Internet Adoption and Usage Patterns are Different: Implications for the Digital Divide

Avi Goldfarb and Jeff Prince*

May 2007

Abstract

There is a well-documented a "digital divide" in internet connection. We ask whether a similar divide exists for internet usage. Using a survey of 18,439 Americans, we find that high-income, educated people were more likely to have adopted the internet by December 2001. However, conditional on adoption, low-income, less-educated people spend more time online. We examine four possible reasons for this pattern: 1) differences in the opportunity cost of leisure time, 2) differences in the usefulness of online activities, 3) differences in the amount of leisure time, and 4) selection. Our evidence suggests this pattern is best explained by differences in the opportunity cost of leisure time. Our results also help to determine the potential effects of internet access subsidies.

JEL Classification: L86, L96
Keywords: internet adoption, digital divide

---

* Rotman School of Management, University of Toronto, 105 St George St, Toronto, ON and Applied Economics and Management, Cornell University, 248 Warren Hall, Ithaca, NY. Address correspondence to agoldfarb@rotman.utoronto.ca and jtp35@cornell.edu. We thank Forrester Research for the data and Chris Forman, Shane Greenstein, the editor, the associate editor, two anonymous referees, and seminar participants at the University of Toronto and Cornell University for helpful comments. This research was partially supported by SSHRC grant #538-02-1013. All opinions and errors are ours alone.

1. Introduction

There is a well-documented "digital divide" in the tendency to connect to the internet (e.g., Chinn and Fairlie, 2006; Fairlie, 2004; Fox, 2005; Hoffman and Novak, 2000). Connection alone, however, is not necessarily the best measure of the benefit of using the technology. Instead, usage generally determines how much value individuals derive from the internet. Prior research analyzing the business benefits of information technology has acknowledged this fact (e.g., Devaraj and Kohli, 2003; Astebro, 2004; Zhu and Kraemer, 2005), but there is less research on the importance of usage to households. In this paper, we find little evidence of a digital divide in usage. We argue that the pricing structure of both fixed connection fees and near-zero usage fees leads to a negative correlation between income and time online among those who have connected.

Using a survey of 18,439 Americans from December 2001, we show that the patterns of internet adoption and usage indeed differ by demographics. Specifically, we find that high-income, educated people were more likely to adopt the internet, but they also spend considerably less time online, conditional on adoption.

We then consider four explanations for this pattern: 1) low-income people have a lower opportunity cost of leisure time due to low wages, 2) low-income people find the internet more useful than others, 3) low-income people have more leisure time, and 4) the low-income people who choose to adopt the internet are those who place a particularly high value on it (i.e., selection). We compare these explanations by correcting for selection, controlling for leisure time levels, and analyzing usage of specific applications (e.g., email, telemedicine). Although data limitations mean we cannot completely rule out the possibility that selection drives the results, we argue that the empirical evidence points most strongly to low-income individuals spending more time online due to lower opportunity costs of leisure time.

These results also have implications for policy discussions on access subsidies. We conduct simulations to determine which applications low-income people would use if given internet access. Our findings indicate that this group would spend a great deal of time online and likely use the internet for activities that policymakers often view positively (e.g., news, health information). This suggests the potential benefits of subsidies; however, we also must consider other issues to determine if subsidies are worth the cost.

Among the relevant internet-usage papers, Lambrecht and Seim (2006) show that adoption of online banking depends on the user's comfort with technology but that usage depends more on the complexity of the user's banking needs. Goldfarb (2006) finds that internet usage for email and chat (rather than e-commerce and information search) was an important driver for internet technology to diffuse beyond the university setting. Sinai and Waldfogel (2004) indirectly examine usage by looking at the importance of online content in the decision to adopt. Here, we aim to show that, in terms of household demographics, adoption and usage patterns differ. We then examine possible explanations for this difference.

The next section describes the empirical strategy and the data. Section 3 shows that, controlling for many factors, internet adoption and usage have different demographic patterns. It then describes four explanations for why we observe this pattern and empirically compares them. Section 4 discusses some policy implications of our results, and Section 5 concludes.

2. Empirical Framework and Data

2.1 Empirical Framework

We model the adoption/usage decision as a two-stage process. In the first stage, households decide whether to adopt the internet; in the second stage, they decide how much time to spend online. Therefore, in the second stage, households that adopt solve the following problem:

$$\underset{I,L,M}{Max}\, u_2(I, L, M) \text{ s.t. } L + I \leq T \text{ and } M + p \leq S \qquad (1)$$

where $u_2(.)$ is utility from usage. It is increasing in *I* (leisure time spent on the internet), *L* (other leisure time), and *M* (money). *T* is total leisure time, *p* is the price of internet access, and *S* is total money available. Equation (1) thus can be restated as:

$$\underset{I}{Max}\, u_2(I, T - I, S - p) \tag{2}$$

Let *I** be the amount of internet usage that solves the above problem. Then, *I** is a function of *T*, *S*, and *p*, as well as any other characteristics that may affect the utility function. In stage one, households adopt if and only if:

$$U_1(I_i^*, T_i - I_i^*, S_i - p) \geq U_1(0, T_i, S_i) \tag{3}$$

where $U_1(.)$ is the utility from adoption. Given this utility framework, we estimate usage and adoption using a Type-II Tobit regression. Individual *i* adopts the internet if and only if:

$$U_1(I_i^*, T_i - I_i^*, S_i - p) - U_1(0, T_i, S_i) = X_{1i}\beta_1 + \alpha_1 T_i + \gamma_1 S_i + \varepsilon_{1i} \geq 0 \tag{4}$$

Therefore, assuming ε_{1i} is an individual-specific normally distributed idiosyncratic error:

$$\Pr_i(adopt) = \Pr(X_{1i}\beta_1 + \alpha_1 T_i + \gamma_1 S_i + \varepsilon_{1i} \geq 0) = \Phi(X_{1i}\beta_1 + \alpha_1 T_i + \gamma_1 S) \tag{5}$$

where X_{1i} is a vector of individual-level controls, including leisure time and demographics. Since we observe usage only if adoption takes place, we estimate the following second-stage usage equation:

$$I^*(X_{2i}, T_i, S_i) = X_{2i}\beta_2 + \alpha_2 T_i + \gamma_2 S_i + \lambda \frac{\hat{\phi}_i}{\Phi_i} + \varepsilon_{2i} \tag{6}$$

where X_{2i} is a sub-vector of the individual-level controls X_{1i}, ε_{2i} is an individual-specific normally distributed error term, and $\frac{\hat{\phi}_i}{\Phi_i}$ is the estimated inverse Mills ratio of the first-stage regression (the

“Heckman correction”). The Heckman correction allows adoption and usage to follow different patterns, assuming that the first-stage errors are normal. To allow identification on more than functional form, we include variables that correlate with adoption but not usage in the first-stage (adoption) equation as recommended by Greene (1997).[1] The Heckman correction resolves the selection problem under either of two assumptions: 1) the instruments truly correlate with adoption but not usage or 2) the first-stage error terms are normal. If both these assumptions are contradicted, then our controls for selection are inadequate. Since we cannot reject the hypothesis that the instruments are invalid, we are unable to completely eliminate selection as the driver of our results. Therefore, we interpret our results with caution.

Using a similar model, we also empirically examine which types of applications people use online. To do this, we again use a Heckman correction, but both stages are now probit regressions. For example, to explore whether people use email, the first stage is a probit regression that examines whether the person adopts the internet. Then, the second stage is also a probit regression that examines whether the person adopts email. As in Equation (6), the second-stage regression has a Heckman correction (the inverse Mills ratio of the first stage) as a covariate. In practice, we estimate this using full information maximum likelihood.

2.2 Data

The data for this study come from a detailed survey of technology choices conducted by Forrester Research. Our data set is a random sub-sample of the Forrester data and contains 18,439 American household respondents, collected in December 2001. Researchers conducted the survey through the mail

[1] Our main instruments (and reasons for choosing them) are: whether a teenager lives in the household (teenagers are more likely to obtain access, leading the parents to have access even if they do not frequently use the internet); whether the respondent or the respondent’s spouse runs a business from home (a home business has a greater need for home connection but not necessarily personal internet usage); whether the respondent telecommutes (which again likely increases the need for connection but not personal usage); whether the respondent brings work home (working from home might increase the need for a home connection but have no relation to personal usage); and the amount of hours spent online for work in the previous year (this might increase the propensity to adopt without altering personal usage propensity). In the online appendix, we also show that results are robust to the use of other instruments.

and entered respondents in a draw for a $500 prize. While we do not have information on the response rate to this particular survey, the general response rate for Forrester technology surveys is between 58% and 68%. The survey includes information on internet adoption, hours online for personal reasons, particular applications used, self-reported leisure time, and a number of demographic variables. Information from a similar survey of the same individuals in the previous year supplements this 2001 survey. Table 1 shows summary statistics for all the variables we use in this study. We list the exact survey questions in the online appendix.

Note that 74% of our sample has adopted.[2] This is higher than the estimate by the National Telecommunications and Information Administration (2002) because Forrester apparently over-sampled high-income individuals.[3] Of those who adopt, 97% use email, making it by far the most popular application. We define internet usage as "hours spent online for personal reasons."[4] The average household uses the internet 8.7 hours per week.

Figure 1 shows internet adoption and usage rates across demographic groups. Here we see that high-income, educated people were more likely to adopt the internet, but they also spend considerably less time online, conditional on adoption. In the econometric results that follow, we show that this general pattern holds, even when using a Heckman correction and including a number of control variables.

3. Results

3.1 Adoption vs. Usage

[2] We count individuals as adopters if, when asked about home connection, they give any response other than "I don't connect from home." An important caveat to this research is that we do not examine people who use the internet exclusively at work. It is possible that this group would mitigate the observed divide in adoption patterns. Our usage results, however, *are* based on using the internet for personal reasons, irrespective of location.

[3] More generally, our sample is slightly older, richer, and more educated than the general population. This survey asks about technology choices by households, leading Forrester to sample high-income households more heavily. When we weight the data to match national demographic distributions, the adoption rate is 62%. This is much nearer the true adoption rate, suggesting that there is not likely to be much selection on unobservables. Specifically, as long as the observed members of demographic groups are representative of their groups in the dimensions of interest, this will not affect our core conclusions.

[4] We present the options in five-hour intervals. In our analysis, we take the midpoint of each interval. If individuals claim 30 or more hours, we assign them a value of 35 hours. This data structure ensures that skewed usage patterns, where a small number of users spend an extraordinary amount of time online, do not unduly alter our results. In the online appendix, we show results with usage defined as "hours spent online from home."

Columns (1) and (2) of Table 2 show that usage and adoption follow very distinct patterns. These columns contain the results of a Type-II tobit regression where we define usage as "usage for personal reasons." The coefficient estimates in Column (2) show that internet adoption is increasing in income and education.[5] internet adoption is higher for younger people, married people, city dwellers, and whites. We find no significant difference in adoption rates by language spoken, gender, or number of children in the household. These results are consistent with previous studies of the digital divide (e.g., Hoffman and Novak, 2000).

This study differs by our ability to examine internet usage. Most strikingly, usage is decreasing in both income and education (Column (1)). Higher income and higher education relate to spending less time online, even with the Heckman correction and controls for leisure time. This is the paper's main result, which is consistent across numerous specifications and modeling techniques.[6]

One possible drawback of our data is that we observe usage only for the respondent but the internet adoption decision for the entire household. Therefore, even though we have controlled for the number of children in a household, concern may remain that a separation between the decision to adopt and the choice of usage level drives our results. To check for this, we run the model separately for one-person and two-person households (shown online in Appendix Table A4). For one-person households, we find identical results for income and university education; the only difference pertains to high school education, which changes sign but is not significant. For two-person households, the results are identical to those for the whole sample. Another potential concern is that internet usage depends on the amount of

[5] Note that graduation from college implies graduation from high school. Therefore, the total "educational impact" on usage for a university grad is the sum of the coefficients on high school and university degree.

[6] First, we run several regressions with various combinations of instruments. Additional instruments included are: moved in the past year, uses a computer at work, owns a cell phone, and Forrester's measure of optimism toward technology. We also include years since the household first used the internet in the second equation as a further check. All these regressions provide similar estimates to those in Columns (1) and (2). We show regressions with all instruments online in Appendix Table A1. We also show online regressions using "home usage" as the dependent variable in Table A2 and regressions for new adopters and small household sizes in Tables A2 and A3. In addition, we run the same regression on subsets of the population to ensure our results do not come from misrepresentation from some groups (e.g., white collar people defining work differently) or selection misspecification. We design subsets based on income, location, education, type of connection, and time since adoption. For virtually all subsets, the results for the remaining regressors are qualitatively unchanged. The only qualitative change is a loss of the significant effect of income when the subset is college graduates.

time a household has been online. For example, if higher-income households adopt earlier and usage declines over time, this could drive our results. We check for this by controlling for time since adoption in our second-stage regression (results are online in Table A1 of the Appendix). We find that usage increases with adoption tenure, and our main results persist even with this control in place.

Overall, the first two columns of Table 2 show that the digital divide exists. Rich, educated people are more likely to adopt. The results also show that if those demographically on the wrong side of the digital divide do adopt, then they spend more time online.

3.2 Why Do Usage and Adoption Patterns Differ?

In this subsection, we identify and empirically compare four main explanations for why we might observe a difference between usage and adoption patterns.

3.2.1 Four Explanations

The four explanations we consider are: 1) low-income people have a lower opportunity cost of leisure time, 2) low-income people find the internet to be more useful than others, 3) low-income people have more leisure time, and 4) the low-income people who choose to adopt the internet are those who place a particularly high value on it (i.e., selection). We discuss each in turn below.

Opportunity cost of leisure time: This setting has a unique pricing structure. First, adoption entails a fixed cost. Second, additional marginal use does not (effectively) incur a marginal monetary cost. Third, the only implicit cost of marginal use is the value of time, as in the standard Becker (e.g., Becker, 1965) model of time allocation. Such a setting has the following implications. First, the positive cost of adoption implies an income elasticity for adoption. Second, conditional on adoption, the implicit price of usage is higher for high-wage users. If high- and low-income groups receive the same benefit per hour of usage, then low-income groups will spend more time online if they have lower opportunity costs (i.e., $\frac{\partial u_2}{\partial L}$ is smaller for low-income groups than high-income groups for each value of L).[7] More generally, this suggests that when consuming a product takes time, researchers must consider the opportunity cost of time and multidimensional consumer types (Wilson, 1993 §8.4) in non-linear pricing strategies.

Usefulness of the internet: Different demographic groups may accrue different benefits from using the internet. In particular, low-income groups may get a particularly large benefit from usage because the internet provides services they cannot get elsewhere. Sinai and Waldfogel (2004) show that blacks who live in white neighborhoods are particularly likely to connect. They argue that this group receives a relatively large benefit from using the internet because they can access content not available locally. In the context of the present paper, if low-income individuals can access content online that is not locally available, they may spend a disproportionate amount of their leisure time online. Lambrecht and Seim (2006) also examine usage. They show suggestive evidence that high-income people derive a greater benefit from using online banking than low-income people. In particular, they find that high-income people have more online banking transactions than low-income people, conditional on adoption.

Quantity of leisure time: Even if low-income groups have the same opportunity costs of leisure time as high-income groups, they may use the internet more simply because they have more spare time. Specifically, suppose low-income and high-income groups have identical utility functions from usage (i.e., their $u_2(.)$ functions from Section 2 are the same). This means they have indistinguishable opportunity costs of leisure time. If low-income groups have more total leisure time (i.e., higher levels of

[7] Prior research finds evidence that the opportunity cost of time positively correlates with income. Calfee and Winston (1998) find that high-income people are willing to pay more to have their commuting time reduced than low-income people. Aguiar and Hurst (2005) find that people reduce food expenditures but not consumption in response to forecastable income changes. In particular, they increase the time spent preparing meals when income falls.

T) and we make very standard assumptions about the utility function from usage,[8] it follows that they will spend more time online.

Selection: Finally, it is possible that the observed difference between adoption and usage patterns is simply a matter of selection. In particular, those who do not adopt likely make that choice because they derive a lower net benefit from internet adoption. If an important barrier to adoption is cost, then most high-income people can afford to own a computer and pay for access. For low-income people, however, the computer purchase and internet access are significant expenses. Therefore, only those low-income people who place an especially high value on internet access will adopt. This may lead to those low-income people who adopt using the internet more.

3.2.2 Comparing the Explanations

In this subsection, we empirically evaluate each of the four explanations posited above. We begin by considering the possibility that selection is driving our result. In addition to showing that income and usage negatively correlate even with the Heckman correction, Table 2 provides further evidence that selection may not be driving the pattern in Figure 1. Columns (3) and (4) present results with no selection correction (i.e., regress usage on the covariates with no inverse Mills ratio). The results are qualitatively the same and perhaps slightly stronger in the selection-corrected model. This suggests that unobservable variables that make an individual more likely to adopt the internet negatively correlate with usage. Despite this suggestive evidence, we are unable to fully dispel concerns about selection.

Next, we examine differences in the amount of leisure time as a possible explanation for the observed demographic difference in usage and adoption. In our data, no substantial difference exists in measured leisure time between high- and low-income people. internet adopters in households with annual

[8] Specifically, if we assume $\frac{\partial u_2}{\partial I} > 0$, $\frac{\partial u_2}{\partial L} > 0$, $\frac{\partial^2 u_2}{\partial I^2} < 0$, and $\frac{\partial^2 u_2}{\partial L^2} < 0$, we hold that the optimal amount of internet usage is increasing in total leisure time.

income below $30,000 report that they have 21.94 hours of leisure time per week on average. Similarly, adopters in households with more than $100,000 report an average of 21.37 hours of leisure time per week. Table 2 provides further evidence against the idea that differences in leisure time are driving our results. Columns (1) and (2) show that the pattern in Figure 1 still holds when we control for stated amounts of leisure time. Also, in Columns (5) and (6), we present the results of the same model with leisure time excluded. The coefficients are almost identical, indicating that our measure of leisure time does not alter the relationship between internet adoption/usage patterns and income. While leisure time is a significant and economically important predictor of usage, it does not explain the differences among income groups.

Finally, we examine whether usage patterns for specific applications are consistent with either or both of the remaining two explanations (usefulness of the internet and opportunity cost of leisure time). Table 3 shows the correlations between demographics and usage of various internet applications, conditional on access (note that access differs from home adoption in that it allows for internet access from any location). In particular, it shows the coefficients of Heckman-corrected probit regressions of various application adoption dummy variables on demographics.[9] Table 4 (Columns (2) through (9)) then uses these coefficient estimates to predict the probability of using each of these applications for the entire sample (adopters and non-adopters) and also breaks these probabilities down by income. Implicitly, these results assume that the Heckman correction fully controls for selection.

Across income and education, Table 3 shows that the probability of using the internet for any of these applications is generally similar; however, interesting differences occur. Controlling for other demographic characteristics, low-income Americans are more likely than others to use the internet for chat, online games, and health information. They are less likely than high-income Americans to use the internet for e-commerce and researching purchases.

The fact that application usage varies according to income provides mild support for the idea that usefulness of the internet varies across demographic groups. However, this does not offer a direct

[9] We present the first stage of these regressions in our online Appendix Table A4.

measure of usefulness, and usage patterns are rather similar for many applications. While we cannot reject the possibility that low-income people find the internet more useful than others, we believe these findings lend greater support to the idea that differences in opportunity cost of leisure time are driving our main result. For example, low-income individuals are much more likely to use the internet for gaming and chat, two relatively inexpensive and often time-consuming internet applications.

4. Policy Implications

In addition to trying to understand the pattern observed in Figure 1, our second contribution is to provide a better understanding of the effect of subsidizing home internet access. Column (1) of Table 4 contains predicted usage for the entire sample and breaks this down by income. The results show that predicted usage among low-income individuals would be high, even higher than their counterparts, and Columns (2) through (9) illustrate that application usage often follows patterns similar to those of high-income individuals. In particular, these findings suggest that a subsidy for internet use would not be wasted. Individuals who have not yet adopted (and who are primarily low-income) would use the internet intensely if given access.

A potential worry is that the relevant benefit of using the internet may be concave. This would imply that the high usage observed in low-income households does not reduce the welfare implications related to the digital divide. We address this question in Table 5. To construct this table, we run a series of probits relating application adoption to time spent online. The estimates we report are the expected changes in the probability of using each application with changes in the amount of time spent online. Our results suggest that the benefits of using the internet are not likely to be concave. At least up to 17 hours per week,[10] increases in hours using the internet relate to significant increases in the use of many valuable online activities, including e-government, researching purchases, telemedicine, and online news.

The National Telecommunications and Information Administration (2002) emphasized online health information and e-government as benefits of the internet in addition to general information, online

[10] Note that 88% of those using the internet in our sample used it for 17 hours or less per week.

commerce, and entertainment. Revisiting Columns (2) through (9) of Table 4, in support of the goals of the NTIA, the simulations suggest that at least half of low-income non-adopters would use the internet for email, researching purchases, e-commerce, health information, and news. Another 46% of low-income individuals would use e-government. On the other hand, many low-income Americans would also be particularly likely to use the internet for chat and online games if given access. This may suggest an argument against subsidies to the extent that it is undesirable to support such activities.

Scholars should interpret the results of this section as suggestive of a subsidy's impact. Ideally, we would have a natural experiment where we could randomly assign subsidies and see the result. In the absence of such an experiment, we rely on the Heckman correction to understand differences between adopters and non-adopters. Furthermore, the simulations do not reflect an equilibrium outcome. Nevertheless, we believe the simulations help elucidate the impact of access subsidies on usage. They suggest that subsidizing internet access to low-income and less-educated Americans would likely achieve many of the goals stated by policymakers, although (perhaps) unintended consequences would also occur.

5. Conclusions

We show that internet adoption and usage follow different patterns. While income and education positively correlate with adoption, they negatively correlate with hours spent online. Given our results, we argue that the most likely explanation for this finding is that low-income individuals spend more time online due to their lower opportunity costs of leisure time. In particular, the pricing structure of the internet, with both fixed connection and near-zero usage fees, leads to a negative correlation between income and time online among those who have connected. We interpret the fact that low-income people are particularly likely to do time-consuming, inexpensive activities online as support for the role of the opportunity cost of leisure time.

Our results also provide a better understanding of access subsidies and the digital divide. If given the opportunity to go online, Americans on the wrong side of the digital divide would likely use the internet a great deal and engage in many of the online activities policymakers have stated as the goals of access subsidies. While this prediction does not necessarily mean that access subsidies are a worthwhile policy (that depends on a full cost/benefit analysis and on any perceived negative benefit of subsidizing activities like online gaming), it does suggest that some important benefits will ensue from such subsidies.

Our study has a few limitations. First, while we control for selection to the extent possible, we cannot entirely reject the possibility that selection drives usage and adoption's differing relationships with income and education. Therefore, we interpret our results as suggestive that the pricing structure of internet access has led to different adoption and usage patterns across demographics. Second, our data (from 2001) are old by internet standards. It is possible that the conclusions based mainly on dial-up connections do not apply to today's internet.[11] Third, in considering this study's implications, it is important to remember that not all users want to adopt the internet. Fox (2005) presents the results of interviews with non-adopters. Many non-adopters do not want to be online. While 5% say that the internet is too expensive, 32% say they "are just not interested" (p. 3). The perceived benefit of the online experience matters in addition to the cost of adoption. If people do not want to go online, then any subsidies would be wasted.

Despite these limitations, we believe our study provides new insight into the nature of the digital divide. The difference in adoption and usage patterns likely depends on the pricing structure. If users are charged per minute, then the pattern likely would be different. For example, in Europe, local calls are typically tolled, meaning dial-up access has a per-minute charge. Mann (2000) argues that this leads to lower overall usage rates in Europe than in the US.

Furthermore, our results suggest the demographic implications of pricing structures with fixed connection fees and free unlimited usage. For example, researchers have shown that television viewing is

[11] Still, we believe that the implications of our results are relevant today. A digital divide remains regarding adoption of internet technology in general and of broadband in particular. In early 2006, 73% of Americans were online, and 62% of these adopters had broadband access. Even by this time, adoption rates varied substantially by income and education. Only 53% of Americans with household income under $30,000 were online, while 91% of households earning $75,000 or more were connected (Source: Pew Internet & American Life Project: http://www.pewinternet.org/trends, visited December 4, 2006.) The differences for broadband adoption across income groups were even larger. While broadband differs in many ways from dial-up, the fundamental pricing structure is unchanged. Further, the incentives of low-income people to spend a lot of time online (conditional on adoption) also are likely unchanged.

negatively correlated with education (Waldfogel, 2002) and income (Hughes, 1980). A low opportunity cost of time for low-income people may be driving these results in the same way we believe it is driving ours.

References

Aguiar, M.A., Hurst, E., 2005. Consumption vs Expenditure. Journal of Political Economy 113, pp. 919-948.

Astebro, T., 2004. Sunk Costs and the Depth and Probability of Technology Adoption. Journal of Industrial Economics 52, pp. 381-399.

Becker, G.S., 1965. A Theory of the Allocation of Time. The Economic Journal 75, pp. 493-517.

Calfee, J., Winston, C., 1998. The Value of Automobile Travel Time: Implications for Congestion Policy. Journal of Public Economics 69, pp. 83–102.

Chinn, M., Fairlie, R., 2006. The Determinants of the Global Digital Divide: A Cross-Country Analysis of Computer and Internet Penetration. Oxford Economic Papers 59, pp. 16-44.

Devaraj, S., Kohli, R., 2003. Performance Impacts of Information Technology: Is Actual Usage the Missing Link? Management Science 49, pp. 273-289.

Fairlie, R., 2004. Race and the Digital Divide. Contributions to Economic Analysis and Policy 3(1), article 15.

Fox, S., 2005. Digital Divisions. Working Paper, Pew Internet and American Life Project.

Goldfarb, A., 2006. The (Teaching) Role of Universities in the Diffusion of the Internet. International Journal of Industrial Organization 24, pp. 203-225.

Greene, W.H., 1997. Econometric Analysis, 3rd editon. 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., Kahin, B. (Eds.), Understanding the Digital Economy: Data, Tools, Research. The MIT Press, Cambridge MA, pp. 245-260.

Hughes, M., 1980. The Fruits of Cultivation Analysis: A Reexamination of Some Effects of Television Watching. The Public Opinion Quarterly 44, pp. 287-302.

Lambrecht, A., Seim, K., 2006. Adoption and Usage of Online Services in the Presence of Complementary Offline Services: Retail Banking. Working Paper 06-27, NET Institute, October.

National Telecommunications and Information Administration (NTIA)., 2002. A Nation Online: How Americans Are Expanding Their Use Of The Internet. http://www.ntia.doc.gov/reports.html.

Sinai, T., Waldfogel, J., 2004. Geography and the Internet: is the internet a substitute or a complement for cities? Journal of Urban Economics 56, pp. 1-24.

Waldfogel, J., 2002. Consumer Substitution among Media, FCC Media Ownership Working Group.

Wilson, R.B., 1993. Nonlinear Pricing. Oxford University Press, New York.

Zhu, K., Kraemer, K.L., 2005. Post-Adoption Variations in Usage and Value of E-Business by Organizations: Cross-Country Evidence from the Retail Industry. Information Systems Research 16, pp. 61-84.

Table 1: Descriptive statistics

Variable	# of observations	Mean	Standard Deviation	Minimum	Maximum
Home usage for adopters (hours/wk.)	14,310	8.725	9.046	0	35
Personal usage for adopters (hours/wk.)	14,453	8.654	8.749	2	35
Internet adopted at home	18,439	0.738	0.440	0	1
Access to internet anywhere	18,439	0.823	0.381	0	1
Personal income	18,439	68,392	51,202	5,000	350,000
High school graduate	18,439	0.918	0.274	0	1
University/college graduate	18,439	0.458	0.498	0	1
Married	18,439	0.736	0.441	0	1
White	18,439	0.905	0.293	0	1
Age	18,439	52.301	13.894	18	99
Female	18,439	0.508	0.500	0	1
English is primary language	18,439	0.977	0.151	0	1
In city with less than 100,000 people	18,439	0.184	0.388	0	1
In city with 100,000–499,999 people	18,439	0.143	0.350	0	1
In city with 500,000–1,999,999 people	18,439	0.203	0.402	0	1
In city with 2,000,000 or more people	18,439	0.470	0.499	0	1
Number of children in household	18,439	0.559	0.921	0	3
Leisure time (By five hour group)	18,439	4.013	2.049	0	7
Use for email	15,035	0.924	0.265	0	1
Use for chat	14,095	0.233	0.423	0	1
Use for online games	13,998	0.135	0.342	0	1
Use for researching purchases	14,377	0.659	0.474	0	1
Use for e-commerce	15,170	0.651	0.477	0	1
Use for health information	14,217	0.486	0.500	0	1
Use for news	14,550	0.477	0.499	0	1
Use for e-government	14,254	0.399	0.490	0	1
Primary Instruments					
Teen in the home	18,439	0.151	0.358	0	1
Operates a business from home	18,439	0.137	0.344	0	1
Telecommutes	18,439	0.0400	0.196	0	1
Brings work home (in 2001)	18,439	0.240	0.427	0	1
Brings work home (in 2000)	18,439	0.204	0.403	0	1
Work usage (in 2000) (by five hour group)	18,439	1.180	1.438	0	7
Secondary Instruments					
Moved in past year	18,439	0.0527	0.223	0	1
Has a computer at work	17,922	0.568	0.495	0	1
Has a cell phone	18,411	0.620	0.485	0	1
Years since first used the internet	18,439	3.901	2.578	0	7
Measure of optimism toward technology	18,336	1.543	0.498	1	2

Table 2: Coefficients of internet adoption and Heckman-corrected usage (in hours)

	Heckman- Usage defined by hours online for personal reasons		Non-selection results		No control for leisure time	
	(1)	(2)	(3)	(4)	(5)	(6)
Covariates	Personal usage	Home adoption	Personal usage	Home adoption	Personal usage	Home adoption
Income ($0,000)	-0.071	0.014	-0.046	0.015	-0.065	0.014
	(0.017)**	(0.003)**	(0.016)**	(0.003)**	(0.017)**	(0.003)**
High school graduate	-1.677	0.673	-0.621	0.615	-1.653	0.67
	(0.395)**	(0.039)**	(0.371)+	(0.037)**	(0.399)**	(0.039)**
University/college graduate	-1.014	0.135	-0.875	0.129	-1.02	0.134
	(0.187)**	(0.029)**	(0.178)**	(0.029)**	(0.190)**	(0.029)**
Married	-2.175	0.292	-1.466	0.256	-2.403	0.294
	(0.190)**	(0.026)**	(0.177)**	(0.025)**	(0.192)**	(0.026)**
White	-0.198	0.473	0.277	0.401	-0.004	0.469
	(0.302)	(0.036)**	(0.280)	(0.035)**	(0.305)	(0.036)**
Age	-0.051	-0.014	-0.056	-0.011	-0.039	-0.014
	(0.008)**	(0.001)**	(0.007)**	(0.001)**	(0.008)**	(0.001)**
Female	-1.728	0.016	-1.634	0.021	-2.228	0.026
	(0.161)**	(0.024)	(0.153)**	(0.024)	(0.160)**	(0.024)
English is primary language	-0.981	-0.031	-1.048	-0.055	-0.82	-0.036
	(0.497)*	(0.072)	(0.476)*	(0.069)	(0.504)	(0.072)
In city with 100,000 to 499,999 people	0.396	0.123	0.583	0.120	0.399	0.123
	(0.265)	(0.038)**	(0.253)*	(0.037)**	(0.269)	(0.038)**
In city with 500,000 to 1,999,999 people	-0.027	0.132	0.190	0.131	-0.038	0.132
	(0.245)	(0.035)**	(0.233)	(0.034)**	(0.248)	(0.035)**
In city with over 2 million people	-0.336	0.114	-0.144	0.115	-0.328	0.113
	(0.215)	(0.030)**	(0.204)	(0.030)**	(0.218)	(0.030)**
# of children in household	-0.616	-0.023	-0.554	-0.020	-0.876	-0.018
	(0.091)**	(0.017)	(0.087)**	(0.017)	(0.091)**	(0.017)
Leisure time	0.696	0.035	0.688	0.025		
	(0.040)**	(0.006)**	(0.038)**	(0.006)**		
Teen in the home		0.165		0.158		0.163
		(0.039)**		(0.039)**		(0.039)**
Operates a business from home		0.284		0.262		0.282
		(0.036)**		(0.035)**		(0.036)**
Brings work home (in 2000)		0.021		0.037		0.018
		(0.037)		(0.036)		(0.037)
Brings work home (in 2001)		0.126		0.122		0.125
		(0.035)**		(0.035)**		(0.035)**
Telecommutes		0.182		0.162		0.181
		(0.070)**		(0.070)*		(0.070)**
Work usage (in 2000)		0.355		0.331		0.355
		(0.012)**		(0.012)**		(0.012)**
ρ		-0.241				-0.273
		(0.024)**				(0.023)**
σ		8.579				8.707
		(0.056)**				(0.058)**
λ		-2.070				-2.378
		(.213)**				(0.207)**
# of observations		18,439	18,439	18,439		18,439
Log likelihood		-56,931.6	-51,451.9	-9,007.5		-57,081.4

All regressions include occupation fixed effects and a constant. Standard errors in parentheses.
+ significant at 10%; * significant at 5%; ** significant at 1%

Table 3: Heckman-corrected probit coefficients of application adoption conditional on internet adoption (first stage in online Appendix)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Email	Chat	Online games	Research purchases	E- commerce	Health information (telemedicine)	News	E-government
Income ($0,000)	-0.002	-0.009	-0.015	0.011	0.014	-0.005	0.002	0.002
	(0.004)	(0.003)**	(0.004)**	(0.003)**	(0.003)**	(0.002)*	(0.002)	(0.002)
High school graduate	0.089	-0.186	-0.155	0.085	0.002	0.006	-0.035	-0.106
	(0.058)	(0.058)**	(0.066)*	(0.052)	(0.051)	(0.054)	(0.054)	(0.054)+
University/college graduate	0.218	-0.12	-0.269	0.168	0.168	0.035	0.185	0.119
	(0.040)**	(0.030)**	(0.035)**	(0.027)**	(0.027)**	(0.026)	(0.027)**	(0.027)**
Married	-0.100	-0.298	-0.128	-0.098	-0.077	-0.001	-0.163	-0.195
	(0.039)**	(0.028)**	(0.034)**	(0.027)**	(0.026)**	(0.026)	(0.026)**	(0.026)**
White	0.161	-0.052	-0.111	0.147	0.164	0.007	-0.200	-0.083
	(0.054)**	(0.046)	(0.052)*	(0.041)**	(0.040)**	(0.041)	(0.041)**	(0.041)*
Age	-0.006	-0.014	-0.011	-0.012	-0.018	0.003	-0.001	0.001
	(0.002)**	(0.001)**	(0.001)**	(0.001)**	(0.001)**	(0.001)**	(0.001)	(0.001)
Female	0.053	-0.111	-0.020	-0.134	-0.083	0.254	-0.172	-0.196
	(0.034)	(0.025)**	(0.030)	(0.023)**	(0.023)**	(0.022)**	(0.022)**	(0.023)**
English is primary language	0.024	-0.215	-0.083	0.045	0.016	-0.053	0.006	-0.082
	(0.097)	(0.075)**	(0.090)	(0.071)	(0.069)	(0.069)	(0.069)	(0.070)
In city with 100,000 to 499,999 people	0.043	0.129	0.127	0.057	0.084	-0.008	-0.034	0.081
	(0.054)	(0.041)**	(0.047)**	(0.038)	(0.037)*	(0.037)	(0.037)	(0.038)*
In city with 500,000 to 1,999,999 people	-0.059	0.062	0.069	0.009	0.065	-0.039	-0.093	0.06
	(0.049)	(0.038)	(0.044)	(0.035)	(0.034)+	(0.034)	(0.034)**	(0.034)+
In city with over 2 million people	-0.056	-0.006	0.003	-0.011	0.126	-0.061	-0.114	0.045
	(0.043)	(0.034)	(0.039)	(0.031)	(0.030)**	(0.030)*	(0.030)**	(0.030)
# of children in household	-0.039	-0.026	0.029	-0.054	-0.027	-0.04	-0.038	-0.081
	(0.019)*	(0.014)+	(0.016)+	(0.013)**	(0.013)*	(0.013)**	(0.013)**	(0.013)**
Leisure time	0.008	0.024	0.019	0.034	0.025	0.018	0.014	0.026
	(0.008)	(0.006)**	(0.007)**	(0.006)**	(0.006)**	(0.006)**	(0.006)*	(0.006)**
ρ	-0.786	-0.522	-0.260	-0.690	-0.715	-0.661	-0.526	-0.687
	(0.0611)**	(0.044)**	(0.0638)**	(0.0348)**	(0.0317)**	(0.0442)**	(0.0449)**	(0.0326)**
# of observations	18,308	17,527	17,433	17,816	18,439	17,649	18,006	17,686
Log likelihood	-9,433.7	-13,046.2	-11,001.7	-14,337.7	-14,722.8	-15,362.2	-15,604.8	-14,861.1

All regressions include occupation fixed effects and a constant. Standard errors in parentheses.
+ significant at 10%; * significant at 5%; ** significant at 1%

Table 4: Predicted usage rates (based on Table 3 regressions)

	(1) Predicted usage rates	(2) Email	(3) Chat	(4) Online games	(5) Research purchases	(6) E-commerce	(7) Health information (telemedicine)	(8) News	(9) E-government
All participants	10.04 (0.0207)	0.935 (0.00028)	0.280 (0.00068)	0.156 (0.000511)	0.688 (0.000687)	0.681 (0.000808)	0.544 (0.000536)	0.520 (0.000565)	0.456 (0.00062)
Non-adopters only	11.02 (0.0413)	0.915 (0.00066)	0.314 (0.00188)	0.187 (0.00137)	0.617 (0.00162)	0.595 (0.00188)	0.568 (0.00134)	0.518 (0.00136)	0.454 (0.00147)
Adopters only (not predicted—actual usage rate)	8.65 (0.0728)	0.924 (0.0022)	0.233 (0.00356)	0.134 (0.00288)	0.658 (0.00395)	0.651 (0.00387)	0.485 (0.00419)	0.477 (0.00414)	0.397 (0.00410)
By income									
Less than $25,000	12.39 (0.0431)	0.921 (0.00055)	0.359 (0.00175)	0.217 (0.00126)	0.620 (0.00151)	0.597 (0.00175)	0.580 (0.00126)	0.521 (0.00121)	0.455 (0.00135)
$25,000 – $50,000	10.93 (0.0398)	0.920 (0.00066)	0.306 (0.00136)	0.184 (0.000998)	0.645 (0.00138)	0.632 (0.00163)	0.550 (0.00127)	0.502 (0.00135)	0.435 (0.00145)
$50,000 – $75,000	9.59 (0.0396)	0.935 (0.00066)	0.276 (0.00115)	0.155 (0.000824)	0.696 (0.00133)	0.697 (0.00156)	0.532 (0.00113)	0.512 (0.00130)	0.443 (0.00138)
$75,000 – $100,000	9.03 (0.0377)	0.946 (0.00056)	0.244 (0.00106)	0.125 (0.00071)	0.722 (0.00118)	0.721 (0.00134)	0.533 (0.00106)	0.528 (0.00120)	0.465 (0.00131)
More than $100,000	8.23 (0.0374)	0.953 (0.00041)	0.213 (0.00101)	0.0949 (0.000593)	0.757 (0.00104)	0.761 (0.00121)	0.526 (0.00102)	0.543 (0.00109)	0.486 (0.00124)
# of observations	18,439	18,439	18,439	18,439	18,439	18,439	18,439	18,439	18,439

Standard errors in parentheses

Table 5: Change in probability of application adoption with changes in total internet usage

	(1) Email	(2) Chat	(3) Online games	(4) Research purchases	(5) E- commerce	(6) Health information (telemedicine)	(7) News	(8) E- government
Usage = 7 hours per week	0.028	0.167	0.068	0.158	0.134	0.143	0.152	0.154
	(0.002)**	(0.011)**	(0.009)**	(0.008)**	(0.009)**	(0.011)**	(0.010)**	(0.011)**
Usage = 12 hours per week	0.022	0.235	0.132	0.153	0.139	0.197	0.178	0.179
	(0.002)**	(0.014)**	(0.012)**	(0.010)**	(0.010)**	(0.012)**	(0.012)**	(0.013)**
Usage = 17 hours per week	0.019	0.318	0.191	0.18	0.17	0.199	0.21	0.234
	(0.002)**	(0.019)**	(0.019)**	(0.012)**	(0.012)**	(0.017)**	(0.016)**	(0.018)**
Usage = 22 hours per week	0.019	0.353	0.216	0.177	0.169	0.195	0.225	0.218
	(0.002)**	(0.020)**	(0.021)**	(0.013)**	(0.013)**	(0.018)**	(0.018)**	(0.020)**
Usage = 27 hours per week	0.018	0.357	0.302	0.173	0.191	0.236	0.202	0.238
	(0.002)**	(0.030)**	(0.031)**	(0.019)**	(0.018)**	(0.026)**	(0.027)**	(0.029)**
Usage = 35 hours per week	0.02	0.478	0.306	0.168	0.16	0.206	0.224	0.247
	(0.002)**	(0.018)**	(0.021)**	(0.013)**	(0.014)**	(0.018)**	(0.017)**	(0.019)**
# of observations	14,324	13,392	13,295	13,677	14,453	13,518	13,837	13,550
Log likelihood	-1,710.6	-6,894.2	-5,139.5	-8,163.3	-8,791.5	-9,134.3	-9,346.7	-8,961.3

Base group is “Usage = 2 hours per week.” Standard errors in parentheses.
+ significant at 10%; * significant at 5%; ** significant at 1%

Figure 1: Internet adoption and usage