregation of all establishments within a firm that have been in the CI database for two consecutive years. The requirement that the establishment be in the database for two consecutive years is required to obtain explanatory variables for the analysis, as prior year variables on organization characteristics are used to determine adoption decisions.

The measure of organization will not, in general, correspond exactly to a firm. This is true for several reasons. First, the organizations in the dataset may consist of only a subset of the industries in a particular firm. An analysis of the organizations in the data set suggests that most are clustered in a small number of SIC codes, however. 94.9% of the organizations in the data set have business activities in only one two-digit SIC code, while another 3.4% conduct business in only two. Second, the CI Technology database does not, in general, sample all the establishments from a firm. Entry and exit of establishments may also change the composition of establishments in an organization across two different years in the sample.

Adoption variables

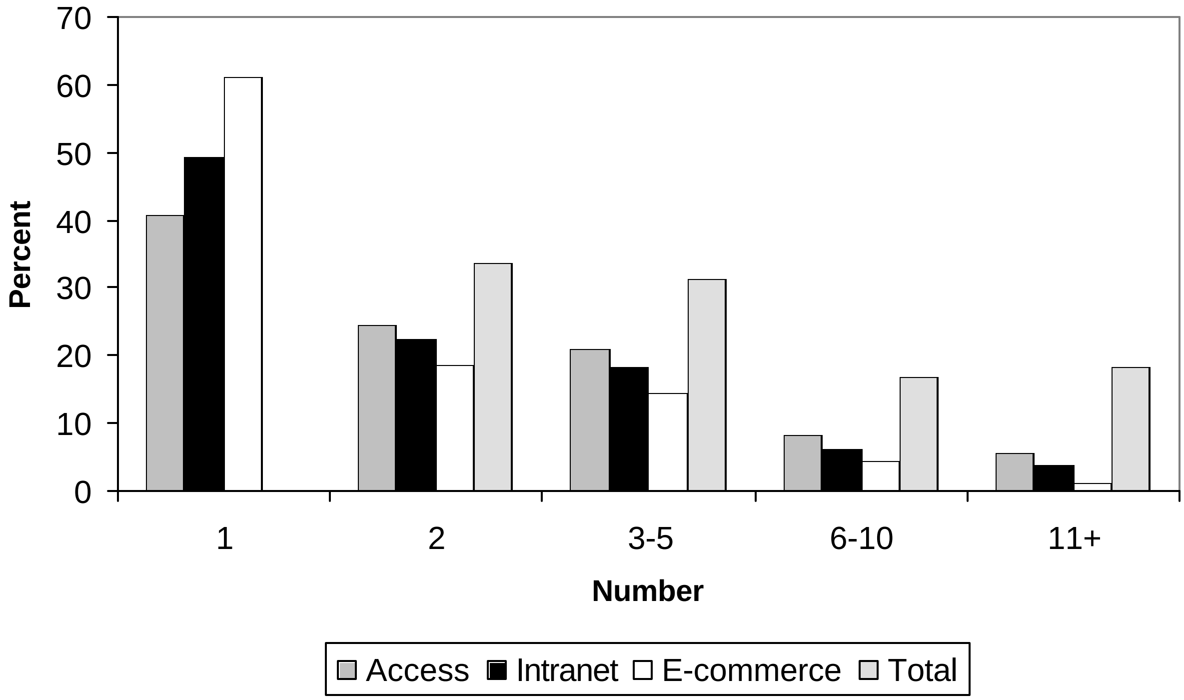
The CI database includes several measures of establishment use of Internet technologies. The first measure of Internet use contains data on an establishment's Internet Service Provider (ISP). Establishments that have responded to this survey by indicating use of an ISP are counted as adopting access. The second measure of Internet use is a direct survey on adoption of Internet technologies. Using this data, establishments are identified as adopting intranet if they responded positively in this survey to adopting either TCP/IP-based e-mail or intranet. An establishment is counted as adopting e-commerce if it responded positively to adopting any of the following: business-to-business electronic commerce, business-to-consumer electronic commerce, electronic commerce, customer service, education, extranet, publishing, purchasing, or technical support.

Table A.1. Establishments per Organization

Number of establishments	Percent
1	88.6
2	4.9
3-5	3.4
6-10	1.6
11-25	1.1
26-50	0.4
51-100	0.1
over 100	0.0

Sources: (1) Harte Hanks Market Intelligence. 2000. CI Technology Database [computer file]. (2) Author's calculations.

Figure A.1. Number of Establishments Adopting Internet in Multi-Establishment Organizations



Sources: (1) Harte Hanks Market Intelligence. 2000. CI Technology Database [computer file]. (2) Author's calculations.

Notes: Sample period is 1996-1998. Based on sample of organizations that adopted basic access. Access, intranet, and e-commerce show distribution of establishments with these technologies among organizations that have adopted the technology. Total indicates the distribution of number of establishments among organizations that have adopted access.