The (Teaching) Role of Universities in the Diffusion of the Internet

Avi Goldfarb*
Rotman School of Management
University of Toronto
105 St. George St.
Toronto, ON M5S 3E6
CANADA
(416) 946-8604
agoldfarb@rotman.utoronto.ca
July 2005

Abstract

This paper provides evidence that students in the mid-1990s learned to use the Internet at university and then continued use after graduation, highlighting a micro-level role for institutions in diffusion. In particular, it shows that university attendance in the mid-1990s had a larger effect on Internet use (especially email) than did university attendance in other periods. It also shows that people who live with students from the mid-1990s are more likely to be Internet users. These effects are largest for low income households. They do not, however, hold for the use of other technologies such as word processing.

(JEL Classification: L86, O33; keywords: Internet, diffusion, universities)

A more polished version is forthcoming in the *International Journal of Industrial Organization*

* Thanks to David Dunne and to Nielsen Canada for providing data; to seminar participants at the University of British Columbia, the University of Toronto, the 2004 IIOC, the 2004 ZEW ICT conference, the 2004 NBER Summer Institute, and the 2005 Internet Use in the Americas conference; and to Ajay Agrawal, Erik Brynjolfsson, Jihui Chen, Maryann Feldman, Chris Forman, Neil Gandal, Shane Greenstein, Ig Horstmann, Mara Lederman, Scott Stern, Joel Waldfogel, Joachim Winter, and two anonymous referees for helpful comments. This research was supported by SSHRC Grant #538-02-1013.

1. Introduction

The role of educational institutions in teaching people about new technologies, while not well researched, is nontrivial. Universities likely played a role in encouraging the use of many technologies including laboratory equipment, software, and databases; however, little empirical work has documented this role or tried to understand how it works. In this paper, I provide evidence that universities taught a generation of students how to use the Internet. These students then brought the technology into their homes after they left school. In addition to playing an important role in the invention of the Internet,[1] universities helped create demand for the technology.

Universities are in a special position for diffusing some technologies. Like many companies, economies of scale mean that they often have cutting edge technologies available. Furthermore, they can force people to use these technologies and can teach skills that are complementary to the technologies. Unlike the situation in many companies, however, there is a considerable flow of people out of universities every year. Together, these factors make universities an ideal conduit for the diffusion of complex technologies. Diffusion through universities may have been part of the reason that the Internet diffused so rapidly through American society.[2]

It is important to understand the causes of this rapid diffusion process. Information technology is often credited with the economic boom of the late 1990s.[3] Understanding the process of its diffusion will help provide insight into the process that led to rapid economic growth in the United States and elsewhere. While it now may seem that the diffusion of the Internet was inevitable, the diffusion process took decades. Invented in the late 1960s, the Internet was used by few people outside of the research community in the early 1990s. Kenney (2001, p. 6) writes, "When the WWW software was first released

[1] Numerous studies and books describe the role of universities in the invention of the network as it exists today (e.g. Berners-Lee, 1999; Kenney, 2001). The Internet browser was invented at the University of Illinois. The University of California at Berkeley was instrumental in connecting the Internet backbone to the west coast of the United States. Many Internet companies have been spun out of Stanford (e.g. Google), MIT (e.g. Akamai), and other schools.

[2] Figure 1 displays the growth of Internet access in the United States. In 1994, the technology was in few homes, but by September 2001 over half of all US households had Internet access.

[3] For example Oliner and Sichel (2000), Baily and Lawrence (2001), and the survey by Jalava and Pohjola (2002).

in 1992, the majority of adopters were in institutions, especially universities." In this paper, I show that universities played a role in creating demand for the technology outside the early adopters in the research community. Network externalities associated with the communications function of the Internet likely meant that the benefits to adoption for students were high: the Internet allowed students to better communicate with their professors. Universities were therefore particularly well-placed to help diffuse the Internet.

While many other papers have documented the role of universities in technology diffusion, this paper argues that it is not just the level of education that leads to adoption, but the attendance (and presumed use of the technology) at school that lead to continued use of the technology well into the future. University attendance during the early years of the Internet is correlated with usage in later years.[4] Using a difference-in-difference framework of education and age, I show that the marginal impact of university education in the mid-1990s on Internet use was much higher than in other periods. This is not true for other computing technologies such as word processing.

This paper is not the first to suggest universities play a role in technology development and diffusion. There is a long literature on the subject of technology spillovers from universities.[5] According to Rogers (1995, p. 358-9), land-grant universities in the United States were in many cases explicitly set up to teach farmers about the latest agricultural techniques. This paper is also not the first to find a link between education levels and technology adoption. It is well documented in the literature that higher education is correlated with adoption (e.g. Hoffman and Novak, 2000). What is new in this paper is the demonstration that university attendance during the early years of the Internet is correlated with usage in later years.

[4] Throughout this paper, a person who continued using the Internet after graduation is called an "adopter". An individual who stops using the Internet becomes a "non-adopter". This is different from the standard language in the literature, but appropriate given the nature of the study. Almost all students would have used the Internet while attending university. What is interesting is adoption on a continuing basis.
[5] Technology spillovers from universities are discussed in many papers (e.g. Acs, Audretsch, and Feldman, 1992; Jaffe, 1989). Monjon and Waelbroeck (2003) discuss the roles of research collaboration between academia and business and of the publication of research results on technology diffusion.

Furthermore, I find that members of households with people who were students in the mid-1990s are more likely to be Internet users. In doing so, I provide a measurement for the magnitude of social learning in a narrow context: learning within the household. Perhaps through direct network externalities or through teaching people the basics of getting online, students may have acted as change agents in facilitating technology diffusion. In this way, educational institutions may have had a role in technology diffusion beyond the direct transmission of knowledge to students.[6] These results highlight a micro-level role for institutions in diffusion.

Most of the results are based on the September 2001 Computer and Internet Use Supplement to the CPS. The data are rich enough to allow modeling the decision to adopt Internet access simultaneously with the decisions to adopt Internet applications such as email, information search, online entertainment, and e-commerce. In this way, I explore the reasons that university attendance led to Internet use. Those who attended university during the mid-1990s and those who live with them are particularly likely to use the Internet for online communication. Combined with the results on other technologies and the high levels of transmission within households, this provides strong evidence that direct network externalities were an important factor in the diffusion of the Internet. People are more likely to adopt if they regularly communicate with earlier adopters.

These results are generally consistent with Goolsbee and Klenow (2002). Using a similar econometric framework, they "find that local spillovers are important for household computer adoption" (p. 340) and they provide strong evidence that these spillovers are a result of computers being part of a communication network. I build on their findings by examining the network in more detail. Furthermore, in contrast to their finding that elementary and high schools are not particular important to personal computer adoption, I find that universities are important for Internet adoption.

[6] There is little evidence in the economics literature on the role of social learning in technology diffusion. One notable exception is Conley and Udry's (2003) study of the role of social learning in the diffusion of the pineapple in Ghana. Their study looks at social learning in the broader context of farm-to-farm. Following Manski (1993), they note that many studies of social learning suffer from an inability to identify actual communication. One way they overcome this problem is through a survey about communication networks. I overcome the problem by narrowing

The next section discusses a number of possible explanations why university education in the mid-1990s might be associated with Internet adoption. Section 3 describes the data and Section 4 presents the basic framework used in the paper to separate the many possible explanations. Section 5 then shows that people who were students in the mid-1990s, and those who live with them, are particularly likely to adopt the Internet, even controlling for age, occupation, industry, and a number of other factors. This does not hold for other computing technologies such as word processing and computer games. Section 6 explores the adoption of particular applications to better understand why university attendance is correlated with adoption. Section 7 concludes.

2. Why might Internet adoption be correlated with university attendance?

A central finding of this paper is that Internet adoption is correlated with university attendance in the mid-1990s. In this section, I will describe a number of possible explanations for this correlation. The first three explanations argue that there was nothing distinctive about university attendance per se. Instead, they suggest that students from the mid-1990s are different from other groups and this unobserved heterogeneity explains the differences in adoption levels. Section 5 shows that these are not complete explanations for the correlation between university attendance in the mid-1990s and Internet adoption. The second four explanations further refine this result by exploring why universities might drive future Internet usage. Each defines a different possible role. All eight explanations are summarized in Table 1.

There are a number of possible sources of unobserved heterogeneity that may drive this correlation. First, there is considerable evidence that younger, educated people are more likely to adopt new technologies in general and the Internet in particular (e.g. Rogers, 1995; Hoffman and Novak, 2000). Second, the occupations available to university graduates during this time period may have been disproportionately Internet-oriented. Therefore, they are more likely to use the Internet because of

---

the scope of investigation to household members. This intra-household diffusion (Stoneman, 2002, p. 7-9) is another topic that has received little empirical exploration.

occupation-related factors independent of university influence. Bresnahan, Brynjolfsson, and Hitt (2002), Doms, Dunne, and Troske (1997), and others have shown that information technology use increases demand for educated workers. If information technology jobs were particularly prevalent in the late 1990s, then this could drive the result that education in this time period leads to technology adoption. Third, it is possible, even controlling for education, age, industry, and occupation, that the group of people being studied are inherently more technology-savvy than others. This idea is suggested by Goolsbee and Klenow (2002, p. 318) as an explanation for local spillovers. In summary, there are three likely sources of unobserved heterogeneity that may cause university graduates from the mid-1990s (and perhaps those who live with them) to be particularly likely to adopt the Internet.

1. People who were students in the mid-1990s, and those who live with them, may use the Internet more because they are younger and more educated than the rest of the population.
2. People who were students in the mid-1990s, and those who live with them, may use the Internet more because they are in occupations and industries that demand Internet skills.
3. People who were students in the mid-1990s, and those who live with them, may use the Internet more because they are particularly technology-savvy.

In section five, I show that these do not provide a complete explanation for the high propensity of people who were educated in the mid-1990s to adopt the Internet. Therefore, in the 1990s universities likely played a role in getting students online. There are a number of possible roles that universities may have played. Universities may have helped the technology diffuse by reducing the costs associated with diffusion. Before its widespread diffusion, the Internet was a difficult technology to sample. Individuals needed access to a computer with a modem as well as the programs that would allow them to access the Internet. Universities may have provided these requirements. They often forced students to use the

Internet; meaning students had no choice but to try the technology.[7] They then might have learned that the technology was not as complex as it may have first appeared. By driving down the cost barriers to adoption, universities may have opened the way for students to use the Internet.

Alternatively, educational institutions could have increased the benefits of Internet adoption by increasing the value of online commerce, online communication, or online information. Universities may have taught students to be more trusting of Internet security and consequently are more willing to purchase products online.[8] Students may then have transferred these attitudes to those who live with them. For online communication, people who attended universities in this time period may have a network of people they know from school who are also online. Since online communication is a technology that displays direct network externalities, this would increase the benefit of the online communication.[9] The people they live with may also be connected to these networks. Universities may also have taught students to be particularly skilled at accessing online information. They may then have transferred these skills to those who live with them. In summary,

> 4. People who were students in the mid-1990s, and those who live with them, may face lower costs to adoption than others.
>
> 5. People who were students in the mid-1990s, and those who live with them, may be more likely to purchase products online.
>
> 6. People who were students in the mid-1990s, and those who live with them, may get a greater net benefit from online communication.
>
> 7. People who were students in the mid-1990s, and those who live with them, may get a greater net benefit from the information available on the Internet.

---

[7] Rogers (1995) calls these the trialability and complexity barriers to adoption. Stoneman (2002, chapter 3) identified possible reasons as risk and uncertainty and learning-by-doing.
[8] The opposite hypothesis is also a possibility: universities may have taught people to distrust Internet security.
[9] Direct network externalities mean that the value of a technology depends directly on the number of other people who have adopted. Forman and Goldfarb (2005) provide a detailed review of direct network externalities in the context of information and communication technologies.

The next two sections describe the data and method used to examine each of these possible explanations.

3. Data

3.1 Data Sets Used

Most of the results are based on the September 2001 Computer and Internet Use Supplement to the CPS. These results are complemented by similar regressions on the December 1998 Computer and Internet Use Supplement, on the October 1994 Computer Use Supplement, and on the *Internet Planner* panel survey conducted by A.C. Nielsen Canada from 1995 to 2000.

The September 2001 CPS supplement contains information about the computer and Internet habits of a representative sample of 142,667 Americans. Most regressions focus on the 104,891 people aged 18 or older in the data. In particular, the survey contains information on demographics, Internet access, Internet application usage, and within household relationships. Unlike the earlier CPS supplements, the September 2001 supplement contains information on the use of non-Internet computer applications including word processing and desktop publishing, computer games, and programs for managing household records. This data allows me to explore whether the results in the study relate to computer use in general or to the Internet in particular. The information on non-Internet computer applications in this supplement makes it especially attractive. The primary weakness of the CPS data for this study is that it is cross-sectional. Consequently, I cannot exactly identify individuals who were students during the mid-1990s. I have to approximate this based on education and age. This means that the coefficient estimates may suffer from errors-in-variables bias.

The *Internet Planner* Survey, conducted by A.C. Nielsen on Canadian Internet usage and repeated annually from 1995 to 2000, does not suffer from this weakness. Due to the panel nature of the study, I can identify individuals who were students in the mid-1990s and then examine their Internet habits in later years. The raw sample consists of 12,100 individuals in Nielsen's Homescan® panel.

Respondents completed the survey using a hand-held computing response system to scan a bar-coded questionnaire. There are 5,519 individuals with data for 2000.

The Nielsen data set, however, has two main weaknesses. The first weakness is the flip side of the main strength. To be identified as a student during the mid-1990s and remain in the data set in 2000, individuals must be in the data set for more than one year. However, to be in the data set in multiple years, individuals cannot have changed residences. Second, they must be the self-identified head of the household. This makes for an unusual population, one that is less mobile and older than the general population. These weaknesses are especially relevant in examining the role of universities as students and young professionals are particularly transient and much less likely to be the head of a household than the general population. This is one of the reasons I focus on the CPS data rather than the Nielsen data. Still, the weaknesses of the CPS and Nielsen data sets are in many ways opposite. Therefore, I conclude that if the same results hold in both then the results are likely independent of the weaknesses.

3.2 Variable Definitions

Internet use is defined in response to straightforward questions such as "Do you use the Internet at any location?" (CPS Sept. 2001 question). All surveys ask whether the respondent used a number of applications over the past year including email, chat, purchase, information (of various kinds), and free entertainment. I divide these into five broad categories: e-communication (email and chat), e-commerce (purchase), e-information (e.g. product information and online news), e-entertainment (e.g. online games), and other (including job search and online banking). Table 2 shows summary statistics.

The main independent variable of interest is whether the individual respondent was a student during the mid-1990s, the time period hypothesized to be key to the diffusion of the Internet through universities. In the Nielsen data, this group of people is easy to identify. *Student 1995-97* is equal to one if the person was a student in 1995, 1996, or 1997. *Student 1995-99* is defined similarly.[10] In the CPS data, however, an approximation is used to define whether an individual was a student in the relevant time period. In particular, this group is defined as those individuals who graduated college or university and were aged 18 to 22 at some point from 1993 to 1997. In other words, all individuals born between 1971 and 1979 who are college or university graduates are in this group.[11] All regressions also include whether the individual has a college or university education and whether the individual was born between 1971 and 1979, creating a difference-in-difference identification of the effect of university attendance. The interaction variable therefore identifies whether the effect of postsecondary education on Internet use was particularly strong for the cohort that went to college or university in the mid-1990s. Table 2 shows that 5.2% of CPS respondents were born between 1971 and 1979 and are college or university graduates.

The CPS data contain household composition information, allowing for the inclusion of a variable for whether respondents live with somebody who was a student in the mid-1990s (and are not such former students themselves).[12] 3.8% of respondents fit in this category. Many regressions also include controls for whether there is someone born from 1971 to 1979 in the household and whether there is someone with a post-secondary education in the household. This data allows me to investigate indirect transmission of the technology from the university to the general public through informal communication networks.

The CPS data also allow controls for race, income group, citizenship, country of origin, homeownership, metropolitan area, state, occupation, industry, employment status, gender, age, marital status, and whether the individual is currently enrolled in postsecondary education (*Current student*).[13] The Nielsen data allow controls for income group, metropolitan area, region, gender, age, number of other Internet users in the sample who live nearby (in the same Forward Sortation Area), and whether the individual is currently enrolled in postsecondary education (*Current student*).

---

[10] Summary statistics for the Nielsen data are available in online appendix Table A.1.
[11] I chose 1993 to 1997 because that is the period between the invention of the browser and the acceleration in the diffusion rate in 1998 when the Internet became a widely used technology (see Figure 1). Furthermore, the December 1998 CPS shows that 18% of 22 year-olds and 27% of 23 year-olds are college graduates while only 29% of 24 year-olds are graduates. Since the surveys took place in December, this suggests that most people graduate by the time they are 23. Combined, these suggest use of the 1971-1979 cohort. If this period is not exactly right then it will introduce errors-in-variables to the results, biasing them to zero if the errors are independent of any measurement errors in the other variables. Without independence the bias is ambiguous. Online Appendix Table A.5 shows that results do hold with minor adjustments to the definition.
[12] The measured spillover effects are larger if this variable includes former students.
[13] Occupations and industries are listed in the online appendix.

Table 3 shows the percentage of the CPS sample adopting the Internet by age cohort and education level. If shows that those born between 1971 and 1979 who attended university are more likely to have adopted than any other cohort. As discussed in Section 2, there are many possible explanations for this. The next section describes the empirical strategy used to separately identify the explanations.

4. Empirical Strategy

4.1 Disentangling the Different Explanations for the Main Effect: Probit Model

People who were students in the mid-1990s are more likely to use the Internet than any other group. In this section, I present the probit model of technology adoption (David, 1969) that I will use to examine whether this result is simply due to unobserved (to the econometrician) heterogeneity or if universities played some sort of role in the diffusion of the Internet.

In the probit model, an individual adopts a technology if the benefits from adoption exceed the costs. Formally, individual *n* will adopt innovation *V* at time *t* if

$$NB_{Vnt}=B_{Vnt}-C_{Vnt}>0, \tag{1}$$

Where NB_{Vnt} is the net benefit of innovation *V* to individual *n* at time *t*, B_{Vnt} is the benefit to adopting *V*, and C_{Vnt} is the cost of adopting *V*.[14] The probit model assumes that an individual will adopt a technology when equation (1) holds. Since benefits and costs cannot be separately identified and the data are cross-sectional, covariates refer to the net benefit:

$$NB_{Vn}= B_{Vn}-C_{Vn} = S_n\gamma+L_n\delta+ Z_{Vn}\beta+\varepsilon_{Vn} \tag{2}$$

Here $S_n=1$ if individual *n* was a student in the mid-1990s and zero otherwise; $L_n=1$ if individual *n* lives with someone who was a student in the mid-1990s and zero otherwise; and Z_{Vn} is a vector of covariates relating to adoption of technology *V* by individual *n*. Z_{Vn} includes age, education, occupation, and industry controls to see if the marginal effect of having been (living with) a student in the mid-1990s on the net

[14] Implicitly, the cost term, C, includes the opportunity cost of not adopting at some other time (Ireland and Stoneman, 1986).

benefit to Internet adoption is positive. If the marginal effect remains positive, then effects based on observed age, education, occupation, and industry cannot be complete explanations.[15]

As described in Section 2, there are many other possible sources of unobserved heterogeneity (e.g. this cohort may be particularly technology-savvy). Ideally, the unobserved heterogeneity would be addressed with an instrument that is correlated with university attendance but not tied to Internet usage. Unfortunately, no such instrument is available. Therefore, I estimate a number of other probit models using different computing technologies and different time periods. In each of these models, the dependent variable is changed but the right hand side variables are not changed. Since the coefficient on having been a student in the mid-1990s is positive for Internet use but not for other technologies, I argue that unobserved heterogeneity is not a complete explanation for the relationship between university attendance in the mid-1990s and Internet adoption. This assumes that there is not some form of unobserved heterogeneity that led people to university in the 1990s and caused them to adopt the Internet, but did not cause them to adopt other forms of computing technology. Section 5 details the results of these probit regressions.

4.2 Understanding the Role of Universities: Nested Model

This section provides a model that allows for a deeper understanding of the role universities may have played in Internet adoption. The Internet, like many other technologies, has no value in itself. It is a platform for other technologies and only has value when combined with software applications. Consequently, the decision by a given individual to adopt the Internet will depend on the value that individual places on the various applications. In particular, an individual will adopt Internet technology if the benefits of adoption outweigh the costs. There are costs in adopting the core technology and in adopting the applications. For example, costs in adopting the core technology include getting a computer and a modem, learning how to access the Internet, signing up with an Internet service provider, and

[15] If there are occupation/age-specific or occupation/education-specific effects that influence demand for Internet-savvy employees, then these controls will be insufficient. Therefore, to show robustness, I estimate a model with

paying the access fee. Costs in adopting email include signing up for an account, learning how to use email, learning people's email addresses, and perhaps typing. Benefits, on the other hand, only accrue to the individual if they adopt applications. Just having Internet access will not generate utility. It is the use of applications like email, information access, and e-commerce that generates utility. Therefore, by determining whether the net benefit to Internet adoption is higher for students in the mid-1990s because of utility from applications or because of utility from the underlying technology, I can separately identify whether these students were more likely to adopt the Internet because of increased benefits or reduced costs. [16]

Jimeniz and Greenstein (1998) also examine the Internet as a nested diffusion process but in a different context. In a largely descriptive paper, they emphasize that Internet applications will only diffuse with the widespread adoption of personal computers and networking technologies. Gandal, Kende, and Rob (2000) estimate the relationship between hardware and software diffusion in the compact disc market. Like most other research on contingent products,[17] they focus on identifying network externalities and the option value of adoption. I address a different question: given that Internet adoption will involve different applications for different people, what are the actual drivers of adoption?

To formalize these concepts, let NB_{In}=B_{In}-C_{In} be the net benefit of using the Internet at all and let NB_{an}=B_{an}-C_{an} be the net benefit of using particular application *a* (such as email, information search, or e-commerce). Individuals derive utility from the Internet only to the extent that they use applications. Therefore $\mathrm{B}_{\mathrm{In}} = \sum_{a \in A_n} NB_{an}$ and $\mathrm{NB}_{\mathrm{In}} = (\sum_{a \in A_n} NB_{an}) - C_{\mathrm{In}}$, where A_n is the set of all Internet applications used by individual *n*. Costs, however, may relate to the act of getting online. Consequently, I can

occupation/age and occupation/education interactions and another model that excludes all at-work users.

[16] I identify net benefits from application usage rather than from intensity of use. While the census did not ask questions about intensity of use, there is some evidence from the Nielsen data that application use is correlated with intensity. 62% of users of all three applications (e-commerce, e-information, and e-communication) use the Internet daily. 44% of users of two applications and 24% of users of just one application use the Internet daily. Furthermore, student in the mid-1990s who use the Internet are more likely than other Internet users to go online daily (59% vs. 51%). This data is not rich enough to conduct econometric analysis.

[17] The extensive research on the diffusion of contingent products in the marketing literature began with Bayus (1987) on CDs.

separately identify B_I from C_I. This allows for separate measurement of the effect of universities on benefits and costs.[18]

In each period, individuals have a nested decision problem. They must decide whether to adopt the Internet, and if they adopt the Internet, they must decide which applications to adopt. There are *N* households indexed by *n* and *A* applications indexed by *a*. The subscript for the Internet in general is *I*. The net benefit from adopting the Internet is then going to be

$$NB_{In} = \left[\sum_{a=1}^{A} \alpha_a \max\left(B_{an} - C_{an}, 0\right) \right] - C_{In} \tag{3}$$

where α_a is the weighting on the net benefit of application *a*. The net benefit from adopting a particular application *a* will be

$$NB_{an} = B_{an} - C_{an} \tag{4}$$

Since I am not able to distinguish benefits from costs at the application level, the net benefit of using a particular application is measured by

$$NB_{an} = S_n \gamma_a + L_n \delta_a + Z_n \beta_a + \varepsilon_{an} \tag{5}$$

As before, $S_n=1$ if individual *n* was a student in the mid-1990s and zero otherwise; $L_n=1$ if individual *n* lives with someone who was a student in the mid-1990s and zero otherwise; and Z_n is the same vector of other covariates defined in Section 4.1. The household varying error term, ε_{an}, is distributed normally with mean zero to give a standard probit specification on application adoption probability.

The net benefit of adopting the Internet in general will be:

$$NB_{In} = X_n \theta_I + \varepsilon_{In} + \sum_{a \in \overline{A_n}} \alpha_a NB_{an} \tag{6}$$

where $\overline{A_n}$ is the set of applications that individual *i* uses and $X_n \theta_I = S_n \gamma_I + L_n \delta_I + Z_n \beta_I$. Again assuming the error term is normally distributed, the probability of adopting the Internet is

[18] I cannot separate B_a from C_a. Therefore it is only at the level of the underlying technology, not at the application level, that I separate benefits from costs (e.g. physical costs, complexity, and trialability). Also, while I control for the application uses in the CPS, there are other applications. Therefore the measured C_{In} includes benefits from uncommon applications.

$$\Pr(y_{In}=1)=\int_{\varepsilon_{In}}\int_{\varepsilon_{1n}}\ldots\int_{\varepsilon_{An}} 1\left[X_n\theta_I+\varepsilon_{In}+\sum_a \alpha_a 1[X_n\theta_a+\varepsilon_{an}]\right]\phi(\varepsilon)d\varepsilon \tag{7}$$

Here, *1[x]* is an indicator function for *x>0*, *A* is the number of applications, ε represents the vector $\{\varepsilon_{An},\ldots,\varepsilon_{1n},\varepsilon_{In}\}$, and $\phi()$ is the P.D.F. of the multivariate normal. Assuming the epsillons are independent, integrating out the application epsillons yields:

$$\Pr(y_{In}=1)=\int_{\varepsilon_{In}} 1\left[X_n\theta_I+\varepsilon_{In}+\sum_a \alpha_a E[X_n\theta_a+\varepsilon_{an}\mid y_{an}=1]\Pr(y_{an}=1)\right]\phi(\varepsilon_{In})d\varepsilon_{In} \tag{8}$$

Let $\tilde{\varepsilon}_{an}=E[\varepsilon_{an}\mid\varepsilon_{an}>-X_n\theta_a]$. This value can be calculated using the distribution of the truncated normal (Greene 1997, p. 207-8). Then (8) becomes:

$$\Pr(y_{In}=1)=\int_{\varepsilon_{In}} 1\left[X_n\theta_I+\varepsilon_{In}+\sum_a \alpha_a (X_n\theta_a+\tilde{\varepsilon}_{an})\Pr(y_{an}=1)\right]\phi(\varepsilon_{In})d\varepsilon_{In} \tag{9}$$

This model was estimated as follows.[19] First, I estimated the decisions over applications in equation (5) as probit models. With estimates of θ_a, I computed $X_n\theta_a$, $\tilde{\varepsilon}_{an}$, and $\Pr(y_{an}=1)$ for every consumer, regardless of whether the consumer adopted the application. We these terms, the coefficients θ_I and α_a were calculated using a probit regression in a second stage. The net benefit of the applications is only identified up to scale. As a result, the absolute value of α_a has no economic meaning. It re-weights the values in order to estimate the net benefit to Internet adoption from them. This is similar to an inclusive value in a nested logit regression.[20]

I focus on five applications: e-commerce, e-information, e-communication, e-entertainment, and an "other" category. Individuals are defined as using e-commerce if they claim to have made an online purchase of any kind that year. Individuals are defined as using e-information if they claim to have visited websites that contain information on news, product information, or health. Individuals are defined as using e-communication if they claim to have used the Internet for email or chat in the past year. E-

[19] I thank an anonymous referee for suggesting this estimation strategy.
[20] Unlike a nested logit regression with a utility maximization model, α does not need to be bounded.

entertainment means the individual played online games or looked at television or movie websites. The "other" category included all other application-related uses included in the survey (online banking, online stock trading, online job search, and online government services).

The coefficients on the first stage probit regressions for each application determine whether universities served to help e-communication, e-commerce, e-information, or some other factor. The coefficients on the second stage regression will examine whether universities helped reduce the underlying costs of adoption.

The next two sections present the results.

5. Internet Adoption

5.1 Students in the Mid-1990s Became Leaders in Internet Adoption

Table 4 shows the core results, which should be interpreted as a difference-in-difference. The coefficient on *Born 1971-79 & postsecondary graduate* (in the first row) shows that post-secondary education has a particularly large effect on people in this age cohort. All results are marginal effects unless otherwise stated.[21] Column (1) presents the main regression using data from the September 2001 CPS Supplement and column (2) presents its coefficients; column (3) drops occupation and industry fixed effects; column (4) excludes individuals who use the Internet at work; and column (5) includes occupation-age and occupation-education interactions. The qualitative results do not vary much from regression to regression, although the marginal effect of having a postsecondary education for those born 1971 to 1979 varies between 2.9% and 5.9%.[22]

The difference between columns (1) and (3) shows the importance of occupation and industry in determining whether an individual uses the Internet. The higher marginal effect on *Born 1971-79 &*

[21] These marginal effects are based on the average levels of the covariates. For example, 43.8% of people without a postsecondary education who were born from 1971-79 adopt the Internet. Column (1) of Table 4 shows the marginal effect of postsecondary education to be 21.6% and the additional effect of education for the cohort born 1971-79 to be 3.5%. Therefore the probability of adoption for someone born 1971-79 with a post-secondary education will be 68.9%, all else equal. Since actual adoption rates are 88.4%, the remaining 19.5% must be due to other factors such as occupation, marital status, citizenship, etc.

postsecondary graduate in column (3) relative to column (1) supports occupation and industry effects as explaining around 40% of the cohort effect. Columns (4) and (5) address the possibility that demand for Internet savvy employees varies not only by occupation, but by occupation-age and occupation-education. Column (4) shows that the result is robust to excluding all people who use the Internet at work. Column (5) includes interactions for each of the 46 occupations, age, and education. With these extra variables, the main result holds and the fit of the model (measured by the log likelihood) changes little.

Many of the other results are likely familiar to readers who have studied technology adoption (e.g. Jimeniz and Greenstein, 1998; Compaine, 2001). Young, wealthy, educated city dwellers are most likely to adopt the technology.[23] The spline regression shown in Figure 2 splits the sample by age cohort, where a separate coefficient is estimated for each age and for each interaction term of age and postsecondary graduation. The group highlighted by the box (age 22-30 in 2001) is the cohort of people that was born between 1971 and 1979. The marginal effect of education for this group is clearly higher than for the groups that immediately preceded and followed it. If instead, the spline regression is done by cohort, the marginal effect of postsecondary graduation is still highest for those born from 1971-79. The effect for this cohort is higher than the effect for the 1960-70 cohort with 95% confidence in a two-tailed test. The difference remains even in an elasticity sense. In particular, the marginal effect for the 1971-79 cohort divided by the average propensity to adopt for that cohort is 0.342. For the 1960-70 and 1950-59 cohorts the comparable numbers are 0.313 and 0.292 respectively.

The result is robust to different definitions of the cohort that attended school in the mid-1990s, but does not extend to the cohort that would have left school before 1993. In particular, the qualitative results hold if the relevant cohort is restricted to 1971-75 or to 1976-79. Furthermore, if instead of 1971-79 the cohort is defined as 1970-79 the results hold. Interestingly, the results do not hold under various

---

[22] Online appendix Tables A.2 and A.3 show the results are robust to a number of other specifications. Throughout Section 5, any results not shown are available in the online appendix.

[23] While the *Born 1971-79* dummy has a negative sign, the overall age trend is still negative. In particular, without the age trend, the *Born 1971-79* dummy would be positive. The spline regressions in Figure 2 and online appendix Table A.4 provide further evidence that it is the inclusion of the age trend that leads to the negative sign on this dummy.

definitions of the cohort that attended university before the mid-1990s. Whether defined as 1960-70, 1968-70, 1966-70, or 1966-69, the cohort that attended university before the mid-1990s are no more likely to use the Internet than might be expected given their age and education.

Table 5 presents results using A.C. Nielsen Canada's *Internet Planner*. The advantage of this data set is that people who were students in the mid-1990s (1995-97) are accurately identified due to the panel nature of the data. The overall results do not change; however the marginal effect of going to university in the mid-1990s is more than 8%. This effect may be larger because people who attended school in the relevant period are more accurately identified. On the other hand, it may also be due to weaker demographic controls or the unusual characteristics of students in this data set.[24] Column (4) shows that the results hold with an age-education interaction term. The main result does not seem to be a function of education being more important in adoption for younger individuals.

Table 6 shows that these effects are particularly strong in households with income under $40,000 and for non-citizens. Furthermore, column (2) shows that there is no significant relationship for those households earning over $60,000. For households with income under $40,000 the 6.0% marginal effect represents 15.9% of the average adoption rate of 37.7%. For non-citizens, the 9.2% represents 25.7% of the average adoption rate of 35.8%. Universities therefore appear to have helped close the income-related digital divide while simultaneously increasing the education-related digital divide in the low-income segment of the population. Furthermore, since low-income people are less likely to ever adopt, universities may have helped the technology diffuse to groups that would not otherwise have adopted it.

In summary, the effect of postsecondary education is especially large for the group that attended university in the mid-1990s. This effect is largest in low income households, and it is only partially explained by age, education, occupation, and industry controls.

[24] As mentioned earlier, all members of this data set are household heads who did not change residences. Unfortunately, there is insufficient data to use household-level controls (either fixed or random effects) in this data. Identification with household-level controls would rely on those people who were not students at the beginning of

5.2 People Living with Students in the Mid-1990s are More Likely to Adopt

The teaching role of universities in the diffusion of the Internet goes beyond direct transmission to students. The regressions described above also show that adults living with people who were students during the relevant time period are more likely to use the Internet. For example in Table 4, row 2 shows the effect of having someone who was a student between 1993 and 1997 in the household. In each model, the marginal effect of having someone who likely attended university in the mid-1990s is positive and significant. In most cases, this indirect effect of universities on adoption is larger than the direct effect on former students, likely reflecting the possibility that these are people who are otherwise unlikely to adopt. As in Section 5.1, Table 6 shows that the results are particularly strong for low income households and for non-citizens. Again, this has interesting implications on the digital divide and it suggests that universities did more than merely speed up the diffusion process by a few months.

There are two likely explanations for this result. First, as a communication technology, the Internet may exhibit direct network externalities. In particular, the value of adopting the technology depends on the number of other people an individual knows who have already adopted. Since household members certainly know each other, the benefit to adoption will increase if a household member is online (because of university attendance in the mid-1990s). Second, former students may have taught household members how to use the Internet. The costs of learning how to use the technology may be lower if a member of the household is already online.

5.3 Other Computing Technologies Give Insignificant Results

The results presented in sections 5.1 and 5.2 do not apply to uses of computer technologies that are not Internet related. Figure 3 presents model (1) of Table 4 employing Internet use ("Connect to the Internet"), "Use a computer at work", "Do word processing or desktop publishing" at home or work, "Play games on the computer", and "Use home computer to manage household records or finances" as

---

the sample, became students in the middle of the sample, and then were not students by the end of the sample. There are only 38 such people in the data.

dependent variables. What is striking about Figure 3 is that there is no significant relationship between attendance at university in the mid-1990s and use of computer applications that are not Internet related. There is also no significant relationship between living with someone who attended university in the mid-1990s and use of computer applications that are not Internet related except for a small relationship in word processing. Furthermore, the direct and indirect effects of university attendance in the mid-1990s on Internet use are much larger than the effects on other computer applications. Figure 3 provides strong evidence that it is not unobserved heterogeneity that is driving the results in Sections 5.1 and 5.2. Since people who attended university in the mid-1990s and those who live with them are not more likely than others to adopt computing technologies besides the Internet, it is unlikely that they are particularly technology-savvy.

Furthermore, an identical experiment with the October 1994 CPS Supplement data on modem use gives no significant results for people aged 22 to 30 at the time. Unlike the cohort born 1971-79, universities did not have an extra impact on Internet adoption for the cohort born 1964-72 even when they were 22 to 30 years old. This provides additional evidence that something different happened in the mid-1990s.

It is important to consider why universities had an impact on Internet use, but not on the use of other computing technologies. The Internet was a technology developed mainly by the U.S. government for research institutions. By 1993 it was widely diffused to universities but few other institutions had adopted it. In contrast, other computing technologies generally diffused simultaneously to businesses and to universities from the for-profit companies that developed them. Furthermore, the Internet is a communications technology. Therefore, direct network externalities may have meant that the benefits to adoption for students were high because it allowed them to communicate with their professors who had already adopted. Word processing and computer games are not communications technologies and should not display network effects in the same way. Overall, universities in the 1990s were particularly well-positioned to play an important role in the diffusion of the Internet.

6. Separating Costs and Benefits of Internet Adoption

Table 7 presents the results of the nested diffusion regressions. Column (1) shows the results of coefficients on the propensity to adopt the Internet in general, controlling for benefits related to known application use. The application-specific columns show the coefficients on propensity to adopt each application. As before, The variables of interest are labeled *Born 1971-79 & postsecondary graduate* and *In household with someone born 71-79 & postsecondary graduate*. The other variables in the regressions serve as controls and are identical to those used in the previous section.

The first row of the first column suggests that, controlling for the benefits from application adoption, Internet adoption is not significantly related to having been a student in the mid-1990s. As explained in Section 4.2, this suggests that the universities did not significantly lower the costs of connecting to the Internet or of learning how to use it. The second row of the first column, however, suggests that living with someone who attended university during this time period did lower these costs.

The second column of Table 7 shows that people who were students in the mid-1990s and those who live with them are especially likely to use e-communication. Combined with the results of the previous section, this provides additional evidence that former students and those who live with them are particularly likely to adopt due to direct network externalities related to knowing more people who are online.

Former students are also slightly more likely to adopt e-information, e-commerce, and "other" applications than the rest of the population, although the magnitude of the difference is generally smaller than for e-communication. Besides communication, people who live with former students are also particularly likely to adopt "other" applications. They are not particularly likely to adopt e- information, e-commerce, or e-entertainment.

In summary, Table 7 suggests that people who were students in the mid-1990s gained tools that were complementary to Internet use. The high net benefit of online communication suggests that the key direct role universities played in encouraging Internet usage was a network of other Internet users. Universities also played an indirect role. These former students then brought Internet use into their homes, reducing the underlying cost of adoption for those who lived with them and generating further network externalities. Perhaps costs fell because with a household member to show them how to use it, the complexity barrier was reduced. Alternatively, seeing the benefits of Internet use to a household member may have reduced observability or trialability barriers. Or maybe having a household member go through the process of setting up the wires reduced physical barriers to access. Choosing between these explanations is beyond the scope of this study.

7. Discussion and Conclusions

The correlation between university education and Internet adoption for those who attended university in the mid-1990s is much higher than for other cohorts. Those who live with these former students are also particularly likely to adopt. This is not true for other computing technologies. Both these former students and those who live with them are particularly likely to use the Internet for communication. Combined, these results suggest that universities helped generate direct network externalities that led to increased Internet adoption.

These results suggest that in addition to creating knowledge and educating the workforce, universities provide a conduit for new innovations to diffuse into society. This role links the other roles of universities together. The Internet is an example of disseminating research through teaching rather than through publications or industry relationships. More than other university-developed technologies, universities were particularly well-situated to diffuse the Internet because of its ability to faciliate communication between professors and students. Furthermore, the technology matured in the university environment long before it became commercial.

The role of higher education in diffusing technology has important implications for the digital divide and inequality.[25] Rogers (1995, p. 125) asserts that "we often find that the diffusion of innovations widens the socioeconomic gap between the higher and lower status segments of a system." This gap will

[25] See Compaine (2001) for a detailed discussion on this topic

be exacerbated if an important mechanism of diffusion is the university.[26] At the same time, the correlation is largest for groups that were generally less likely to adopt the Internet such as low income households and non-citizens. Therefore, the implications for the digital divide are a little more nuanced. In terms of Internet adoption, low income households benefit from university education more than high income households. Future attempts to diffuse technologies through schools should keep in mind that (at least in this case) schools help change the behavior of low income students and their households much more than they help change the behavior of high income students. Furthermore, this suggests that universities did not merely speed the technology diffusion process by a few months. Instead, universities appear to have diffused the Internet to groups that may not otherwise have adopted it on the same scale.

Many companies already appear to recognize the role of universities in teaching technology. For example, in 2001 Microsoft donated several million dollars to the University of Waterloo. As part of the agreement, the university agreed to start teaching Microsoft's C# programming language instead of Sun Microsystems' Java. Even though both parties were criticized widely in the Canadian media for the transaction, Microsoft likely hoped that University of Waterloo students would take the technology learned at school with them to their workplaces after they graduated (Restivo, 2001). Similarly, Apple has a history since the late 1970s of generously donating its technology to the education system. Apple founder Steven Jobs has acknowledged that these donations were partly aimed at getting former students to demand Apple computers once they entered the workforce (Kheit, 2003).

Marketers of other technologies should ask whether their products could be effectively integrated in a classroom setting. If a new technology can significantly increase the effectiveness of classroom teaching (e.g., through better communication or more effective learning techniques) then managers have an incentive to subsidize adoption of the technology in the classroom. Still, the negative press Microsoft received for its donation to the University of Waterloo shows that caution needs to be used in forcing

[26] Preliminary evidence, however, suggests that in the case of the Internet the divide between the educated and uneducated will disappear over time. University education does not appear to be highly correlated with Internet use for the cohort that entered university after 1997.

students to adopt a new technology. A new technology in the classroom should not replace a competitor's equally valuable technology. The new technology needs to be both genuinely novel and useful.

This study suggests two particular areas for future exploration. First, the results suggest a useful instrument for studying of the effects of the Internet on individual behavior, especially for low income groups. Second, many details of this role remain unexplained. How does this role of universities work? Did universities play a key role in diffusing other technologies? What are the most effective ways to diffuse technology through universities? Are there other institutions that play a similar role?

References

Acs, Z.J., Audretsch, D.B., Feldman, M.P., 1992. Real Effects of Academic Research: Comment. American Economic Review 82, 363-367.

Baily, M.H., Lawrence, R. Z., 2001. Do we have a new e-conomy? American Economic Review 9, 308-312.

Bayus, B.L., 1987. Forecasting Sales of New Contingent Products: An Application to the Compact Disc Market. Journal of Product Innovation Management 4, 243-255.

Berners-Lee, T. with Fischetti, M., 1999, Weaving the Web, HarperCollins, San Francisco.

Bresnahan, T., Brynjolfsson, E., Hitt, L., 2002. Information Technology, Work Organization, and the Demand for Skilled Labor: Firm-Level Evidence. Quarterly Journal of Economics 117, 339-376.

Compaine, B., 2001, Re-Examining the Digital Divide In: Compaine, B. and Greenstein. S. (Eds), Communications Policy in Transition: The Internet and Beyond, MIT Press, Cambridge MA.

Conley, T.G., Udry, C.R., 2003. Learning About a New Technology: Pineapple in Ghana. Mimeographed, University of Chicago.

David, P.A., 1969. A contribution to the Theory of Diffusion. Stanford Center for Research in Economic Growth, Memorandum No. 71, Stanford University.

Doms, M., Dunne, T., Troske, K.R., 1997. Workers, Wages, and Technology. Quarterly Journal of Economics 112, 253-290.

Forman, C., Goldfarb, A., 2005. The Diffusion of Information and Communication Technologies to Businesses In: Hendershott, T., (Ed), Handbook on Economics and Information Systems, Elsevier, Amsterdam, Forthcoming.

Gandal, N., Kende, M., Rob, R., 2000. The dynamics of technological adoption in hardware/software systems: the case of compact disc players. RAND Journal of Economics 31, 43-61.

Goolsbee, A., Klenow, P.J., 2002. Evidence on learning and network externalities in the diffusion of home computers. Journal of Law and Economics 45, 317-343.

Greene, W.H., 1997, Econometric Analysis, Third Edition, Prentice Hall, Upper Saddle River, NJ.

Hoffman, D.L., Novak, T.P., 2000. The Growing Digital Divide: Implications for an Open Research Agenda In: Brynjolfsson, E. and Kahin B., (Eds), Understanding the Digital Economy: Data, Tools, Research, The MIT Press, Cambridge MA, 245-260.

Ireland, N., Stoneman, P., 1986. Technological diffusion, expectations and welfare. Oxford Economic Papers 38, 283-304.

Jaffe, A.B., 1989. Real Effects of Academic Research. American Economic Review 79, 957-970.

Jalava, J., Pohjola, M., 2002. Economic Growth in the New Economy: evidence from advanced economies. Information Economics and Policy 14, 189-210.

Jimeniz, E., Greenstein, S., 1998. The Emerging Internet Retailing Market as a Nested Diffusion Process. International Journal of Innovation Management 2, 281-308.

Kenney, M., 2003. The Growth and Development of the Internet in the United States: In B. Kogut (Ed), The Global Internet Economy, Cambridge, MIT Press, 69-108.

Kheit, J., 2003. The Devil's Advocate: Rabbits From Mad Hatters: Steve Jobs and Apple, Then And Now. Mac Observer. http://www.macobserver.com/columns/devilsadvocate/2003/20030422.shtml.

Manski, C.F., 1993. Identification of Endogenous Social Effects: The Reflection Problem. Review of Economics Studies 60, 531-542.

Monjon, S., Waelbroeck, P., 2003. Assessing spillovers from universities to firms: evidence from French firm-level data. International Journal of Industrial Organization 21, 1255-1270.

National Telecommunications Information Administration, US Department of Commerce., 2002. A Nation Online. February.

Oliner, S.D., Sichel, D.E., 2000. The Resurgence of Growth in the Late 1990s: Is Information Technology the Story? Journal of Economic Perspectives 14, 3-22.

Restivo, K., 2001. Much ado about Microsoft grant: U of Waterloo $10M research fun. National Post, Aug 16, FP1**.**

Rogers, E., 1995, The Diffusion of Innovations, Fourth Edition, Free Press, New York.

Saloner, G., Spence, A.M., 2002, Creating and Capturing Value: Perspectives and Cases on Electronic Commerce, John Wiley and Sons, Inc., New York.

Stoneman, P., 2002, The Economics of Technological Diffusion, Blackwell Publishers, Oxford UK.

Table 1: Competing explanations for why people who were students in the mid-1990s are more likely than others to use the Internet

	Possible Explanations	**Source of evidence**
Explanations based on Unobserved Heterogeneity (Probit Regressions in Section 5)	1. They may be younger and more educated.	Age and education controls.
	2. They may be in occupations and industries that demand Internet skills.	Occupation and industry controls and interactions.
	3. They may be particularly technology-savvy.	Comparison to similar technologies such as word processing and computer games. Also repeat study with 1994 modem use.
Explanations based on University Attendance (Nested Regressions in Section 6)	4. They may face lower costs to adoption than others.	See if the net benefits to Internet use are particularly strong for these groups, controlling for the net benefits of applications.
	5. They may be more likely to purchase products online.	See if the benefits of online purchasing (e-commerce) are particularly strong for these groups.
	6. They may get a greater net benefit from online communication.	See if the communication benefits of the Internet are particularly strong for these groups.
	7. They may get a greater net benefit from the information available on the Internet.	See if the benefits to online information are particularly strong for these groups.

Table 2: Summary Statistics of CPS September 2001 Supplement
(only those aged 18 and above)

Variable	Mean	Standard Deviation
Born 1971-79 & postsecondary education	0.0521	0.222
Born 1971-79 (Age 18-22 from 1993-97)	0.151	0.358
Current student	0.0549	0.228
Born 1971-79 with postsec. education lives in the household	0.0382	0.192
“Connect to the Internet” (home or work)	0.563	0.496
“Use a computer at work”	0.360	0.480
“Do word processing or desktop publishing” (home or work)	0.392	0.488
“Play games on the computer” (home)	0.283	0.450
“Use home computer to manage household records or finances”	0.155	0.362
Use e-communication	0.496	0.500
Use e-information	0.488	0.500
Use e-commerce	0.362	0.480
Use e-entertainment	0.239	0.427
Use other online activity	0.273	0.446
Income <US$20,000	0.160	0.367
US$20,000<= Income<US$40,000	0.221	0.415
US$40,000<= Income<US$60,000	0.167	0.373
Income >US$60K	0.291	0.454
US citizen	0.934	0.249
Foreign Born	0.126	0.331
Metropolitan area	0.754	0.431
Homeowner	0.732	0.443
Employed	0.634	0.476
Unemployed	0.0288	0.167
Age	45.92	17.45
Female	0.527	0.499
High school diploma	0.333	0.471
Some college or university	0.191	0.393
College or university graduate	0.325	0.467
US Midwest	0.251	0.433
US Suth	0.290	0.454
US West	0.242	0.428
White	0.852	0.355
Black	0.952	0.293
Married	0.586	0.493
Never married	0.223	0.416
# Observations	104,891	

Table 3: Percent Adopting by Age and Education

	Non-Graduate	Post-Secondary Graduate
Born 1900-19	5.4%	17.9%
Born 1920-29	11.6%	38.2%
Born 1930-39	23.4%	58.3%
Born 1940-49	41.8%	78.6%
Born 1950-59	52.1%	84.0%
Born 1960-70	53.6%	86.8%
Born 1971-79	54.6%	88.4%
Born 1980-89	74.2%	81.0%

Table 4: Factors Driving Internet Use
(Marginal Effects unless otherwise stated—standard errors in parentheses)

	(1)	(2)	(3)	(4)	(5)
	Base Model	Base Model Coefficients	No Occupation or Industry Fixed Effects	Excluding At-Work Users	Interactions of Occupation, Age, and Education
Born 1971-79 & postsecondary graduate^	0.0352	0.0919	0.0592	0.0287	0.0424
	(0.0116)**	(0.0306)**	(0.0110)**	(0.0136)*	(0.0120)**
In household with someone born 71-79 & postsecondary graduate	0.0506	0.133	0.0546	0.0607	0.0515
	(0.00924)**	(0.0248)**	(0.00899)**	(0.0109)**	(0.00925)**
Age	-0.00925	-0.0239	-0.00852	-0.00934	Varies by occupation
	(0.000157)**	(0.000403)**	(0.000153)**	(0.000164)**	
Born 1971-79	-0.0512	-0.131	-0.0535	-0.0406	-0.0510
	(0.00655)**	(0.0166)**	(0.00641)**	(0.00668)**	(0.00671)**
Postsecondary graduate	0.216	0.582	0.288	0.206	Varies by occupation
	(0.00427)**	(0.0122)**	(0.00370)**	(0.00533)**	
Student	0.272	0.841	0.292	0.344	0.271
	(0.00608)**	(0.0252)**	(0.00594)**	(0.00888)**	(0.00622)**
Female	0.00807	0.0208	0.0643	0.00629	0.00793
	(0.00420)+	(0.0108)+	(0.00357)**	(0.00457)	(0.00422)+
White##	-0.00425	-0.0110	-0.0125	0.000728	-0.00420
	(0.00890)	(0.0230)	(0.00800)	(0.00970)	(0.00893)
Black##	-0.160	-0.405	-0.178	-0.153	-0.159
	(0.0109)**	(0.0276)**	(0.00964)**	(0.00999)**	(0.0109)**
Employed###	0.0888	0.228	0.188	-0.0557	0.0907
	(0.0182)**	(0.0465)**	(0.00440)**	(0.0177)**	(0.0182)**
Unemployed###	0.0488	0.128	0.120	0.0498	0.0490
	(0.0192)*	(0.0515)*	(0.00954)**	(0.0200)*	(0.0192)*
Homeowner	0.0210	0.0540	0.0172	0.0248	0.0211
	(0.00461)**	(0.0118)**	(0.00447)**	(0.00495)**	(0.00463)**
Metropolitan Area	0.0433	0.111	0.0488	0.0405	0.0423
	(0.00483)**	(0.0123)**	(0.00418)**	(0.00509)**	(0.00484)**
Married ####	0.0529	0.136	0.0502	0.0722	0.0522
	(0.00501)**	(0.0129)**	(0.00489)**	(0.00543)**	(0.00503)**
Never Married####	-0.0361	-0.0927	-0.0414	-0.0266	-0.0375
	(0.00656)**	(0.0168)**	(0.00636)**	(0.00711)**	(0.00662)**
US Citizen	0.109	0.277	0.148	0.108	0.111
	(0.0104)**	(0.0262)**	(0.00991)**	(0.00996)**	(0.0105)**
Foreign Born	-0.132	-0.335	-0.149	-0.123	-0.130
	(0.00809)**	(0.0204)**	(0.00764)**	(0.00780)**	(0.00811)**
Household Income: Less than $20,000#	-0.127	-0.322	-0.150	-0.102	-0.125
	(0.00669)**	(0.0169)**	(0.00646)**	(0.00651)**	(0.00672)**
Household Income: $20,000-$40,000#	0.00101	0.00260	-0.00767	0.0131	0.00180
	(0.00557)	(0.0144)	(0.00545)	(0.00607)*	(0.00558)
Household Income: $40,000-$60,000#	0.103	0.275	0.111	0.120	0.104
	(0.00564)**	(0.0155)**	(0.00550)**	(0.00681)**	(0.00565)**
Household Income: $60,000 or More#	0.188	0.504	0.221	0.198	0.186
	(0.00523)**	(0.0149)**	(0.00498)**	(0.00652)**	(0.00525)**
Observations	104,891	104,891	104,891	77,603	104,891
Log Likelihood	-49,127	-49,127	-52,381	-41,086	-48,971

+ significant at 10%; * significant at 5%; ** significant at 1%
Unless otherwise stated, regressions include a constant and state, industry, and occupation fixed effects.
^I sometimes refer to this group as "Former Students" in the text. They are in the age cohort that attended university from 1993-97.
#Base=refused to answer; ##Base=Other; ###Base=out of labor force; ####Base=widowed, divorced, or separated

Table 5: Probit Regressions using the Nielsen data
(Marginal Effects unless otherwise stated—standard errors in parentheses)

	(1)	(2)	(3)	(4)
Variable	Year=2000$	Year=2000$ Coefficients	Year=2000$	Year=2000$
Student 1995-97	0.0835	0.235		0.104
	(0.0387)*	(0.116)*		(0.048)*
Student 1995-99			0.110	
			(0.0305)**	
Number of other users in FSA	0.00501	0.0135	0.00495	0.005
	(0.00144)**	(0.00387)**	(0.00144)**	(0.001)**
Current student	0.176	0.545	0.157	0.174
	(0.0551)**	(0.208)**	(0.0598)**	(0.055)**
Personal income C$30-C$69K#	0.148	0.402	0.150	0.149
	(0.0154)**	(0.0422)**	(0.0154)**	(0.015)**
Personal income > C$70K#	0.298	0.938	0.299	0.298
	(0.0147)**	(0.0585)**	(0.0147)**	(0.015)**
Postsecondary education	0.167	0.451	0.166	0.248
	(0.0139)**	(0.0379)**	(0.0139)**	(0.053)**
Age	-0.175	-0.470	-0.172	-0.159
	(0.00973)**	(0.0261)**	(0.00979)**	(0.013)**
Age*Post-secondary educ.				-0.029
				(0.019)
Canadian west##	0.0990	0.271	0.0983	0.097
	(0.0219)**	(0.0612)**	(0.0219)**	(0.022)**
Ontario##	0.116	0.323	0.116	0.114
	(0.0227)**	(0.0662)**	(0.0227)**	(0.023)**
Quebec##	0.00698	0.0188	0.00733	0.006
	(0.0235)	(0.0633)	(0.0235)	(0.023)
Female	-0.0164	-0.0441	-0.0158	-0.015
	(0.0140)	(0.0376)	(0.0140)	(0.014)
N	5,519	5,519	5,519	5,519
Log likelihood	-3,074	-3,074	-3,082	-3,081

+ significant at 10%; * significant at 5%; ** significant at 1%
$includes a constant and dummy variables on whether have data on the individuals for 1995 through 1999.
$$includes a constant and dummy variables on whether have data on the individuals for 1995, 1996, and 1997.
#Base=income<C$30K; ##Base=Atlantic Canada

Table 6: Is the effect particularly large for any subgroups of the population?
(Marginal Effects—standard errors in parentheses)

	(1)	(2)	(3)
	Income Below $40,000	Income $60,000 and Higher	Not a US Citizen
Born 1971-79 & postsecondary graduate^	0.0602	0.0135	0.0922
	(0.0196)**	(0.0127)	(0.0392)*
In household with someone born 71-79 & postsecondary graduate	0.129	-0.00227	0.0595
	(0.0199)**	(0.00848)	(0.0297)*
Age	-0.00849	-0.00601	-0.00725
	(0.000226)**	(0.000204)**	(0.000706)**
Born 1971-79	-0.0228	-0.0469	-0.0437
	(0.00858)**	(0.0105)**	(0.0194)*
Postsecondary graduate	0.230	0.122	0.287
	(0.00859)**	(0.00516)**	(0.0205)**
Student	0.390	0.100	0.440
	(0.0127)**	(0.00558)**	(0.0281)**
Female	0.0101	-0.00237	-0.0438
	(0.00632)	(0.00496)	(0.0168)**
White##	-0.00872	0.0138	-0.0833
	(0.0130)	(0.0109)	(0.0190)**
Black##	-0.148	-0.0649	-0.0820
	(0.0128)**	(0.0164)**	(0.0262)**
Employed###	0.0548	0.0604	0.0955
	(0.0255)*	(0.0247)*	(0.0680)
Unemployed###	0.0290	0.0449	0.109
	(0.0283)	(0.0190)*	(0.0806)
Homeowner	0.0118	0.0176	0.0493
	(0.00630)+	(0.00740)*	(0.0158)**
Metropolitan Area	0.0415	0.0313	0.00680
	(0.00679)**	(0.00674)**	(0.0300)
Married ####	0.0223	0.0446	0.00576
	(0.00713)**	(0.00775)**	(0.0230)
Never Married####	-0.0387	-0.0247	-0.0330
	(0.00857)**	(0.0103)*	(0.0267)
Us Citizen	0.106	0.0436	
	(0.0133)**	(0.0146)**	
Foreign Born	-0.118	-0.0702	
	(0.0109)**	(0.0109)**	
Household Income: Less than $20,000#	-0.124		-0.0824
	(0.00587)**		(0.0215)**
Household Income: $20,000-$40,000#			0.00228
			(0.0213)
Household Income: $40,000-$60,000#			0.0829
			(0.0259)**
Household Income: $60,000 or More#			0.208
			(0.0266)**
Observations	39,937	30,503	6,905
Log Likelihood	-19,193	-11,530	-2,795

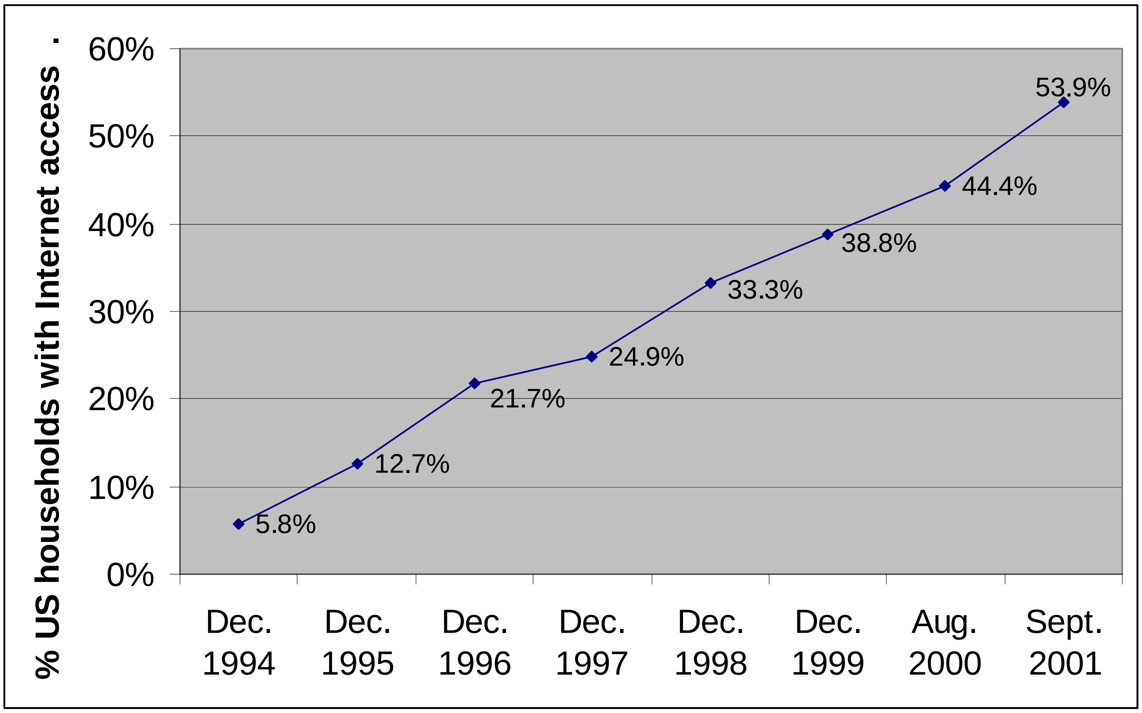
+ significant at 10%; * significant at 5%; ** significant at 1%
All regressions include a constant and state, industry, and occupation fixed effects.
^I sometimes refer to this group as "Former Students" in the text. They are in the age cohort that attended university from 1993-97.
#Base=refused to answer; ##Base=Other; ###Base=out of labor force; ####Base=widowed, divorced, or separated

Table 7: Nested Diffusion Regression
(Marginal Effects—standard errors in parentheses)

	(1)	(2)	(3)	(4)	(5)	(6)
	Use Internet	Use E-Communication	Use E-Information	Use E-Commerce	Use E-Entertainment	Use Other
Born 1971-79 &	0.0241	0.0351	0.0130	0.0253	0.00828	0.0463
postsecondary graduate^	(0.0276)	(0.00574)**	(0.00701)+	(0.0105)*	(0.0109)	(0.0113)**
In household with someone born	0.0401	0.0118	0.00599	0.00721	-0.0143	0.0233
71-79 & postsecondary graduate	(0.0118)**	(0.00570)*	(0.00624)	(0.00938)	(0.00971)	(0.00996)*
Age	-0.0154	-0.000195	-0.00192	-0.00360	-0.00496	-0.00140
	(0.00113)**	(0.000128)	(0.000129)**	(0.000203)**	(0.000215)**	(0.000217)**
Born 1971-79	-0.0543	-0.000427	0.00309	-0.00307	0.000425	0.0270
	(0.00767)**	(0.00465)	(0.00500)	(0.00775)	(0.00798)	(0.00818)**
Postsecondary graduate	0.213	0.0380	0.0449	0.0393	-0.0446	0.11797
	(0.0234)**	(0.00321)**	(0.00325)**	(0.00503)**	(0.00528)**	(0.00530)**
Student	0.433	0.0360	0.0957	-0.0137	0.00468	-0.0567
	(0.0105)**	(0.00467)**	(0.00312)**	(0.00928)	(0.00944)	(0.00958)**
Female	-0.0867	0.0129	-0.0293	-0.0265	-0.0811	-0.0432
	(0.0198)**	(0.00303)**	(0.00309)**	(0.00466)**	(0.00483)**	(0.00494)**
White##	0.0162	0.0140	0.0214	0.0247	-0.0257	0.0131
	(0.0141)	(0.00691)*	(0.00738)**	(0.0104)*	(0.0108)*	(0.0110)
Black##	-0.205	-0.0498	-0.0333	-0.0535	0.00515	-0.00451
	(0.0271)**	(0.00983)**	(0.00946)**	(0.0132)**	(0.0133)	(0.0136)
Employed###	0.0335	-0.0197	-0.0279	-0.0339	-0.0502	-0.0719
	(0.0244)	(0.0125)	(0.0136)*	(0.0200)+	(0.0213)*	(0.0213)**
Unemployed###	-0.0898	-0.0113	-0.00160	-0.0141	-0.0133	0.160
	(0.0342)**	(0.0161)	(0.0168)	(0.0233)	(0.0235)	(0.0229)**
Homeowner	0.0387	-0.00465	-0.00798	-0.0225	-0.0238	-0.0735
	(0.0114)**	(0.00336)	(0.00348)*	(0.00537)**	(0.00563)**	(0.00574)**
Metropolitan Area	0.0219	0.01492	0.00264	-0.00444	-0.00592	0.0388
	(0.0105)*	(0.00368)**	(0.00369)	(0.00553)	(0.00581)	(0.00589)**
Married ####	0.0620	0.00170	0.00209	0.00881	-0.0134	-0.0272
	(0.00747)**	(0.00396)	(0.00401)	(0.00628)	(0.00662)*	(0.00672)**
Never Married####	-0.0746	0.0000224	-0.0319	-0.00480	0.0107	-0.0429
	(0.0137)**	(0.00492)	(0.00552)**	(0.00787)	(0.00820)	(0.00834)**
US Citizen	0.140	-0.0113	0.00368	0.00882	0.0128	-0.0152
	(0.0135)**	(0.00769)	(0.00867)	(0.0133)	(0.0138)	(0.0142)
Foreign Born	-0.182	-0.0228	-0.0241	-0.0583	-0.0292	-0.0157
	(0.0200)**	(0.00670)**	(0.00688)**	(0.00998)**	(0.0100)**	(0.0103)
Household Income:	-0.170	-0.0327	-0.0136	-0.0261	-0.00532	0.0310
Less than $20,000#	(0.0189)**	(0.00669)**	(0.00635)*	(0.00961)**	(0.00988)	(0.0101)**
Household Income:	0.0126	-0.0102	0.00466	0.00315	0.0146	0.0160
$20,000-$40,000#	(0.00899)	(0.00469)*	(0.00448)	(0.00724)	(0.00767)+	(0.00783)*
Household Income:	0.131	0.00295	0.0163	0.0287	0.0211	0.0303
$40,000-$60,000#	(0.00944)**	(0.00442)	(0.00426)**	(0.00697)**	(0.00750)**	(0.00762)**
Household Income:	0.223	0.0232	0.0337	0.0482	0.00137	0.0570
$60,000 or More#	(0.0163)**	(0.00405)**	(0.00405)**	(0.00639)**	(0.00679)	(0.00690)**
α		0.0399	-0.783	0.203	-0.777	0.764
(from "Use Internet" regression)		(0.165)	(0.121)**	(0.303)	(0.188)**	(0.147)**
Observations	104,884	59,081	59,085	59,085	59,085	59,081
Log Likelihood	-49,079	-20,355	-21,438	-37,557	-38,859	-38,599

+ significant at 10%; * significant at 5%; ** significant at 1%
All regressions include a constant and state, industry, and occupation fixed effects.
^I sometimes refer to this group as "Former Students" in the text. They are in the age cohort that attended university from 1993-97.
#Base=refused to answer; ##Base=Other; ###Base=out of labor force; ####Base=widowed, divorced, or separated

Figure 1: Percentage of US Households with Internet Access 1994-2001



Source: Up to 1999 from Saloner and Spence (2002) citing U.S. Internet Council, "State of the Internet: USIC's Report on Use and Threats in 1999" citing Forrester Reports; 2000 and 2001 from National Telecommunications Information Administration (2002). The National Telecommunications Information Administration (2002) reports Dec. 1998 access as 32.7%.

Figure 2: Marginal Effect of Education by Age

Regression has separate covariates for each age and the interaction of post-secondary education with each age

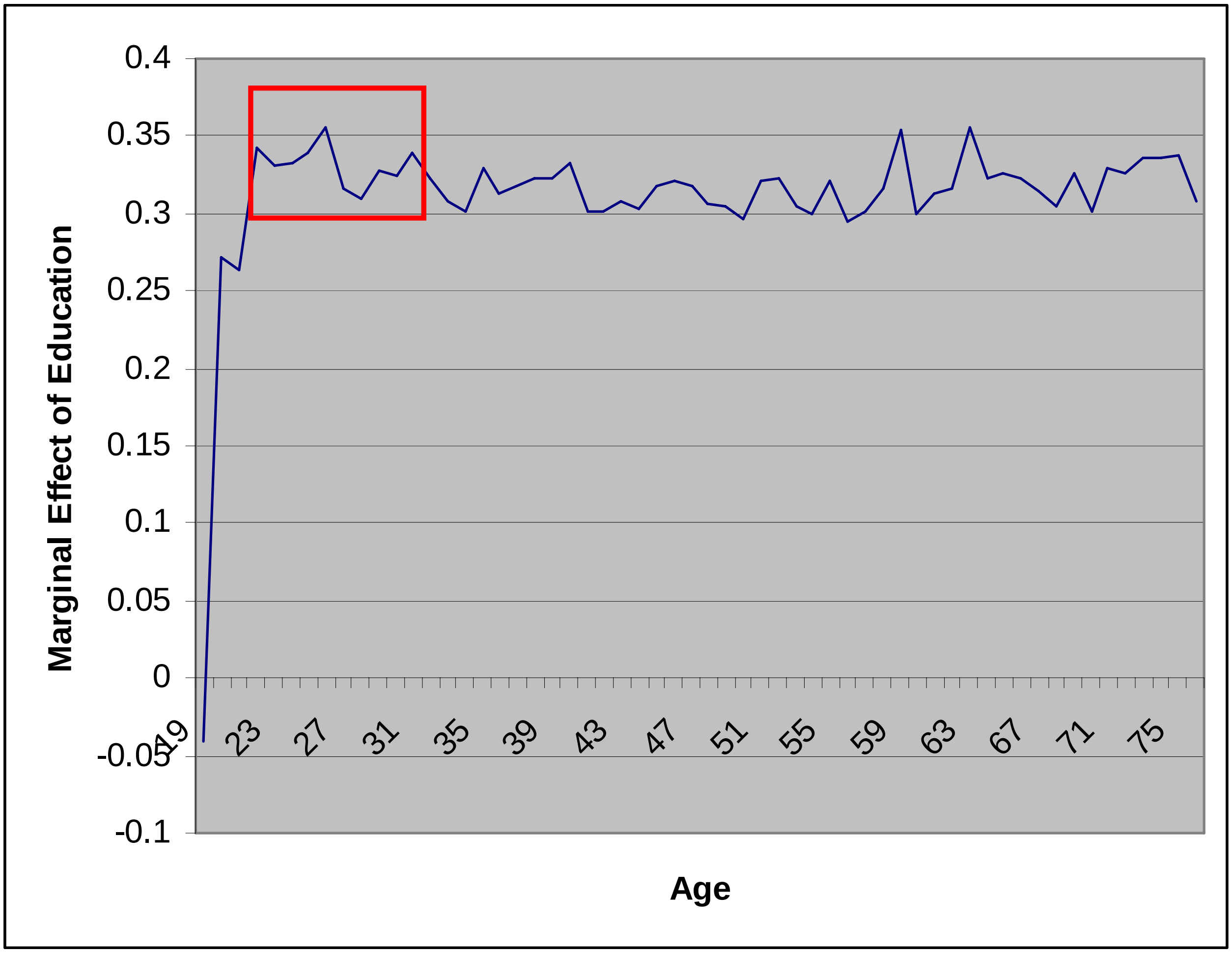
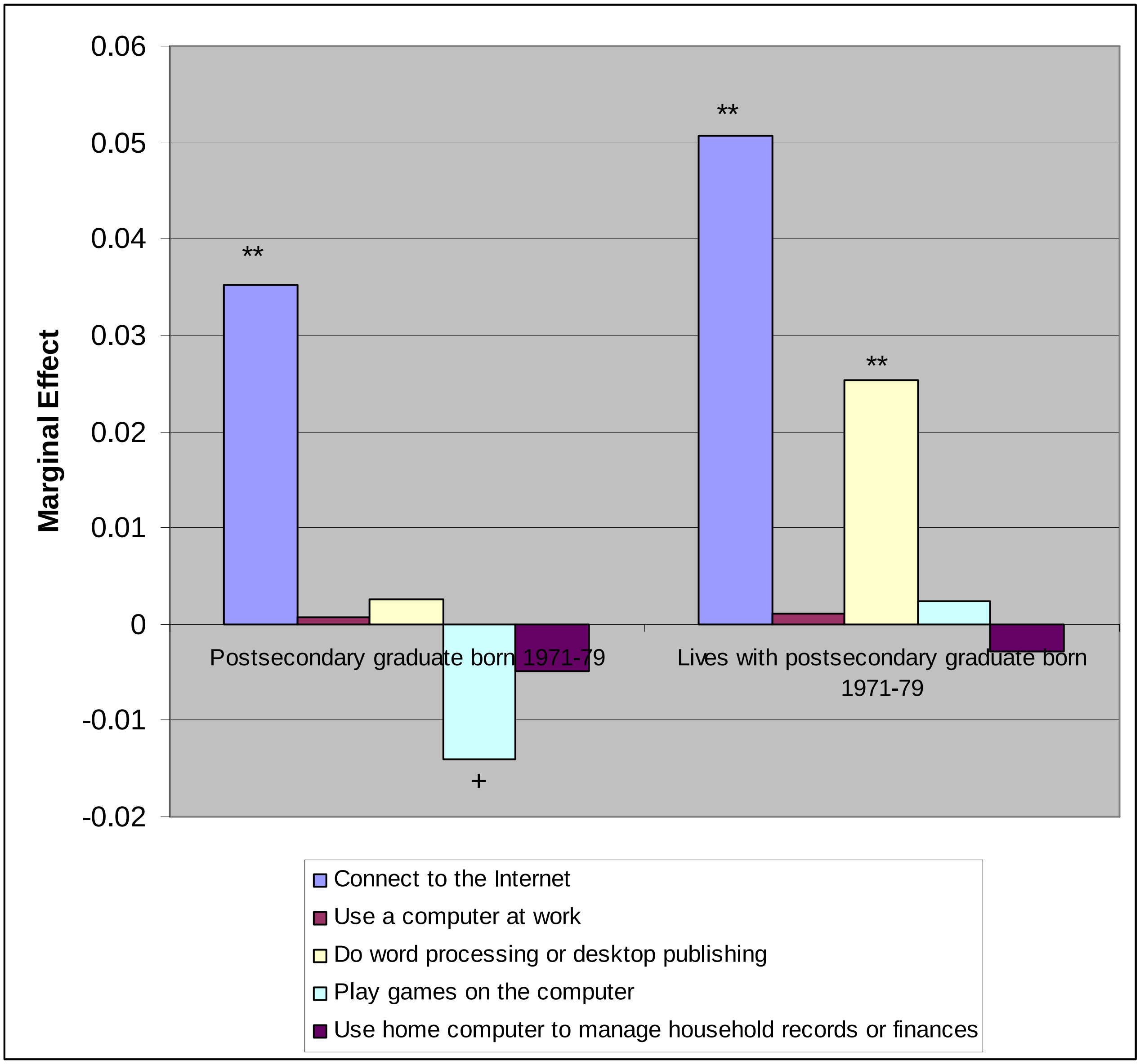


Figure 3: Internet use is different from other computer uses



+ significant at 10%; ** significant at 1%