Communications of the Association for Information Systems

Volume 28 Article 7

2-2011

Home Internet Access and Usage in the USA: Trends in the Socio-Economic Digital Divide

Debabrata Talukdar
School of Management, SUNY Buffalo, dtalukda@buffalo.edu

Dinesh K. Gauri
Whitman School of Management, Syracuse University

Communications of the Association for Information Systems

CAIS

Home Internet Access and Usage in the USA: Trends in the Socio-Economic Digital Divide

Debabrata Talukdar

School of Management, SUNY Buffalo

dtalukda@buffalo.edu

Dinesh K. Gauri

Whitman School of Management, Syracuse University

Abstract:

Past research has found evidence of the existence of a significant digital divide among various socio-economic segments in terms of Internet access and usage. In this research, we use two annual (2002 and 2008) cohorts of demographically representative national random samples to investigate the relative levels of digital divide along the typical socio-economic fault lines. We find from our statistical analyses that not only does a deep digital divide still persist along key dimensions (like education) of the socio-economic fault lines, it has in fact widened along several key dimensions (like income and urban-rural divide).

Keywords: Internet, digital divide, public policy issues, logistic regression

Volume 28, Article 7, pp. 85-98, February 2011

The manuscript was received 9/14/2009 and was with the authors 6 months for 2 revisions.

Communications of the Association for Information Systems

I. INTRODUCTION

The early 1990s saw the introduction of the Internet in the public domain. By the end of that decade, the Internet has become an "indispensable" part in the daily lives for many people in the USA [Hoffman et al., 2004]. It continues to fundamentally alter the key facets of the society in most countries [Chinn and Fairliey, 2006]. With an ever increasing operational reliance on the Internet by society in general, a particularly important research and public policy issue has centered on the notion of the "digital divide" [Compaine, 2001; Dewan and Riggins, 2005; Norris, 2001].

The term *digital divide* was originally coined by the *National Telecommunication and Information Administration* (NTIA) to indicate the gap between societal haves and have-nots in their access to the powerful new information technologies, particularly the Internet [Buente and Robbin, 2008; Robinson et al., 2003]. In this respect, one view holds that access to the Internet should be treated as "a basic human right," warranting government involvement and public policy initiatives from social equity perspective. For instance, a recent survey done by the *British Broadcasting Corporation* [2010] finds that almost four in five people around the world believe that access to the Internet is "a fundamental right." An alternative view holds that the Internet generates positive externalities for the society and potential adopters might not be fully aware of its benefits, which would also warrant government involvement and public policy initiatives from economic efficiency perspective [Robinson et al., 2003].

The general importance and policy relevance of the digital divide issue have led several studies to investigate Internet adoption patterns among distinct socio-economic segments of the U.S. population [Bannerji et al., 2005; Buente and Robbin, 2008; Compaine, 2001; Lorence et al., 2006]. These studies have typically found evidence of significant digital divide among such segments in terms of Internet adoption patterns. At the same time, the overall Internet penetration at the aggregate population level in the U.S. has increased steadily during the first decade of the twenty-first century. For instance, the percentage of U.S. adults using the Internet has reached about 75 percent by late 2009 from about 48 percent in 2000 [Pew Internet and American Life Project, 2010]. That raises the following important questions.

Has the increased Internet penetration at the aggregate population level in the U.S. changed the relative levels of digital divide along typical socio-economic fault lines? If so, such change is with respect to which socio-economic variables and to what extent? However, as we discuss in more detail in the next section, very limited research exists to provide the answer that is statistically rigorous or current-time relevant., The goal of our study is to provide systematic empirical insights into the aforesaid question, using two recent annual (2002 and 2008) cohorts of demographically representative national random samples. We investigate the nature of the digital divide in terms of both Internet access and its usage level when access is gained. The findings from our study are of considerable importance to anyone, especially computing professionals, interested in understanding the open challenges posed by the digital divide to the societal diffusion of one of the most heralded IT innovations. Such understanding, in turn, is imperative from both economic efficiency and social equity perspectives.

The economic efficiency relevance of understanding the challenges posed by the digital divide is quite apparent from the ever expanding operational role of the Internet in the market place. According to the U.S. Department of Commerce, the total e-commerce sales in the U.S. for 2009 were $134.9 billion, an increase of 4.4 percent over 2008 even in the midst of a recession [Grau, 2010]. Recent trends also suggest that it will grow at about 9 percent compound annual rate from 2009 to 2014 [Grau, 2010]. There is hardly any business firm today which does not depend on the Internet as a transaction and communication medium to sell its products and to target its customers with product information. Our research is directly relevant to their strategic goal of matching media with target markets by evaluating the relative communication effectiveness of alternative media across different target market segments [Ratchford et al., 2003]. In particular, our research offers strategic market insights to businesses that enable them to understand how the effectiveness of the Internet in reaching any target segment will be hindered or facilitated by key socio-economic traits of the segment [Bannerji et al., 2005]. Such insights are especially relevant as businesses debate how best to allocate billions of dollars between the Internet and other conventional communication and sales channels [Baker and Hempel, 2006].

The social equity relevance of our research findings stems from an increasingly important societal goal to achieve an effective and equitable "cyber society," where the Internet serves as a conduit for the provision of a myriad of civic services as well as a tool to stimulate a more engaged and participatory civic community [Jensen et al., 2007; Sims

et al., 2008; Sipior et al., 2004]. In fact, the U.S. Congress has passed various legislations that seeks to provide federal financial assistance in the form of grants and tax credits to encourage infrastructure investments in high speed Internet service, especially in rural areas [Fox, 2005; Robinson et al., 2003]. Naturally, from that public policy perspective, insights into changing trends in Internet access and usage level across important socio-economic segments of society are of significant value. Such insights are critical in order to evaluate existing policy initiatives and to prioritize future initiatives to address the Internet based digital divide [Robinson et al., 2003]. They are also critical to understanding how one of the economic benefits that the Internet provides (by enabling consumers to find lower prices for products purchased) is likely to be enjoyed by those segments in our society who probably need it the most [Baye et al., 2003].

While the Internet penetration level *within* the aggregate population in the U.S. has more than doubled between 1998–2009 [Pew Internet and American Life Project, 2010], such increase in itself sheds little insight into the underlying trends in penetration levels *across* distinct socio-economic segments. It is analogous to the fact than an increase in the national average of household income says very little about its effect on the income distribution across households. For instance, an increase in the Internet penetration at the aggregate level can as well be just an upshot of more people in the same socio-economic groups enjoying Internet access—thus leaving the digital divide unchanged or even widening. In other words, whether an increase in the aggregate penetration level is indeed leading to a systematic decline in relative digital divides across distinct socio-economic segments remains an open and important empirical question, unless of course the aggregate penetration reaches nearly 100 percent. However, the recent data (early 2010) shows a plateau in the Internet penetration growth trend in the U.S. since 2006 at a level of about 75 percent [Pew Internet and American Life Project, 2010].

In the next section of this article, we briefly discuss the relevant existing literature to which our study contributes. In Section III, we present our conceptual framework and the empirical model. Section IV discusses the data used and the findings from our empirical analysis, while Section V presents conclusions.

II. RELEVANT EXISTING RESEARCH CONTEXT

As the Internet has assumed an increasingly "indispensable" role in modern life [Hoffman et al., 2004], a natural and important issue that researchers continue to focus on is an in-depth understanding of the factors that facilitate or hinder the Internet adoption process through the general population [Dewan and Riggins, 2005; Ratchford et al., 2003; Venkatesh and Brown, 2001]. The analysis approach used in this research endeavor is essentially the "positive economics" methodology in order to gain systematic insights into "what is" the empirical reality for the research question being investigated. Instead of using any explicit theoretical model, almost all such studies use some form of conceptual framework to guide their empirical analysis. However, a notable exception is the study by Ratchford et al. [2003] which develops and empirically tests an explicit theoretical model of Internet adoption and usage. The model is based on expected differences in relative cost-benefits among consumers to search for information across alternative sources, including the Internet.

In terms of the types of data used in their empirical analyses, the aforesaid research studies used either aggregate or individual level data. Those using aggregate level data focus primarily on analyzing the Internet-based digital divide across countries. Specifically, they focus on aggregate country level diffusion process for Internet adoption and investigate the determinants of cross-country differences in Internet penetration levels [Chinn and Fairliey, 2006; Cuervo and Menendez, 2006]. For instance, the study by Chinn and Fairliey [2006] analyzes the factors that explain the variation in Internet penetration levels across 161 countries. It finds that income per capita, average years of schooling, telephone penetration level, and urbanization rates are some of the key factors in explaining cross-country differences in Internet penetration levels. For European Union member countries, Cuervo and Menendez [2006], find that information and communication technology infrastructure and Internet access costs were the two main factors in explaining the digital divide across member countries.

As for those studies using individual level data, the goal has been to investigate the role of various determinants of an individual's likelihood of adoption of the Internet within a given country—either in the general context [Goldfarb and Prince, 2007] or in specific contexts, such as a search for information on new cars or on health care [Lorence, et al 2006; Ratchford et al., 2003]. While the focus of some of these studies is on the general determinants of Internet use [Ratchford, et al. 2003], it is on specific determinants from the digital divide perspective for other studies [Bannerji et al., 2005; Buente and Robbin, 2008; Goldfarb and Prince, 2007; Lenhart et al., 2003]. Our current study is most directly related to this latter stream of research within the aforementioned broader group of research studies. While these existing studies have definitely advanced our understanding of the important issue of the Internet-based digital divide in the U.S., they also point to the need for additional research to address some of their weaknesses. One weakness: Many of the existing studies are based on univariate rather than multivariate statistical analysis [Fox 2005; Lenhart et al., 2003; McGregor, 1999]. As such, the findings from such studies provide statistically weak

evidence on the presence of the digital divide and thus fuel the debate about the true extent of the digital divide [Compaine, 2001].

There are also very few studies that use a multivariate statistical approach as well as repeated national sampling data over multiple time periods. However, such analysis approach and data are critical to gaining statistically reliable insights into the following key issue: Whether and how the increased Internet penetration at the aggregate population level is altering the relative levels of the digital divide along typical socio-economic fault lines in the U.S. The goal and contribution of our study is to use such approach and data to augment this very limited existing research stream by addressing the above key question on the digital divide issue. Our study also extends the insights from this limited research stream by investigating over a more recent time period, viz., 2002–2008.

III. CONCEPTUAL FRAMEWORK AND EMPIRICAL MODEL

As is typical of studies in this research area, the analysis approach in our study also follows the "positive economics" methodology to gain systematic empirical insights into the research questions that we investigate. In other words, the focus of our study is on the "positive" or "what is" aspects, rather than on the "normative" or "what should be" aspects, of the digital divide issue. At the same time, for the policy deliberations on the normative aspects of the digital divide issue to be effective, they need to be based on systematic insights from positive analyses. So the insights or "outputs" of our study offer part of the complementary "inputs" necessary for any policy deliberations dealing with the normative aspects of the digital divide. Figure 1 provides a conceptual framework for our study. The focus in our study is on investigating the digital divide in terms of Internet adoption patterns of people at their own homes, since it remains less understood than adoptions patterns at public locations like schools, libraries, and workplaces [Compaine, 2001, p. 88].

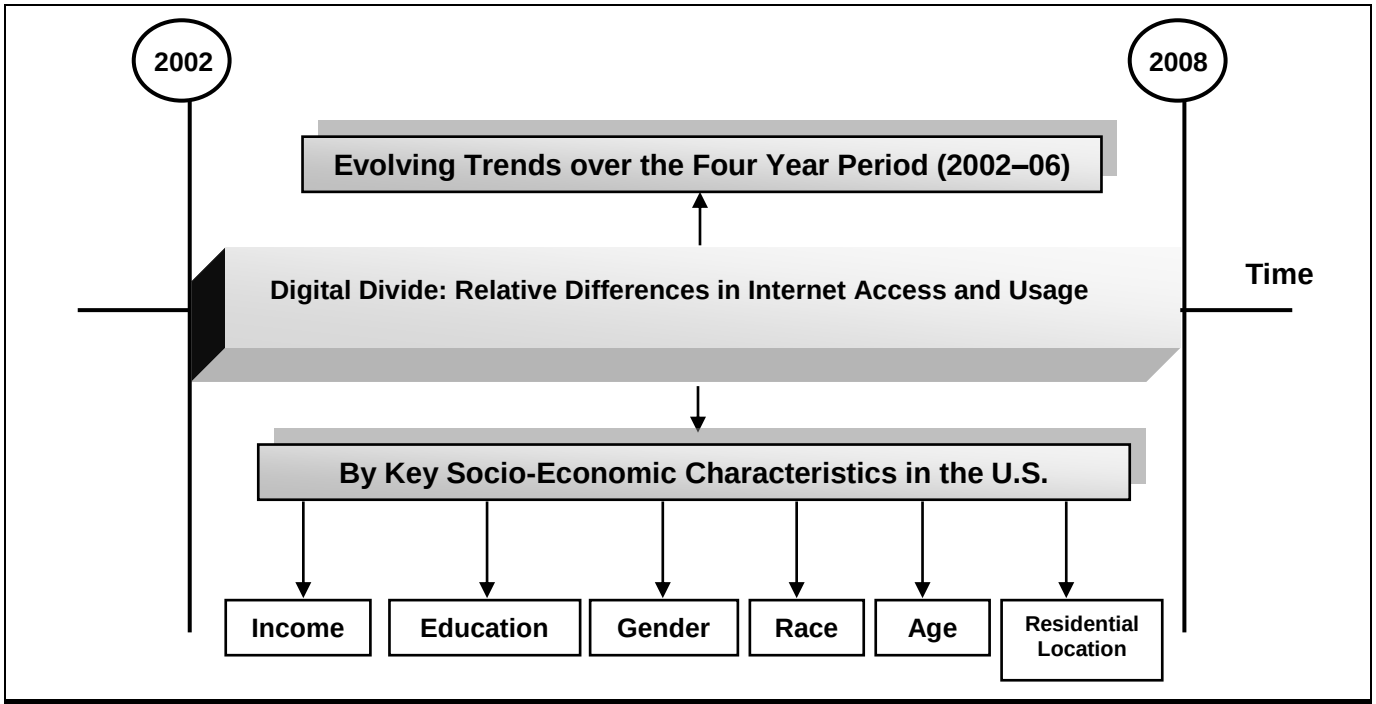

Figure 1. Conceptual Framework

As Goldfarb and Prince [2007] note, an understanding of the Internet-based digital divide across individuals requires analysis for differences in both Internet access and usage patterns. The idea is to recognize the distinction between access per se and "effective access" [Buente and Robbin, 2008; Robinson et al., 2003]. In other words, enjoying the benefits from Internet adoption depends not only on having the means to access it *per se*, but also on the inclination and ability to use it on a regular basis to make such access effective. So, a fuller understanding of the Internet based digital divide in the society requires investigation along both dimensions of the digital divide—Internet access and usage level.

In this context, it is pertinent to recognize that Internet-usage results are naturally expected to have a clear relationship to Internet access results—after all, "no access" implies "no usage." At the same time, a comparison of the level of "gaps" or "divides" in Internet access versus usage stands to shed important insights. For instance, if there is a big gap in access but not in usage, that would indirectly imply that the "Internet have-not" group in fact

consists mostly of heavy users once they do get access. In other words, lack of Internet access would then be more reflective of economic or infrastructure constraints than of any lack of inclination or ability to use once access is secured.

The specific empirical estimation model used in our study is the binary logistic regression model which can be specified as [Allison, 1999]:

$$Pr\ (Y = 1) = F(\beta X_i)$$

where $F(\beta X_i) = e^{\beta X_i} / (1 + e^{\beta X_i})$ is the cumulative logistic distribution. The term "Pr (Y = 1)" denotes the probability of observing a "Yes" outcome ("1" versus a "No" outcome of "0") for the dependent variable of interest, *Y*. The vector of explanatory variables for each respondent *i* is denoted by X_i, while β denotes the vector of corresponding regression coefficients. A positive (negative) value for the estimated regression coefficient of an explanatory variable indicates that the variables increase (decrease) the probability of observing a "Yes" outcome in the dependent variable of interest, *Y*.

We should also note that "e" to the power the corresponding estimated coefficient value for any explanatory variable gives the "odds ratio" with respect to that variable. The odds ratio of an explanatory variable (e.g., gender) represents the likelihood of observing a "Yes" outcome in the dependent variable of interest (e.g., Internet access) for a category (e.g., male) of that explanatory variable relative to that of its base category (e.g., female). For instance, let *X* denotes gender, and it equals "1' if male and "0" if female. Also, let *Y* be the binary dependent variable representing Internet access, and it equals "1" if one has Internet access and "0" otherwise. In that case, if β denotes the estimated coefficient value for gender, the odds ratio of having access to the Internet by gender is characterized as follows:

$$e^{\beta} = \frac{\Pr(Y=1 \mid X=1) / \Pr(Y=0 \mid X=1)}{\Pr(Y=1 \mid X=0) / \Pr(Y=0 \mid X=0)}$$

As noted earlier, for our study we have two dependent variables of interest corresponding to the two distinct aspects of people's Internet adoption behavior—Internet access at home and its daily usage level contingent on access. So, the binary outcomes for our two dependent variables are whether or not a respondent has home Internet access, and whether or not she uses the Internet at home on a daily basis contingent on access. We use six key socio-economic characteristics of the U.S. population that have been typically considered in the existing literature to analyze the Internet-based digital divide [Fox, 2005; Lenhart et al., 2003], and they are as shown in Figure 1.

The two time periods for which we investigate the extent of digital divide in the U.S. population are 2002 and 2008. We estimate the binary logistic regression model separately for each year to get estimates of the relative levels of digital divide for each time period. To see whether there is any distinct change in such relative levels between 2002 and 2008, we test for statistical significance (at .05 level or less) of the difference in estimated coefficient values from the two time periods for each of the explanatory variables [Allison, 1999].

IV. DATA AND ANALYSIS FINDINGS

Data for Analysis

For our empirical analyses, we obtained data through the Pew Internet and American Life Project (PIALP) at the Pew Research Center. Similar data collected through the PILAP has also been used in other recent academic studies [Buente and Robbin, 2008; Lorence et al., 2006]. The specific data used in our study is based on two separate telephone surveys to collect information on Internet access and usage patterns among adult Americans. Both surveys were conducted among a national random sample of adults, eighteen and older in households with telephone. The first survey was done in 2002 between October 30 and November 24 [Kohut and Rainie, 2003], and the second survey was done in 2008 between November 20 and December 4 [Smith, 2009]. The respondents for the surveys were based on a random digit sample of telephone numbers selected from telephone exchanges in the continental United States. The random digit based sampling avoids "listing" bias and provides representation of both listed and unlisted numbers. The survey calls were staggered over times of day and days of the week to maximize the chances of making contact with a potential respondent.

The final sample sizes from the two surveys are 2,745 individuals for 2002 and 2,254 for 2008. Both survey data sets include sample weights for each respondent to address nonresponse patterns inherent in such national survey sampling [Kohut and Rainie, 2003; Smith, 2009]. For each survey, the sample weights were constructed using the

demographic weighting parameters from the most recently available U.S. Census Bureau's *Annual Social and Economic Supplement.*[1] The availability of such sample weights in the data set ensures that, for any statistical analysis based on the data, the demographic characteristics of each survey sample closely represent the demographic characteristics of the target U.S. national population (see Table 1). The demographically representative national random samples give us high confidence in the statistical reliability and generalizability of our findings to capture national trends in each of the survey years (2002 and 2008) and over the six year period between the survey years.

Table 1: Summary Statistics for our Analysis Data

Variables	Year 2002			Year 2008		
	Sample (unweighted)	Sample (weighted)	Target US Population[1]	Sample (unweighted)	Sample (weighted)	Target US Population[1]
Internet adoption level at home (%)						
Access	62.2	60.7	—[2]	69.6	73.1	—[2]
Daily usage with access	51.9	51.1	—[2]	56.1	57.5	—[2]
Mean annual household income (Th. $)	47.2	46.1	45.9	49.7	47.8	47.9
Education level (%)						
Less than high school	10.1	14.6	15.5	8.8	13.3	13.1
High school or more (but no college degree)	59.7	60.3	59.2	57.3	59.5	59.7
College degree or more	30.2	25.1	25.3	33.9	27.2	27.2
Male (%)	49.1	48.2	48.3	47.2	48.9	48.6
Race (%)						
White	84.6	82.0	81.9	84.7	77.8	78.6
African-American	9.8	11.6	11.6	10.3	13.1	12.9
Others	5.6	6.4	6.5	5.0	9.0	8.5
Mean age in years	48.3	48.6	48.3	50.9	45.9	46.4
Rural dwellers (%)	22.1	22.6	22.4	22.2	17.5	17.8

[1] These summary statistic values are based on the U.S. Census Bureau's 2001 (for the 2002 survey) and 2007 (for the 2008 survey) *Annual Social and Economic Supplements* (ASEC). The ASEC for any given year includes all households in the continental United States that had a telephone.

[2] Not collected by the U.S. Census Bureau's 2001 and 2007 *Annual Social and Economic Supplements*.

Our data set contains information from each respondent that allows us to identify whether or not a respondent has access to the Internet at home, and whether or not a respondent who has access to the Internet at home uses it on a daily basis. Our two binary dependent variables are based on this information. From our data, we find that the percentages (weighted) of the respondents who have access to the Internet at home in 2002 and 2008 are 60.7 and 73.1 respectively. From usage level perspective, the percentages (weighted) of the respondents who are found to have access to the Internet at home and to use it on a daily basis are 51.1 in 2002 and 57.5 in 2008. Our data set also contains information on gender, race, age, education, and household income from each respondent. This information as well as the information about whether a respondent lives in an urban or rural area is used to create the relevant explanatory variables in our empirical analysis. As noted earlier, Table 1 shows the summary statistics of all the dependent and explanatory variables used in our analyses.

Analysis Findings

Before we discuss our results, it is important to note that all our regression analysis use the sampling weights of the respondents. As noted earlier, use of such sample weights ensures that, for all our statistical analysis, the sample demographic characteristics closely represent the demographic characteristics of the target U.S. national population.

Internet Access at Home

Our first analysis focuses on the determinants of whether someone has Internet access at home or not. The binary logistic regression results of the likelihood of an individual having Internet access at home is shown in Table 2. From both 2002 and 2008 data analyses, our results show strong evidence that deep digital divide persists along the usual socio-economic fault lines, and it in fact has widened in some cases between 2002 and 2008. For instance, while

[1] For more details on the survey sample design, data collection, and construction of sample weights, please refer to the Appendix.

deep divisions remain by age, education, and gender, the trends in such divisions are found to be "flat." Specifically, we find that the likelihood of a person having Internet access at home falls by about 5 percent ($p < .01$) for each year increase in her age, with no statistically significant difference between 2002 and 2008. Similarly, while persons with college or higher level education are found to have seven times ($p < .01$) more likelihood of enjoying home Internet access than that of those with less than high school education, that relative difference remained statistically the same between 2002 and 2008. We also find no statistically significant difference by gender over the 2002–2008 time period, with men in fact having a 30 percent ($p < .01$) less likelihood than women in having home Internet access. Interestingly, as we discuss later, this gender divide direction gets reversed with respect to Internet *usage* level rather than *access*.

Table 2: Logistic Regression Analysis Results for Internet Access at Home

Explanatory Variables	Coefficient Estimate[1, 2]	
	Year 2002	Year 2008
Annual household income	.353*** (.030)	.479 *** (.044)
Education level (Base is "Less than high school")		
High school or more (but no college degree)	1.092*** (.223)	0.904 *** (.263)
College degree or more	1.954*** (.251)	1.885 *** (.311)
Male	-.339*** (.116)	-.268** (.128)
Race (Base is "White")		
African-American	-.216 (.192)	-.967*** (.264)
Others	-.424* (.247)	-.543 * (.351)
Age	-.053*** (.004)	-.060 *** (.006)
Whether lives in an urban area	.174 (.132)	.355 * (.186)
Intercept	.386 (.282)	1.830*** (.424)
U^2	.26	.28
Prob.> F Value	.00	.00
N	2745	2254

[1] Standard errors of coefficient estimates are given in the parentheses.

[2] Statistical significance coefficient estimates: *** $p < .01$; ** $p < .05$; * $p < .1$.

In contrast, the digital divide with respect to Internet access is found to have in fact widened between 2002 and 2008 in terms of several socio-economic characteristics. Specifically, we find that a richer person in 2008 enjoyed about 60 percent ($p < .01$) more likelihood of home Internet access than that of those in the next lower income group. It reflects a statistically significant ($p < .01$) increase from the 40 percent more likelihood in 2002. As for the racial divide, we find that the odds of an African-American, after controlling for other key socio-economic characteristics, having Internet access at home in 2008 to be about 60 percent ($p < .01$) less than that of a White American. It represents a statistically significant ($p < .01$) widening of the racial digital divide since 2002. We also find similar pattern in the trends of the urban-rural dimension of the digital divide. In 2008, the odds of having home Internet access for urban dwellers was found to be about 40 percent ($p < .05$) more than that of rural dwellers, which again represents a statistically significant ($p < .01$) widening of the rural-urban divide since 2002.

Internet Usage Level

Table 3 presents the binary logistic regression results for the likelihood of an individual who has home Internet access of using the Internet on a daily basis. As with our earlier analysis results for Internet access at home, our results for Internet usage level also show strong evidence of a persistent digital divide along the typical socio-economic fault lines in both 2002 and 2008. In terms of time trends, the digital divide is found to be in fact widening along education and urban-rural fault lines while remaining unchanged along income, racial, gender, and age lines. Specific details of the related analysis results from Table 3 are discussed below.

Table 3: Logistic Regression Analysis Results for Daily Internet Usage at Home		
Explanatory Variables	**Coefficient Estimate**[1, 2]	
	Year 2002	**Year 2008**
Annual household income	.159*** (.028)	.228*** (.033)
Education level (Base is "Less than high school")		
High school or more (but no college degree)	.866 *** (.252)	.933*** (.336)
College degree or more	1.411*** (.266)	1.542*** (.352)
Male	.208 * (.108)	.173* (.093)
Race (Base is "White")		
African-American	-.448** (.204)	-.527** (.216)
Others	.124 (.211)	-.404 (.264)
Age	-.031*** (.003)	-.035*** (.004)
Whether lives in an urban area	.255* (.138)	0.579*** (0.164)
Intercept	-1.447 *** (.313)	-1.303*** (.402)
U^2	.10	.18
Prob.> F Value	.00	.00
N	1707	1568

[1] Standard errors of coefficient estimates are given in the parentheses.
[2] Statistical significance coefficient estimates: *** p< .01; ** p< .05; * p< .1.

As with Internet access, in 2008, people with college or higher level education saw their likelihood of daily Internet use reach more than five times ($p < .01$) than those with less than high school education, which represents a significant ($p < .05$) jump from the four times ($p < .01$) more likelihood in 2002. Urban dwellers are also found to have significantly ($p < .05$) expanded their relative advantage with their odds of using the Internet on daily basis rising from about 30 percent ($p < .01$) higher than that of those living in rural areas in 2002 to about 80 percent ($p < .01$) higher in 2008.

We find that the digital divide by income remains significant but unchanged between 2002 and 2008 with respect to Internet usage level. Specifically, while richer persons are found to have 25 percent ($p < .01$) more likelihood of using the Internet on daily basis than that of those in the next lower income group, that relative difference remained statistically the same between 2002 and 2008. Similarly, the relative odds of an African-American using the Internet on daily basis are found to be about 65 percent ($p < .05$) of that of a White-American, with no statistical difference between 2002 and 2008. In terms of gender divide, the likelihood of daily Internet usage by a male is found to be about 25 percent ($p < .05$) more than that of a female in 2008, but not statistically different from the level in 2002. Finally, we find that the likelihood of an adult person, who has Internet access at home, of using the Internet on daily basis falls by about 3 percent ($p < .01$) for each year increase in her age, again with no statistical difference again between 2002 and 2008.

Summary of Our Findings

The primary insights from our empirical analyses are summarized in Table 4 and Table 5. In summary, we find that an individual's socio-economic profile continues to have statistically significant effects on her likelihood of both Internet access and usage level in the U.S. In fact, not only do we find evidence for the persistence of a deep digital divide along key dimensions of socio-economic fault lines, we also find evidence that the divide has indeed widened between 2002 and 2008 along some dimensions like income and urban-rural divide.

Interestingly, as Table 5 shows, a comparison of the levels of digital divide between Internet access and usage shows that they are relatively higher in the case of access. For instance, our findings show that the digital divide by income and education is much more pronounced in case of Internet access than Internet usage. These findings

Table 4: Digital Divide Trends: Summary of the Findings from Our Empirical Analyses						
	Internet Access			Daily Internet Usage		
Socio-Economic Variable	Evidence of Digital Divide in 2002	Evidence of Digital Divide in 2008	Change in the Levels of Digital Divide between 2002–2008	Evidence of Digital Divide in 2002	Evidence of Digital Divide in 2008	Change in the Levels of Digital Divide between 2002–2008
Income	Yes	Yes	Yes; Increased	Yes	Yes	No
Education	Yes	Yes	No	Yes	Yes	Yes; Increased
Gender	Yes	Yes	No	Yes	Yes	No
Race	Yes	Yes	Yes; Increased	Yes	Yes	No
Age	Yes	Yes	No	Yes	Yes	No
Residential location	Yes	Yes	Yes; Increased	Yes	Yes	Yes; Increased

Table 5: Relative Likelihoods for Internet Access and Daily Usage between 2002 and 2008: Summary of the Findings from Our Empirical Analyses		
Socio-Economic Variable	**Internet Access**	**Daily Internet Usage**
Income	▪ The likelihood for a richer person increased from about 40% more to about 60% more	▪ The likelihood for a richer person remained the same at about 25% more
Education	▪ The likelihood for a college graduate remained the same at about 7 times more than that of one without high school degree	▪ The likelihood for a college graduate jumped from about 4 times more to about 5 times more than that of one without high school degree
Gender	▪ The likelihood for a male remained the same at about 30% lower	▪ The likelihood for a male remained the same at about 25% higher
Race	▪ The likelihood for an African American decreased from similar level to about 60% less than that of a White American	▪ The likelihood for an African American remained the same at about 35% less than that of a White American
Age	▪ The likelihood remained the same at about 5% less for each year increase in age	▪ The likelihood remained the same at about 3% less for each year increase in age
Residential location	▪ The likelihood for an urban resident increased from similar to about 40% more than that of a rural resident	▪ The likelihood for an urban resident increased from about 30% more to about 80% more than that of a rural resident

suggest that the lack of Internet access is more of a reflection of economic or infrastructure constraints than of any lack of inclination or ability to use, once access is secured, by socio-economic characteristics. Also, they indirectly imply that the "Internet have-not" group in fact mostly consists of people who make good use of the Internet once they do get access. Such implication is consistent with the findings and insights from the study by Goldfarb and Prince [2008] that investigated the link between Internet access and usage behavior of individuals.

V. CONCLUSION

Since its advent in the 1990s, the Internet has increasingly emerged as a key source of information for the general public in many domains including business, politics, and general news. Not surprisingly, it has led to the concurrent emergence of a natural and important academic research stream that has focused on an in-depth understanding of the various factors that facilitate or hinder people's adoption of the Internet [Dewan and Riggins, 2005; Ratchford et al., 2003]. A key question in that respect is: Has the increased Internet penetration at the aggregate population level in the U.S. changed the relative levels of digital divide along the typical socio-economic fault lines, and if so, how? A rigorous research approach to address the question requires systematic empirical analysis using repeated national sampling data over multiple time periods. Unfortunately, the existing literature offers very few studies in that regard. The contribution of our study is to use such approach and data to augment this very limited existing research by addressing the above key question on the digital divide issue.

Using individual level survey data from two annual (2002 and 2008) cohorts of demographically representative national random samples, we presented systematic empirical insights into the aforesaid question. We investigated the trends in the digital divide in terms of both home Internet access and its usage level given access. The scope

and nature of our data lend high confidence in the statistical reliability and generalizability of our empirical findings. Among our main findings, both the Internet access and usage level analyses provide mutually reinforcing empirical evidence to indicate that a significant increase in the overall Internet penetration level over the last decade has done little to reduce the digital divide in the U.S. Specifically, we find that an individual's income, education, gender, race, and age, as well as whether she lives in an urban area, continue to have statistically significant effects on her likelihood of both Internet access and usage level conditioned on access. Not only does a deep digital divide persist along such key dimensions of socio-economic fault lines, we find that it is in fact widening over our analysis period (2002–2008), along such dimensions as income and urban-rural divide.

Taken together, our findings suggest that the increase in the Internet penetration at the aggregate population level continues to be driven mostly by still deeper penetration levels among the already "privileged" socio-economic segments of the population. Like the issue of equitable access to basic immunization vaccinations being of prime interest to medical professionals, the research issue and findings of our study on the Internet-based digital divide are of particular interest to computing professionals. Our findings are also of significant practical relevance to businesses as well as to public and civic interest groups.

For businesses, our study is directly relevant to their strategic goal of matching media with target markets by evaluating the relative communication effectiveness of alternative media across different target market segments. As their use of the Internet increases not only as an alternative communication medium to reach their target markets, but also as an alternative transaction channel to sell goods [Baker and Hempel 2006; Chatterjee, et al. 2003; Ghose and Yang 2009], our findings offer strategic insights to make such target marketing more effective. For instance, our empirically estimated "odds-ratios" or relative likelihoods of Internet access and usage by key socio-economic characteristics provide very useful metrics for businesses in evaluating the expected differential effectiveness of using the Internet to target specific demographic markets for advertisement communication and product sales. Also, our overall findings imply that businesses should not expect it to be any easier to reach the "have-nots" through the Internet in 2010 than it was about a decade ago. At the same time, they also underscore the significant opportunities that exist for businesses in more effectively reaching and serving the market at the bottom of the "economic pyramid" as relevant public policy initiatives increasingly aim to break down the digital divide with respect to the Internet [Stelter and Wortham, 2010].

As for the public policy goal to achieve an effective and equitable "cyber society" [Jensen et al., 2007], a study like ours that offers insights into evolving trends in digital divides across important segments of the society is naturally of significant importance in prioritizing relevant policy targets and initiatives. For instance, in the U.S. context, our findings underscore the issue of the urban–rural digital divide as a key area for evaluation of past policy initiatives [Lenhart et al., 2003] and for prioritization of future initiatives. In this context, it is quite interesting and relevant to note that recent policy initiatives in the U.S. give priority to bridging the urban–rural digital divide [Stelter and Wortham, 2010]. Some of these initiatives aim to provide subsidy for Internet providers to wire rural parts of the country that are without Internet access by planning to auction some broadcast spectrum to free space for wireless devices.

As with typical empirical studies using secondary data, we should note that a limitation of our study is that our data lacks some interesting individual demographic characteristic variables like occupation as well as policy relevant variables like price of Internet access and use. An especially interesting area of future research would be to study the role of national and regional level differences in IT-related regulatory and public policy initiatives on digital divide trends within and across countries. In similar vein, another interesting area of future research would be to compare the efficacy of local community based initiatives in reducing digital divide so as to identify "best practices" for designing such initiatives.

ACKNOWLEDGMENTS

The authors thank Pew Internet and American Life Project for the data used in our study. They are grateful to the Associate Editor and two anonymous reviewers for their very helpful suggestions.

REFERENCES

Editor's Note: The following reference list contains hyperlinks to World Wide Web pages. Readers who have the ability to access the Web directly from their word processor or are reading the article on the Web, can gain direct access to these linked references. Readers are warned, however, that:

1. These links existed as of the date of publication but are not guaranteed to be working thereafter.
2. The contents of Web pages may change over time. Where version information is provided in the References, different versions may not contain the information or the conclusions referenced.
3. The author(s) of the Web pages, not AIS, is (are) responsible for the accuracy of their content.
4. The author(s) of this article, not AIS, is (are) responsible for the accuracy of the URL and version information.

Allison, P.D. (1999) "Comparing Logit and Probit Coefficients Across Groups", *Sociological Methods and Research* (28)2, pp. 186–208.

Baker, S. and J. Hempel (2006) "Wiser About the Web", *Business Week* (March 27), http://www.businessweek.com/magazine/content/06_13/b3977401.htm (current Sep. 20, 2010).

Bannerji, S. et al. (2005) "Gender Divide in the Use of the Internet Applications", *The International Journal of e-Business Research* (1)2.

Baye, M., J. Morgan, and P. Scholten (2003) "The Value of Information in an Online Consumer Electronics Market", *Journal of Public Policy & Marketing* (22)1, pp. 17–25.

British Broadcasting Corporation (2010) "Internet Access 'A Human Right,'" http://news.bbc.co.uk/2/hi/technology/8548190.stm (current Mar. 20, 2010).

Buente, W. and A. Robbin (2008) "Trends in Internet Information Behavior, 2000–2004", *Journal of the American Society for Information Science and Technology* (59)11, pp.1743–1760.

Chatterjee, P., D.L. Hoffman, and T.P. Novak (2003) "Modeling the Clickstream: Implications for Web-Based Advertising Efforts", *Marketing Science* (22)4, pp. 520–541.

Chinn, M.D. and R.W. Fairliey (2006) "The Determinants of the Global Digital Divide: A Cross-Country Analysis of Computer and Internet Penetration", *Oxford Economic Papers* (59), pp.16–44.

Compaine, B.M. (ed.) (2001) *The Digital Divide: Facing a Crisis or Creating a Myth*, Cambridge, MA: MIT Press.

Cuervo, M.R.V. and A.J.L. Menendez (2006) "A Multivariate Framework for the Analysis of the Digital Divide: Evidence from European Union-15", *Information and Management* (43)6, pp.756–766.

Dewan, S. and F.J. Riggins (2005) "The Digital Divide: Current and Future Research Directions", *Journal of the Association for Information Systems* (6)12, pp. 298–337.

Fox, S. (2005) "Digital Divisions", Research Report, Pew Internet and American Life Project, Washington, DC 20036.

Ghose, A. and S. Yang (2009) "An Empirical Analysis of Search Engine Advertising: Sponsored Search in Electronic Markets", *Management Science* (55)10, pp. 1605–1622.

Goldfarb, A. and J. Prince (2007) "Internet Adoption and Usage Patterns Are Different: Implications for the Digital Divide", *Information Economics and Policy* (20), pp. 2–15.

Grau, J. (2010) "US Retail e-Commerce Forecast", eMarketer.com, http://www.emarketer.com/Report.aspx?code=emarketer_2000672 (current Oct. 1, 2010).

Hoffman, D.L., P.N. Tomas, and A. Venkatesh (2004) "Has the Internet Become Indispensible?" *Communications of the ACM* (47)7, pp.37–42.

Jensen, M.J., N.D. James, and A. Venkatesh (2007) "Civil Society and Cyber Society: The Role of the Internet in Community Associations and Democratic Politics", *Information Society* (23), pp. 39–50.

Kohut, A. and R. Lee (2003) "Modest Increase in Internet Use for Campaign 2002", Research Report, Pew Internet and American Life Project, Washington, DC 20036.

Lenhart, A. et al. (2003) "The Ever-Shifting Internet Population", Research Report, Pew Internet and American Life Project, Washington, DC 20036.

Lorence, D.P., H. Park, and S. Fox (2006) "Racial Disparities in Health Information Access: Resilience of the Digital Divide", *Journal of Medical Systems* (30)4, pp. 241–249.

McGregor, D. (1999) "Rural America Falls Through the Net", *Financial Times*, U.S. edition (Feb. 18), p. 5.

Norris, P. (2001) *Digital Divide: Civic Engagement, Information Poverty, and the Internet Worldwide*, New York, NY: Cambridge University Press.

Pew Internet and American Life Project (2010) "Internet Adoption over Time", http://www.pewinternet.org/Static-Pages/Trend-Data/Internet-Adoption.aspx (current Oct. 1, 2010).

Ratchford, B.T., M. Lee, and D. Talukdar (2003) "The Impact of the Internet on Search for Automobiles", *Journal of Marketing Research* (40)May, pp. 193–209.

Robinson, J.P., P. DiMaggio, and E. Hargittai (2003) "New Social Survey Perspectives on the Digital Divide", *IT & Society* (1)5, pp. 1–22.

Sims, J., R. Vidgen, and P. Powell (2008) "e-Learning and the Digital Divide: Perpetuating Cultural and Socio-Economic Elitism in Higher Education", *Communications of the Association for Information Systems* (22), pp. 429–442.

Sipior J. et al. (2004) "A Community Initiative that Diminished the Digital Divide", *Communications of the Association for Information Systems* (13), pp. 29–56.

Smith, A. (2009) "The Internet's Role in Campaign 2008", Research Report, Pew Internet and American Life Project, Washington, DC 20036.

Stelter, B. and J. Wortham (2010) "Effort to Widen U.S. Internet Access Sets Up Battle", *The New York Times* (March 12), http://www.nytimes.com/2010/03/13/business/media/13fcc.html?th&emc=th (current Mar. 26, 2010).

Venkatesh, V. and S. Brown (2001) "A Longitudinal Investigation of Personal Computers in Homes: Adoption Determinants and Emerging Challenges", *MIS Quarterly* (25)1, pp. 71–102.

APPENDIX

Analysis Data Source

For our empirical analyses, we obtained data through the *Pew Internet and American Life Project* (PIALP). The PILAP is a nonprofit, nonpartisan initiative of the *Pew Research Center*. The Project is a think tank that explores the impact of the Internet on children, families, communities, the work place, schools, health care, and civic/political life. Support for the project is provided by *The Pew Charitable Trusts*. The project's website is: www.pewinternet.org. The specific data used in our study is based on two separate telephone surveys in 2002 and 2008 conducted by the *Princeton Survey Research Associates International* (PSRAI) on behalf of the PILAP. The PSRAI is an independent research company, based in New Jersey, specializing in social and policy work. It designs, conducts, and analyzes surveys worldwide. Brief details on sample design, sample contact, and sample weighting procedures for the data used in our analyses are noted below. More details are available in Kohut and Rainie [2003] and in Smith [2009]. Similar data collected by the PSRAI on behalf of the PILAP has also been used in other recent academic studies [Buente and Robbin, 2008; Lorence et al., 2006].

Sample Design

The specific data used in our study is based on two separate telephone surveys to collect information on Internet access and usage patterns among Americans. The first survey was done in 2002 between October 30 and November 24 [Kohut and Rainie, 2003], and the second survey was done in 2008 between November 20 and December 4 [Smith, 2009]. The sample for both surveys was designed to represent all continental U.S. telephone households. The sample was drawn using standard list-assisted random digit-dialing (RDD) methodology. Active blocks of telephone numbers (area code + exchange+ two-digit block number) that contained three or more residential directory listings were selected with probabilities in proportion to their share of listed telephone households; after selection two more digits were added randomly to complete the number. This method guarantees coverage of every assigned phone number regardless of whether that number is directory listed, purposely unlisted, or too new to be listed. After selection, the numbers were compared against business directories and matching numbers purged.

Sample Contact

As many as ten attempts were made to contact every sampled telephone number. Sample was released for interviewing in replicates, which are representative subsamples of the larger sample. Using replicates to control the release of sample ensures that complete call procedures are followed for the entire sample. Calls were staggered over times of day and days of the week to maximize the chance of making contact with potential respondents. Each household received at least one daytime call in an attempt to find someone at home. In each contacted household, interviewers asked to speak with the youngest adult (eighteen years or older) male currently at home. If no male adult was available, interviewers asked to speak with the oldest female adult at home. This systematic respondent selection technique has been shown to produce samples that closely mirror the population in terms of age and gender.

Sample Weighting

Weighting is generally used in survey analysis to compensate for patterns of non-response that might bias results. The interviewed sample of all adults in both the surveys was weighted to match national parameters for demographic variables like sex, age, education, and race. These parameters came from an analysis of the U.S. Census Bureau's 2001 (for the 2002 survey) and 2007 (for the 2008 survey) *Annual Social and Economic Supplements* (ASEC). The ASEC for any given year includes all households in the continental United States that had a telephone. Weighting was accomplished using Sample Balancing, a special iterative sample weighting program that simultaneously balances the distributions of all variables using a statistical technique called the *Deming Algorithm*. The use of these sample weights in statistical analysis ensures that the demographic characteristics of the sample closely approximate the demographic characteristics of the target national population.

ABOUT THE AUTHORS

Debabrata Talukdar is Professor of Marketing at State University of New York at Buffalo. He received his Ph.D. in Marketing and Applied Economics from the University of Rochester. He also holds master's degrees in Operations Research from the University of Rochester and in International Development from Massachusetts Institute of Technology. Before joining academia, he worked on international development and environmental policy issues at the World Bank. His primary research interests lie with quantitative modeling and empirical analyses of issues related to economic development, and marketing of both "private" and "public" goods to consumers. His research works have been presented and published in various international conferences and leading journals in the economic development and marketing areas. The works continue to be widely cited in leading academic journals from a wide spectrum of areas like economics and public policy, marketing, management science, computer information systems, law, and public health.

Dinesh K. Gauri is an Assistant Professor of Marketing in the Whitman School of Management at Syracuse University. He received his Ph.D. from the University at Buffalo. He also holds master's degrees in Economics from University at Buffalo and in Mathematics and Computer Applications from Indian Institute of Technology, Delhi, India. His research interests include consumers' price search, retail strategy, promotion effectiveness, retail store performance, and online store loyalty related issues. His research has been published in various journals including *Journal of Marketing, Journal of Marketing Research, Marketing Science, Journal of Retailing* and *Communications of ACM*. He serves on the editorial board of *Journal of Retailing* and *Journal of Business Research*.

Communications of the Association for Information Systems

ISSN: 1529-3181