



## **MORE THAN THE DIGITAL DIVIDE? INVESTIGATING THE DIFFERENCES BETWEEN INTERNET AND NON-INTERNET USERS ON ATTITUDES AND BEHAVIORS**

---

Chan Zhang

*University of Michigan, Program in Survey Methodology. Ann Arbor, MI*

Mario Callegaro

*Knowledge Networks, Inc., Menlo Park, CA*

Melanie Thomas

*Knowledge Networks, Inc., Menlo Park, CA*

Prepared for the Midwest Association for Public Opinion Research (MAPOR)  
2008 Conference

*March 2009 version*

### **ABSTRACT**

Even before the internet became popular among the public, researchers had identified the potential problems involved in using a web-based data collection method for survey research. Probability-based web surveys are generally believed to produce data of higher quality than those from non-probability web surveys. Both of these methods, however, are subject to coverage error. This is shown by the fact that in 71% of American households, at least one member uses the internet at any location (including at home), and that only 62% of households have an internet connection from home as of October of 2007. Moreover, some subgroups of the population are more likely to be offline than others. This phenomenon is often referred to as the digital divide. Evidence also suggests that this divide may not be closing quickly enough to be dismissed by researchers who want to study a representative sample of the population. As a solution, Knowledge Networks covers the ISP charges and provides WebTVs and in some cases PC's to its panelists (i.e., KnowledgePanel® members) who live in a household without internet access, thus closing the digital divide gap. As a result, all panel members recruited by landline RDD are able to complete surveys online, which minimizes the potential bias that can result from ignoring non-internet users. Data from KnowledgePanel® are examined to build a profile of these non-internet users to find out who they are and how they behave. In this paper, demographic characteristics of the offline population are first explored and compared to other research findings. Then, responses to a series of survey questions are analyzed using (multinomial) logistic regressions, to identify whether internet access status still makes a difference when controlled by relevant demographic variables. For example, the analysis reveals that differences in environmental behaviors between internet and non-internet users (e.g., recycling paper and glass) still exist after controlling for the relevant demographic variables.

## BACKGROUND ON THE DIGITAL DIVIDE DEBATE

### *From No Access to Access*

The debate on inequality of internet access in the U.S. has been discussed for over a decade. Since 1995 when only three percent of Americans had ever used the World Wide Web (Pew Center for the People and the Press, 1995), researchers had identified differences in access depending on a number of variables. In the same year, the National Telecommunication and Information Administration (NTIA) issued their first report using the Internet and computer use data supplement of the Current Population Survey (CPS) with an emblematic title: "Falling through the net: A survey of the 'have nots' in rural and urban America" (Department of Commerce, 1995). Since then, NTIA and the Pew have continued to provide publicly available data on internet access in the U.S.

The year 2002 was a turning point in the debate about the digital divide. The NTIA reports, up to that year, focused on inequality of access and differences in demographics. The title of almost every document started with "Falling through the net." For further reading, DiMaggio et al. (2004) provide a detailed literature review of digital divide studies up to 2002.

*Breaking with tradition, the fifth report of the NTIA published in 2002, based on data collected in 2001, was called: "A nation online: How Americans are expanding their use of internet" (Department of Commerce, 2002). The title of this last report no longer focused on differences. Rather, it highlighted the fact that more and more Americans were going online" (Hargittai, 2008, p. 937).*

A quick read of the report (e.g. executive summary) could easily influence the reader to think that the internet had reached a universal level and that usage had become more equally distributed across households, but if the reader were to peruse the report, it would be evident that the level of equality was misleading. The strongest critique came from Steven Martin (2003) who reanalyzed the CPS data used by NTIA and provided evidence that computer and internet usage spread much slower among poorer households than richer ones. This new report directly contradicted the NTIA suggestion that the digital divide was quickly closing between the poor and other disadvantaged groups, and the richer households. The findings of Martin were echoed in the conclusions of a Pew Internet and American Life project report (Horrigan et al., 2003) that used their own data collected with RDD telephone surveys:

*Between April 2000 and the spring of 2002, the Internet population grew across all demographic groups. But the gaps between rich and poor, well-educated and less-well educated, rural and suburban, black and white, the disabled and non-handicapped, and old and young persist (p. 33).*

The sixth NTIA report (Department of Commerce, 2004) shifted the focus of the digital divide to broadband access and types of online behaviors. Disparities in internet penetration were not discussed in the text, although two tables in the appendix show how income, education and race are still key variables to explain the differences in dialup versus broadband access. Income was the focus of a new analysis by Martin & Robinson (2007) using the 1997-2003 Internet and computer use data supplements of the CPS. The authors found that through the years, the increasing rate of the odds of accessing the internet slowed down for lower income households. The trend is similar for broadband access. Other disparities in internet access for minorities, specifically Latinos, were highlighted by Fairle (2007).

The latest data from the CPS are the October 2007 Internet and computer use supplement (Current Population Survey, 2008). According to the CPS, 61.7% of U.S. households connect to the internet from home, while in 71% of households at least one of the members connects to the internet from any location. The National Telecommunications and Information Administration (2008a) issued a report called "Networked nation: Broadband in America 2007". In this report the focus is only on broadband connection and progress in broadband access. Data on non internet households are available only on some tables published in a separate document (National Telecommunications and Information Administration, 2008b). Analysis done on this new dataset show again evidence of digital divide among minorities, low educated, low income, and households with language barriers (e.g. Latinos) (Fairle, 2008).

## *Dialup versus Broadband Access*

The adoption rate of broadband versus dialup internet connection also reveals profound discrepancies in user demographics. Broadband subscription in the United States has steadily increased since 1999 when there were fewer than three million lines, to 2006 when there were more than 60 million lines (Prieger & Hu, 2008). Access to these lines, however, has been shown to be unevenly distributed. In 2002, AT&T was sued in Florida for allegedly bypassing minority neighborhoods when deploying broadband. Areas with high-concentrations of minorities or poorer people are less likely to have access to broadband high-speed internet than whites or richer people. Additionally, the Congressional Research Service identified in 2008 that "the rate of broadband deployment in urban and high income areas appears to be outpacing deployment in rural and low-income areas" (Kruger & Gilroy, 2008). While this disparity is being recognized, the explanations are unsatisfactory.

The 2008 NTIA report discusses how home broadband usage in America has "soared" (9.1% in 2001 to 50.8% in 2007), and that even rural America has experienced "impressive growth" (5.6% in 2001 to 38.8% in 2007) (National Telecommunications and Information Administration, 2008). The rural areas, however, continue to lag in broadband adoption, and almost two times as many continue to use a dial-up connection. The report shows that 10.7% of home internet users in the United States use a dial-up connection, compared to 19.3% of rural internet users (National Telecommunications and Information Administration, 2008a). While broadband usage continues to increase, the reasons for adopting or not adopting the faster technology are not always intuitive.

Of people who adopted broadband service up to 2006, 57% said they wanted a faster connection, and only 4% mentioned price as being a factor in their decision (Horrigan, 2006). When dial-up users are asked if they want to switch to broadband, 39% express interest, while 60% express no interest. A reason used to explain why 60% of dial-up users do not want to switch to a faster connection is that 22% percent of them already have a high speed connection at work. Another reason used to explain why the rest of dialup users do not want to switch is that cost is an impediment, because many are older or have low-incomes (Horrigan, 2006). The cost for dialup (\$18) is on average cheaper than broadband and among broadband, DSL is cheaper than cable. In 2005 the average DSL bill was reported as \$32 per month, while the average cable modem bill was \$41 per month. This cost discrepancy is useful to explain the increase in broadband, but does not explain the persistent use of dialup by rural areas.

The phenomenon of uneven geographical distribution of broadband lines has also been explained by the demand for broadband, specifically that minority groups do not demand broadband access in as high rates as whites. Two of the main contributors to the

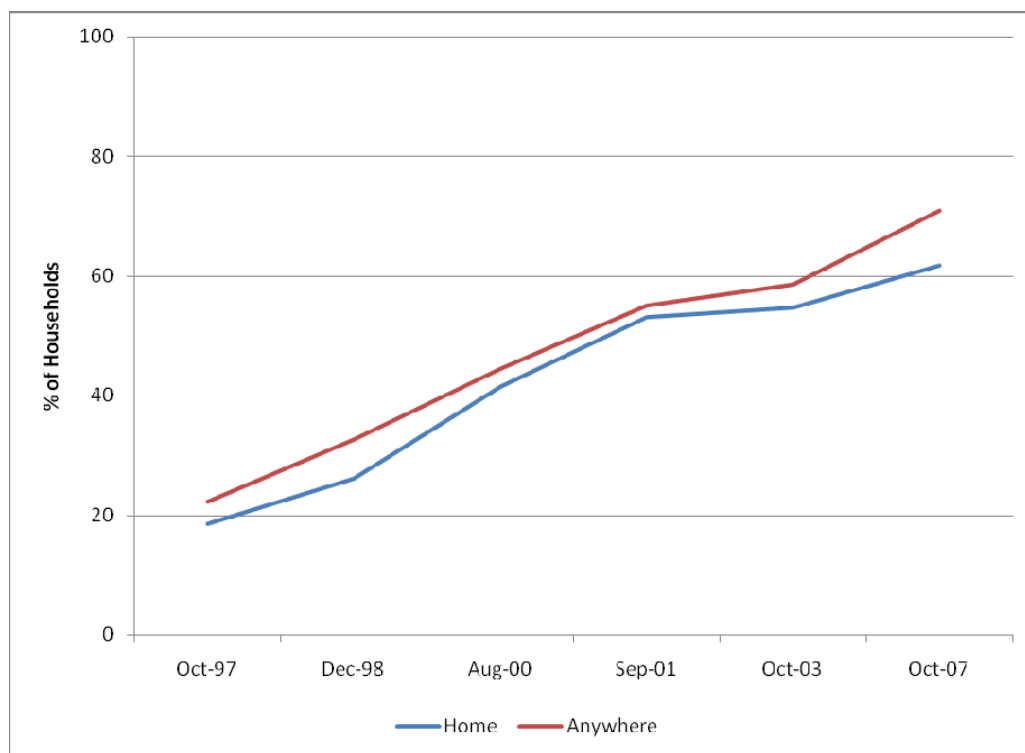
problem are that “competition, or its lack, is an important driver of the Digital Divide,” and that service quality is a “large detriment of demand” (Prieger & Hu, 2008). Unfortunately, there is a lot of ignorance surrounding rates of connection speed to the internet. It has been found that one-sixth of all Americans do not know if broadband service is available in their area, and even among internet users there is unawareness of broadband availability. The ignorance in availability also differs by race in that 24% of Caucasian internet users do not know if broadband is available in their area, while 27% of African American internet users and 43% of Hispanic internet users do not know (Prieger & Hu, 2008). Another study showed that 81% of internet users did not know their home-speed connection rate (Horrigan, 2006). Not knowing what kind of service is available in their area, or even the kind of service they currently use, prevents these minority groups from influencing the perceived demand for the product. Without an accurate measure of demand, competition is less likely to increase in those areas, which would both lower prices and increase quality of internet service.

To put these data into perspective we now provide some trends using three different sources.

### Latest Trends

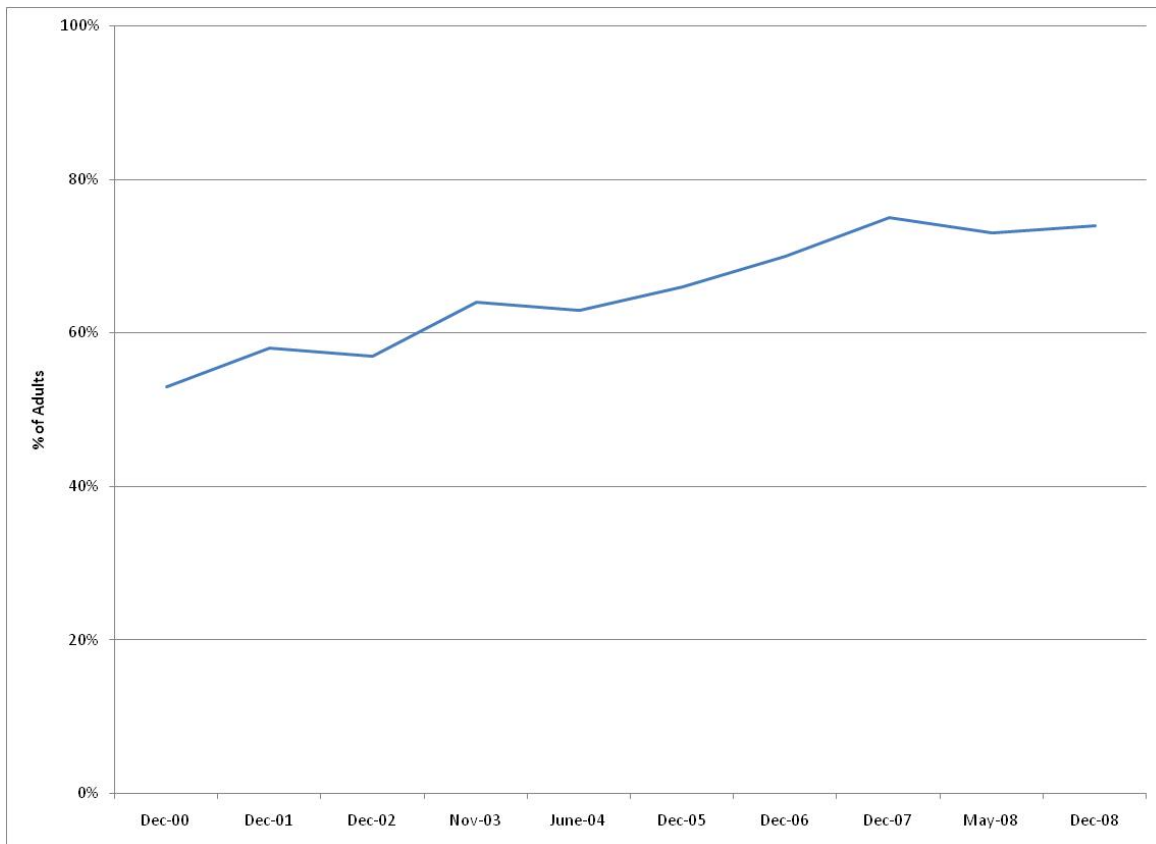
In order to understand how internet penetration is developing, it is useful to see some historical trends. In order to do so we provide three measures from three different agencies and their specific definition of access.

The data from CPS come from a survey conducted primarily face-to-face with an area-based sample. The estimates were based on the following question wording: “Do you/Does anyone in this household use the Internet at any location?” and “Do you/Does anyone in this household connect to the internet from home?”.



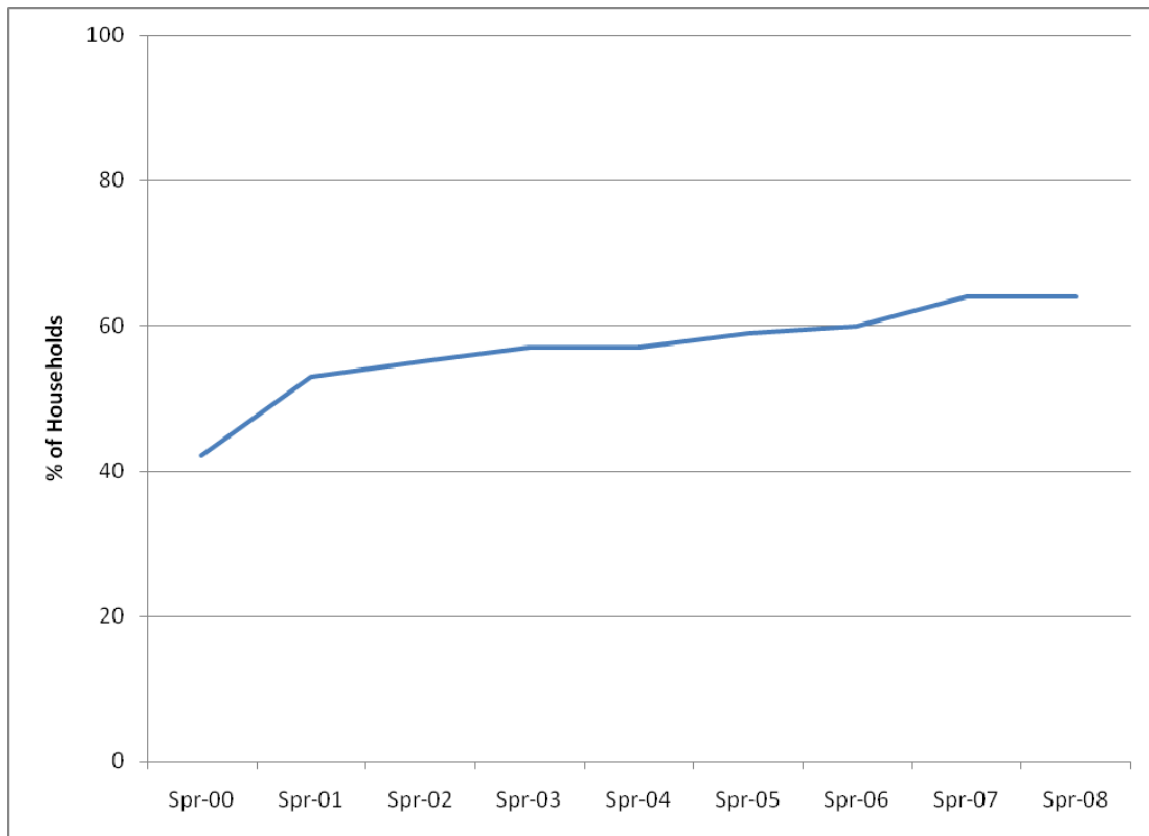
**Figure 1.** Trend in internet access from home and from anywhere, CPS Internet and computer usage supplements. Data at a household level.

In October 2003, at least one household member connected to the internet from any location in 58.7% of American households, while 54.6% of American households had internet access from home (Department of Commerce, 2004). Four years later these two numbers jumped to 71 and 61.7% respectively (Current Population Survey, 2008). The Pew Internet and American Life Project uses a RDD telephone survey to monitor internet penetration in the U.S. (Pew Internet & American Life Project, 2009). To compose the chart, data collected at the end of the year were selected. The definition of internet use changed during the years, with the common denominator of using the internet at least occasionally from any location. Pew provides the data on adults. The most recent estimates were that 74% of U.S. adults went online in December 2008 with a flattening trend since December 2007.



**Figure 2.** Trend in internet access from anywhere, Pew Internet and American Life Project. Data at a person level (18+).

Lastly we provide data from the Knowledge Networks HomeTechnology Monitor™ (Knowledge Networks, 2008). Data were collected every year in the Spring with a RDD telephone survey. The question wording was as follows: “Can you connect to an on-line service or the internet, or send and receive email on (any of) your home computer(s)?”. During the last two measurements in 2007 and 2008, the internet penetration at a household level was 64%.



**Figure 3.** Trend in internet access from home, Knowledge Networks' Home Technology Monitor™. Data at a household level.

These three data sources were in agreement, showing how in the past few years the pace of internet penetration has slowed down. About 40% of U.S. households do not have an internet connection from home and about 30% of adults do not go online even occasionally.

### **THE DIGITAL DIVIDE AS COVERAGE ERROR FOR INTERNET SURVEYS**

As internet surveys are becoming more and more popular to collect data, especially with the use of online panels, the question about the generalizability of results to the entire target population remains crucial. Online panels are increasingly being used for market (Comley, 2007; Postoaca, 2006), social (Tortora, 2009), psychological (Göritz, 2007) and medical research (Couper, 2007). Market research is the sector relying most heavily on online panels. Comley (2007) citing data from Inside Research, reported that in 2006 one third of all market research in the U.S. was conducted online.

By definition, online panels do not include non-internet household members, and thus are not representative of non-internet households. This paper explores differences in responses to survey questions between internet and non-internet users using a sample of KnowledgePanel members from internet and non-internet households.

According to our definition, non-internet households are households that at the time of recruitment to the panel, did not have internet access, and therefore were given a WebTV or more recently a laptop in order to complete surveys online. Details about KnowledgePanel recruitment are described in the section about datasets.

## Previous Studies

Few studies have investigated the differences in survey responses between internet and non-internet households.

In one of the frequently cited books on the digital divide, Pippa Norris (2001) used data from the 1999 Values Update Survey collected by an RDD telephone survey at the Pew Center for the People and the Press. The author found evidence of a distinct cyber culture in America. Internet users (defined as accessing the web from anywhere or sending and receiving emails) were found to be significantly more secular towards traditional morality, more tolerant towards alternative lifestyles such as homosexuality, and less approving of censorship. Online users were also found to be more laissez-faire towards the role of the state. These differences were not simply the byproduct of the younger age and higher education of the online population, but remained statistically significant even after controlling for age, gender, race, income, and education using OLS regression models (chapter 10).

Robinson, Neustadtl and Kestnbaum (2002) used the 2000 general Social Survey (GSS), a face-to-face survey, for their analysis. They classified internet users as people who used the internet from any location including WebTV, and with four levels of usage depending on the number of hours per week (p. 287). Items analyzed were attitudinal questions measuring openness, tolerance, and diversity. The authors performed the analysis on each attitudinal item in a simple bivariate way before and after adjusting the differences by age, education, gender, race, and income using Multiple Classification Analysis (MCA). Internet users were found to be more supportive and tolerant than non-internet users, even after controlling for the demographic differences. The authors dubbed this phenomenon as diversity divide. On other topics, differences were reduced or disappeared after controlling for demographic variables. Moreover, the relationship found was not monotonic, for example it was not found that heavy internet users were increasingly more different than non-internet users. Internet users were also found to have more optimistic views of their lives and more trust in people than non-internet users. On many other variables, however, the differences disappeared after controlling for the previously described variables using MCA.

In a later study, Robinson and Martin (2005) analyzed the 2000, 2002, and 2004 GSS data. The authors replicated the results from their previous study finding internet users more tolerant than non-internet users (e.g. on political minorities). Given the three datasets, it was also possible to compare the differences between internet and non-internet users over time. Regarding the tolerance index, the research found that "each survey year seems to leave being a less politically tolerant group of nonusers" (p. 22). On attitudes of personal trust, racial views, gender issues, sex attitudes, confidence on science, and self-perception of health, internet users were found to have higher scores than non-internet users (see Table 10 on page 24). The authors conclude that:

*...the results have strong methodological implications for surveys conducted over the Internet, or only with Internet users, in that certain attitude groups are overrepresented in such samples, in a way that cannot be adjusted by simple demographic factors...* (p. 32).

Valliant and Lee (2005) used data from six months of the 2003 Survey of Consumer Attitudes (SCA), a monthly RDD telephone survey conducted by the University of Michigan. Internet users were defined as such if they answered positively to the question: "Do you have internet ownership from home?". The analysis concentrated on a battery of questions about their personal outlook and the country's economic outlook.

Internet users were found to be consistently more optimistic than non-users about the economic outlook. The authors re-ran the analysis, this time controlling for income, home ownership, region, race, marital status, sex, education, and age with a series of survey-weighted logistic regressions. For 12 of the 13 items under investigation the differences between internet and non-internet users disappeared after controlling for the aforementioned eight variables. The authors concluded that for the variables analyzed, statistical adjustments to web-survey weights using calibrations to population totals for socioeconomic variables can reduce the bias of surveys using only the online population.

Similar to this finding, Dever, Rafferty and Valliant (2008) analyzed the data from the 2003 Michigan Behavioral Risk Factor Surveillance System (BRFSS), in which an internet module was attached to the core questionnaire administered by phone using a list-assisted RDD sample. Internet access was measured by the question: "Do you have access to the Internet from home?". The authors analyzed 25 health characteristics variables by presence of internet at home. Using similar methods to the studies described above, with a straight comparison, 20 of the 29 characteristics had statistically significant differences between the two groups. When controlling for up to eight demographics variables (income, children in the household, age, race, gender, education, employment and marital status) in a logistic regression setting, most differences between the two groups disappeared. The authors' 2008 addition to this paper simulated a weighting procedure of the internet-users subsample using general regression weights (GREG), a technique motivated by a linear relationship between an analysis variable (internet/non-internet) and a set of covariates (demographics variables). The assumption for the procedure to work is that the same model should hold for both groups, and that population values are available. They were taken from the Michigan Current Population Survey. The GREG weights effectively compensated for the coverage error when seven covariates were used (all of the above demographics minus income). The authors concluded by saying that the results were promising: "there is some hope for using well-designed internet surveys to make estimates for the general population"(p. 61), although they did not discuss the size of the weights used in the analysis and their impact on final estimates. In addition, their results do not apply to volunteer or opt-in web surveys.

A different approach was taken by Piekarski and colleagues (2008). In the first part of the study they compared demographic differences using a face-to-face interview (Survey of the American Consumer or SAC) from Mediamark Research using four groups: respondents with no internet access; respondents with access but no use in the past month; low internet users ( $\leq 5$  times a day); and high internet users ( $\geq 5$  times a day). When looking at users versus non-users (non-internet households), the results were in line with the literature on the digital divide: internet households being disproportionately younger, higher income earners, and more highly educated. Differences were almost eliminated when controlling for the usual demographics variables but persisted between low and high users of the internet. In the second part of the study, the authors compared responses to attitudinal questions from the SAC to results from a sample of the Survey Sampling International (SSI) Survey Spot™ opt-in panel controlled by gender, age and census region. The SSI respondents were by definition internet households. Several differences were found between the Survey Spot sample and the SAC sample even after controlling for basic socio-demographic factors. Specifically, differences in product consumption, services used, and attitudes that were related to technology and/or internet use were identified. The last step was to verify if differences between the two samples still existed after weighing the Survey Spot sample. The authors concluded that: "Internet users that belong to Access Panels [like the Survey Spot example] are different from all internet users and frequent users."

Rookey, Hanway and Dillman (2008) analyzed data from the Gallup panel comparing internet versus non-internet users, the latter being surveyed via a mail questionnaire. The authors found that, not unexpectedly, non-internet households were older, had lower income and less likely to be married than internet households. On survey items, the differences among the two groups did not disappear after standard weighting procedures (10 demographic variables) for one fifth of all items. To rule out possible mode effects between mail and web, the authors surveyed a subset of the respondents by telephone (third mode) and found some evidence that mode was not a factor in explaining different answer between internet vs. non-internet users

From the studies reviewed so far it appears that differences between internet and non-internet users can sometimes be eliminated by controlling for a set of demographics predictors, but it is not always the case. In many instances, the differences in attitudes or behaviors still remained after taking demographic variables into account. The research identified so far, excluding the last study, entailed analyzing telephone interviews and using an internet status variable to run the analysis for probability-based sampling design surveys. The last two studies reviewed are different. In the study Piekarski et al's (2008) answers were collected on an online panel, and secondly, the sample was non-probability based (opt-in). The Gallup panel study uses the same sample and compares internet and non-Internet households' answers to the same questions, with the difference that the non-internet household used a mail instrument.

The research presented is different from the previous studies, because it entails a sample where all respondents, internet and non-internet households, replied to survey questions using the same instrument (web). The sample also comes from a probability-based RDD design, which allows for the computation of confidence intervals and sampling statistics.

## **ABOUT KNOWLEDGEPANEL<sup>®</sup>**

Knowledge Networks initially selects households using random digit dialing (RDD) sampling methodology. The sampling is done without replacement, to ensure that numbers already fielded by Knowledge Networks do not get fielded again. Once a household is contacted by phone and household members are recruited to the panel by obtaining their e-mail addresses or setting up e-mail addresses, panel members are sent surveys over the internet using e-mail. As of August 2002, those RDD households that inform interviewers that they have a home computer and internet access have been recruited to the panel and asked to take their surveys using their own equipment and internet connections. If the household does not have a PC and access to the internet, they are told that in return for completing a short survey weekly, the household will be given a WebTV set-top box or a PC and free monthly internet access.

Households are classified as internet or non-internet households upon recruitment, and these data are confirmed and or corrected every with a specific computer usage and technology profile survey.

## DATASET AND UNIVARIATE STATISTICS

The dataset used was comprised of two different surveys sent to all panel members from late 2007 to April 2008. The total sample size used was 34,577 respondents. In order to simplify the calculation of some variables (e.g. education) and avoid clustering effects, one random adult per household was selected. The first survey, titled the Core Adult profile, was sent to all initial panel recruits. It collected basic demographic information for each panel member, and was a requirement to join KnowledgePanel. The second study chosen was the Public Affairs profile, which had a completion rate of 92.5% (Callegaro & Disogra, 2008).

In Table 1 we compare weighted demographics characteristics from the KnowledgePanel adult sample by internet and non-internet status to the weighted October 2007 CP Internet and computer usage supplement (Current Population Survey, 2008). Our two surveys were chosen to reflect data collected in more or less the same period of the CPS supplement. The CPS data are considered a gold standard, given the very high sample size (50,000 households), the in-person administration, and the extremely high response rate.

**Table 1.** Weighted characteristics of the KnowledgePanel adult sample and weighted CPS October 2007 Internet and computer use supplement data 18 years of age and older.

	Non-Internet HH		Internet HH	
	KN	CPS2007	KN	CPS2007
<b>Income</b>				
<\$10,000	14.56	13.01	3.35	2.63
\$10,000~\$50,000	61.53	66.45	34.78	31.97
\$50,000~\$75,000	14.35	12.39	23.34	23.81
>\$75,000	9.56	8.15	38.54	41.59
<b>Own</b>				
Own	59.12	58.51	79.14	78.51
Rent or Other	40.88	41.49	20.86	21.49
<b>Region</b>				
Northeast	18.14	16.86	18.68	19.47
Midwest	20.91	22.08	20.91	22.03
South	39.40	40.29	36.00	34.35
West	21.55	20.77	24.41	24.15
<b>Gender</b>				
Male	49.67	47.08	44.65	48.92
Female	50.33	52.92	55.35	51.08
<b>Race</b>				
White, Non-hispanic	56.36	57.55	74.93	74.57
Black, Non-hispanic	19.94	16.67	7.66	8.52
Hispanic	17.84	21.00	10.66	9.64
Other	5.85	4.77	6.74	7.27
<b>Education</b>				
<HS	25.08	26.89	6.43	6.60
HS	42.34	41.00	23.62	26.75
Some College	21.14	21.06	32.15	31.59
Bachelor or More	11.45	11.06	37.80	35.06
<b>Age</b>				
18~29	18.91	20.72	20.78	22.51
30~44	23.35	22.56	30.79	30.05
45~59	30.24	23.01	28.08	30.16
60+	27.51	33.72	20.34	17.28
<b>MSA</b>				
Yes	80.78	79.91	85.41	86.52
No	19.22	20.09	14.59	13.48

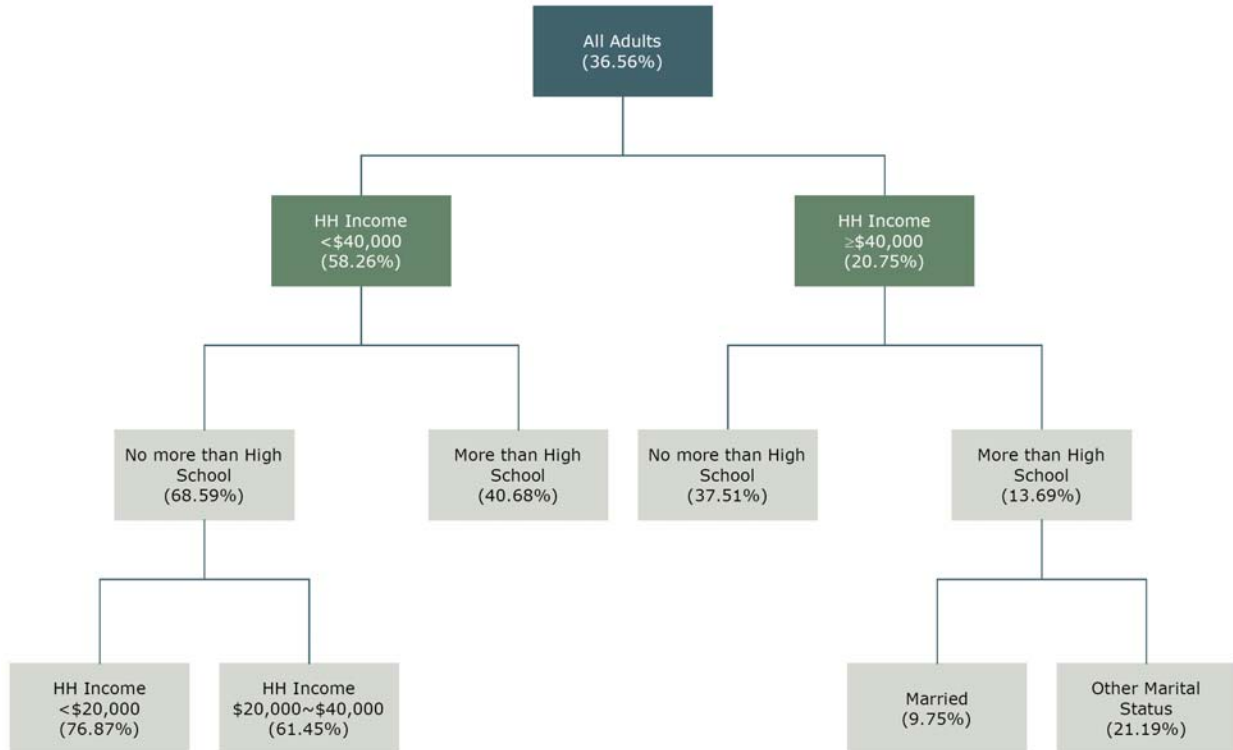
**Table 1.** Weighted characteristics of the KnowledgePanel adult sample and weighted CPS October 2007 Internet and computer use supplement data 18 years of age and older (Continued).

	Non-Internet HH		Internet HH	
	KN	CPS2007	KN	CPS2007
<b>Employment Status</b>				
Employed	49.96	51.15	66.63	71.60
Unemployed	9.41	3.30	4.92	2.60
Not working-retired	18.15	24.90	14.04	11.53
Not working-disabled	15.56	8.99	4.77	2.89
Not working-other	6.92	11.66	9.64	11.38
<b>Marital Status</b>				
Married	34.56	44.92	59.94	61.51
Never Married	27.00	26.63	18.79	24.70
Other	38.44	28.44	21.26	13.80
<b>Presence of Own Kids in the HH</b>				
Yes	25.91	22.7	36.72	32.82
No	74.09	77.3	63.28	67.18

The comparison of this particular sample with the CPS sample is generally showing minor differences with a few exceptions due to different question wording and sampling error.

We then performed a SEARCH procedure (Morgan, Solenberger, & Nagara, 2005; Sonquist, Baker, & Morgan, 1973) of demographic characteristics associated with non-internet households. The analysis was run using a freeware version of SEARCH for SAS, available at the University of Michigan (Morgan, Solenberger, Van Eck, & Nagara, 2007). The procedure was used to identify the demographic characteristics that had relatively strong associations with household internet-access status. The procedure started with the entire sample and searched among all the possible binary splits of each predictor variable. The program then chose the split that best maximized the differences in distribution of the dependent variable between the two subgroups, measured by the likelihood-ratio chi-square statistics. The analysis was repeated for each resulting group, and the branches of the “tree” grew until the reduction in error variance was less than 0.4 percent. The SEARCH procedure, therefore, did not assume specific relationships between predictors and the dependent variable, and among predictors.

As predictors, we used education, age, income, gender, household size, ethnicity, employment status, housing status, marital status, an indicator of being a parent or guardian, and an indicator of rural/urban area. Figure 1 shows the results of the analysis for non-internet households.



**Figure 4.** SEARCH analysis of demographic characteristics associated with non-internet households. Percentage represents the percentage of adults in that demographic group living in a non-internet household.<sup>1</sup>

The strongest and second strongest predictors of being a non-internet household are income and education respectively. The results from table 1 and Figure 1 confirm that non-internet households in KnowledgePanel have similar characteristics to non-internet households in the general population. The SEARCH procedure highlighted again how income and education are the key drivers of the digital divide, as the previous literature suggested.

## MULTIVARIATE RESULTS

In order to compare answers to survey questions between internet and non-internet households, controlling for demographics characteristics, we ran a series of multivariate logistic regressions with internet status as the explanatory variable (non-internet households 0; internet households 1). According to our definition, internet users had internet access at home at the time of recruitment, while non-internet households did not; therefore they were provided a WebTV unit and an internet connection. Each web survey was programmed to be perfectly compatible with the Microsoft WebTV browser so the PC and WebTV respondents were exposed to the same screen by screen questionnaire. For this preliminary analysis, we selected a subset from the battery of questions.

<sup>1</sup> For example 76.78% of our respondents whose household income is below \$20,000 and who have no more than high school education live in a non-Internet household.

The analysis was done using four different models; in the univariate model only the internet status variable was a covariate in the model, and no other variables were used as controls. The univariate model was therefore the baseline measurement, comparing answers from internet and non-internet households. In the basic model we controlled for the most influential variables on internet status: income, age, education, and ethnicity. In the intermediate model, gender, region and metropolitan status were added as predictors. The full model was then built based on the intermediate model, with employment status, marital status, presence of kids in the household, and home ownership also controlled as covariates. This approach gives a total of 12 explanatory variables including the internet access variable.

Table 2 summarizes the analysis results using the four regression models from a series of questions selected from KnowledgePanel's Core Adult profile survey and Public Affairs profile survey. For each survey question, four estimated odds ratios (OR) are reported, corresponding to the four regression models. The odds ratio compares the odds of certain responses for people living in internet households to those living in non-internet households, conditional on other covariates controlled in the model. An odds ratio greater than one indicates that people with internet access at home are more likely to report that behavior or attitude than those without internet access, and vice versa. It is also indicated in the table whether the estimated odds ratios are significantly different from one at the 95% level.

**Table 2.** Survey questions odds ratios for the four logistic regression models with internet status as explanatory variable.

Questions:	Probability modeled	Univariate	Basic Model	Intermediate Model	Full Model
<i>In the past 12 months, have you...?</i>					
Given money to an environmental organization	"Yes"	1.669*	1.065	1.054	1.094
Contributed time or expertise to an environmental group	"Yes"	1.521*	1.076	1.076	1.138
Recycled newspapers or other papers	"Yes"	1.993*	1.339*	1.327*	1.275*
Recycled cans	"Yes"	1.642*	1.203*	1.197*	1.145*
Recycled glass	"Yes"	1.911*	1.258*	1.281*	1.260*
Recycled plastic	"Yes"	1.897*	1.275*	1.275*	1.256*
Purchased a product specifically because it was made from recycled materials	"Yes"	1.551*	1.245*	1.197*	1.228*
Taken steps to reduce the use of energy(oil, gasoline, electricity)	"Yes"	1.437*	1.188*	1.167*	1.149*
Taken other actions to improve the environment	"Yes"	1.235*	1.061	1.049	1.083
Any of these actions	"Yes"	1.895*	1.330*	1.280*	1.231*

\* Significant at the 95% level

**Table 2.** Survey questions odds ratios for the four logistic regression models with internet status as explanatory variable. (Continued)

Questions:	Probability modeled	Univariate	Basic Model	Intermediate Model	Full Model
<i>In your opinion, what is the most important issue facing the U.S. today?</i>					
Taxes	"Yes"	0.936	0.892	0.951	0.913
Education	"Yes"	1.584*	1.296*	1.268*	1.272*
War on Terrorism	"Yes"	0.895*	0.915	0.941	0.919
Situation in Iraq	"Yes"	0.944	0.879*	0.871*	0.912*
Economy /Jobs	"Yes"	1.023	1.074	1.084	1.105*
Moral Issues	"Yes"	1.229*	1.233*	1.218*	1.085
Healthcare	"Yes"	0.886*	0.986	0.967	0.993
Other	"Yes"	0.987	0.938	0.945	0.985
<i>How interested are you in politics and public affairs?</i>	"Very Interested"	1.304*	1.105*	1.146*	1.189*
<i>Do you agree or disagree with the following statements about political affairs in this country?</i>					
Most public officials are not really interested in the problems of the average person	"Strongly Agree"	0.859*	1.097*	1.117*	1.117*
It is a citizen's duty to serve on a jury even if it interferes with his/her private life.	"Strongly Agree"	1.309*	1.077	1.089*	1.098*
It is a citizen's duty to report a crime even if it might put him or her in some jeopardy.	"Strongly Agree"	0.933	0.893*	0.903*	0.887*
It is a citizen's duty to pay taxes even if they seem unfair or too high.	"Strongly Agree"	1.154*	0.918	0.914	0.906*
Someone like me can't really influence government decisions.	"Strongly Agree"	0.580*	0.795*	0.799*	0.811*
It is a citizen's duty to keep informed about politics even if it is time-consuming	"Strongly Agree"	1.039	1.026	1.039	1.055

\* Significant at the 95% level

**Table 2.** Survey questions odds ratios for the four logistic regression models with internet status as explanatory variable. (Continued)

Questions:	Probability modeled	Univariate	Basic Model	Intermediate Model	Full Model
<i>How often do you...during a week?</i>					
Watch national network news	"3 times or more"	0.832*	1.023	1.013	1.029
Watch local news programs on television	"3 times or more"	0.880*	1.102*	1.085*	1.052
Watch cable news networks	"3 times or more"	1.166*	1.228*	1.254*	1.256*
Search for news on the internet	"3 times or more"	3.667*	2.643*	2.745*	2.884*
<i>Do you have any friends or relatives who you know to be gay, lesbian, or bisexual?</i>	"Yes"	1.768*	1.416*	1.365*	1.474*
<i>Are you yourself gay, lesbian, or bisexual?</i>	"Yes"	1.103	1.120	1.168	1.662*
<i>Do you actively participate in any of the following types of organizations or groups?</i>					
Hobby, sports team, or youth group	"Yes"	2.091*	1.470*	1.499*	1.425*
Neighborhood association or community group	"Yes"	1.653*	1.342*	1.317*	1.280*
<i>Do you happen to have in your home or garage any guns or revolvers?</i>	"Yes"	0.766*	0.840*	0.798*	0.938
<i>Does anyone in the household own a pet?</i>	"Yes"	1.596*	1.383*	1.352*	1.231*

\* Significant at the 95% level

In Appendix A we present a more comprehensive table, which contains the 95% confidence intervals for the odds ratios, and two R squared measures for the full model. In Appendix B we report the question wording of the results.

### *Environment Questions*

There are ten questions about environmental-related behaviors during the past 12 months. For those questions, we found that the effect of having internet access at home on the survey response was significantly reduced, once demographic variables were controlled in the model. The odds ratio, however, varied little across the three multivariate models, which meant that the additional demographic information did not explain the association between having internet access and the environmental-related behaviors. For the six questions regarding recycling and energy use, the odds ratios are significant even in the full model.

## *Most Important Issue Facing the Country*

When asked to name the most important issue facing the country, for most of the issues (including "other"), the odds ratio of choosing it as the most important issue changed little among the four models. Exceptions were the issues "education" and "moral". For "education," the odds ratio in the multivariate models was much smaller than the one in the univariate model. One possible explanation is that education was controlled in the three multivariate models and respondents' education levels may be correlated with choosing 'education' as the most important issue. For "moral," the odds ratio was significantly reduced in the full model. Compared to the other two multivariate models (i.e., basic and intermediate model), the full model had additional controls on employment status, marital status, presence of kids in the households, and home ownership.

It is not clear how these four variables could theoretically be related to the consideration of "moral" as the most important issue facing the U.S.

## *Political Attitudes*

For the six statements on political attitudes, the probabilities of strongly agreeing with the statements were modeled in the logistic models. There was variation in the pattern of how the odds ratio changed among the models. For two of the statements, the estimated odds ratio was about the same among the models. This indicates that the association between internet access status and the political attitude does not depend on the demographic differences between people with internet access at home and those without home access. For the other four statements, the association between the internet access status and the responses were weakened in the multivariate models, compared to the univariate analysis. However, the extent of the association tended to be the same among the three multivariate models.

## *Media Use and News*

There were four questions used in the analysis that relate to how people obtain the news. The probability modeled in the logistic regression was based on using a particular news medium more than 3 times per week. The univariate analysis showed that individuals in internet households watched national network news programs and local

news programs less frequently than those in non-internet households. Moreover, this association disappeared in the multivariate models. In contrast, both the univariate and multivariate analysis showed that people with internet access at home had a higher probability of watching cable news than those without home access. One explanation for this is that internet households may be more likely to have cable services than non-internet households. The large odds ratios of "searching for news on the internet" between people of internet households and those of non-internet households was not surprising.

## *Homosexuality*

The estimated odds of having friends or relatives who were homosexual was higher among individuals with internet access than those without access. The differences in odds reduced when demographic characteristics were controlled in the model, but the differences were still significant. The situation was quite different for the question asking

the homosexuality of respondents themselves. It was the only question in the analysis for which odds ratios were observed to increase significantly in the full model. One speculation is that when the marital status was controlled in the full model, the positive association of internet access and self-perception of being homosexual was revealed, which was once neutralized by the association of the opposite direction through the marital status. Married people are less likely to view themselves as homosexual, and also tend to have internet access at home. Thus, people living in internet-access households appear to be less likely to be homosexual, if marital status is not controlled in the analysis.

### *Participations in Groups and Organizations*

For both of the questions regarding participation in groups and organizations, the odds of actively participating in the organizations were greater for people with internet access at home than those without access. The odds ratio was reduced in the multivariate models but remained to be significant. The readers may have noticed that this pattern of changes in odds ratios among the four models occurred for many of the questions analyzed in this study.

### **OTHER QUESTIONS**

The odds ratio of having guns or revolvers at home became insignificant in the full model. For the question on having a pet, the odds ratio reduced steadily as more demographic characteristics were controlled in the analysis.

### **DISCUSSION**

Our study was able to compare answers to the same questions between internet and non-internet households by using the same instrument (web survey). From the analysis on a set of attitudinal and behavioral questions, it appears that the difference between internet and non-internet households still exists even after controlling for the relevant demographic variables. By using odds ratios, the multinomial logistic regression clearly highlights the improvement in the model. The full model is generally able to reduce the odds ratios but not always to the point of getting close to 1. We remind the reader that in our definition, non-internet households have a WebTV that enables them to answer online surveys.

### **FUTURE WORK**

This study focused on a subset of attitudinal and behavioral questions. It is the goal of future studies to expand the variance and quantity of variables used in the analysis to portray a clearer and more extensive depiction of the persistent digital divide. It is also the goal of future studies to simulate this phenomenon with different weighting techniques to determine how weighting can compensate for the exclusion of non-internet households, while minimizing the detrimental impact on data quality (e.g., avoiding extremely high weights).

## Acknowledgements

We thank Elisa Chan and the Statistics team for having compiled the dataset and weighted the data. Tom Wells provided assistance for the surveys selected for the study. Erica Demme and Patricia Graham provided useful comments on the text increasing readability and clarity of the exposition.

## REFERENCES

- Callegaro, M. & DiSogra, C. (2008). Computing response metrics for online panels. *Public Opinon Quarterly*, 72, 1008-1032.
- Comley, P. (2007). Online market research. In ESOMAR (Ed.), *Market research handbook* (5th ed., pp. 401-420). Hoboken, NJ: Wiley.
- Couper, M. (2007). Issues of representation in eHealth research (with a focus on web surveys). *American Journal of Preventive Medicine*, 32, S83-S89.
- Current Population Survey. (2008). October 2007 Internet and computer use supplement. Retrieved August 2008, from [http://www.bls.census.gov/cps\\_ftp.html](http://www.bls.census.gov/cps_ftp.html)
- Department of Commerce (1995). Falling through the net: A survey of the "have nots" in rural and urban America. Electronic Article Retrieved August 2008 from <http://www.ntia.doc.gov/ntiahome/fallingthru.html>
- Department of Commerce. (2002). A nation online: How Americans are expanding their use of internet. Washington D.C.: National Telecommunications and Information Administration.
- Department of Commerce. (2004). A nation online: Entering the broadband age. Washington D.C.: National Telecommunications and Information Administration.
- Dever, J. A., Rafferty, A., & Valliant, R. (2008). Internet surveys: *Can statistical adjustment eliminate coverage bias?* *Survey Research Methods*, 2, 47-62.
- DiMaggio, P., Hargittai, E., Celeste, C., & Shafer, S. (2004). Digital inequality: From unequal access to differentiated use. In K. Neckerman (Ed.), *Social inequality* (pp. 355-400). New York: Russell Sage Foundation.
- Göriz, A. S. (2007). Using online panels in psychological research. In A. N. Joinson, K. Y. A. McKenna, T. Postmes & U.-D. Reips (Eds.), *The Oxford handbook of internet psychology* (pp. 473-485). Norfolk: Oxford University Press.
- Hargittai, E. (2008). The digital reproduction of inequality. In D. B. Grusky (Ed.), *Social stratification. Class, race, and gender in sociological perspective* (3rd ed., pp. 936-944). Boulder, CO: Westview Press.
- Horrigan, J. B. (2006). Home Broadband Adoption 2006. Electronic Article Retrieved August 2008 from <http://www.pewinternet.org/>
- Horrigan, J. B., Rainie, L., Allen, K., Boyce, A., Madden, M., & O' Grady, E. (2003). The ever-shifting Internet population: A new look at Internet access and the digital divide.

Electronic Article Retrieved August 2008 from  
[http://www.pewinternet.org/pdfs/PIP\\_Shifting\\_Net\\_Pop\\_Report.pdf](http://www.pewinternet.org/pdfs/PIP_Shifting_Net_Pop_Report.pdf)

Fairle, R. W. (2007). Explaining differences in access to home computers and the Internet: A comparison of Latino groups to other ethnic and racial groups. *Electronic Commerce Research*, 7, 265-291.

Fairle, R. W. (2008, November 20th). *The educational consequences of the digital divide*. Paper presented at the Annual lecture on science, technology & society. The digital divide: Why do we care?, Center for Human Potential and Public Policy. The Harris School. Chicago. Retrieved December 2008, from  
<http://harrisschool.uchicago.edu/centers/chppp/sts2008/>

Knowledge Networks. (2008). Home Technology Monitor. Retrieved September 2008, from <http://www.knowledgenetworks.com/htm/index.html>

Kruger, L. G., & Gilroy, A. A. (2008). Broadband Internet access and the digital divide: Federal assistance programs. *CRS Report for Congress*. Electronic Article Retrieved from <http://www.nationalaglawcenter.org/assets/crs/RL30719.pdf>

Martin, S. P. (2003). Is the digital divide really closing? A critique of inequality measurement in a Nation Online. *IT & Society*, 1, 1-13.

Martin, S. P., & Robinson, J. P. (2007). The income digital divide: Trends and predictions for levels of Internet use. *Social problems*, 54, 1-22.

Morgan, J. N., Solenberger, P. W., & Nagara, P. R. (2005, Last updated 3 January 2005). History and future of binary segmentation programs Retrieved June 2008, from [http://www.isr.umich.edu/src/smp/search/search\\_paper.html](http://www.isr.umich.edu/src/smp/search/search_paper.html)

Morgan, J. N., Solenberger, P. W., Van Eck, N. A., & Nagara, P. R. (2007). Search: Software for exploring data structure. Retrieved June 2008, from <http://www.isr.umich.edu/src/smp/search/>

National Telecommunications and Information Administration. (2008a). *Networked nation: Broadband in America 2007*. Washington D.C.: Department of Commerce.

National Telecommunications and Information Administration (2008b). Households using the Internet in and outside the home, by selected characteristics: Total, Urban, Rural, Principal City, October 2007. Electronic Article Retrieved October 2008 from [http://www.ntia.doc.gov/reports/2008/Table\\_HouseholdInternet2007.pdf](http://www.ntia.doc.gov/reports/2008/Table_HouseholdInternet2007.pdf)

Norris, P. (2001). *Digital divide. Civic engagement, information poverty, and the Internet worldwide*. Cambridge: Cambridge University Press.

Pew Center for the People and the Press (1995). Technology in the American household; Americans going online... Explosive growth, uncertain destinations. Electronic Article Retrieved August 2008 from <http://people-press.org/reports/pdf/136.pdf>

Pew Internet & American Life Project. (2009). Usage over time (Publication. Retrieved March 2009, from Pew Internet & American Life Project:  
[http://www.pewinternet.org/trends/UsageOverTime\\_Jan\\_07\\_2009.xls](http://www.pewinternet.org/trends/UsageOverTime_Jan_07_2009.xls)

Piekarski, L., Galin, M., Baim, J., Frankel, M., Augemberg, K., & Prince, S. (2008, May 15-18). *Internet access panels and public opinion and attitude estimates*. Paper presented at the 63rd Annual conference of the American Association for Public Opinion Research, New Orleans, LA.

Postoaca, A. (2006). *The anonymous elect. Market research through online access panels*. Berlin: Springer.

Prieger, J. E., & Hu, W.-M. (2008). The broadband digital divide and the nexus of race, competition, and quality. *Information Economics and Policy*, 20, 150-167.

Robinson, J. P., & Martin, S. P. (2005). IT and social change, 2000-2004: Behavioral and attitudinal evidence from the General Social Survey. *Webuse&Society*, 1, 1-33.

Robinson, J. P., Neustadtl, A., & Kestnbaum, M. (2002). The online "diversity divide": Public opinion differences among internet users and nonusers. *IT & Society*, 1, 284-302.

Rookey, B. D., Hanway, S., & Dillman, D. A. (2008). Does a probability-based household panel benefit from assignment to postal response as an alternative to Internet-only? *Public Opinion Quarterly*, 72, 962-984.

Sonquist, J. A., Baker, E. L., & Morgan, J. N. (1973). *Search for structure* (Revised ed.). Ann Arbor, MI: The Institute for Social Research, The University of Michigan.

Tortora, R. (2009). Attrition in consumer panel. In P. Lynn (Ed.), *Methodology of longitudinal surveys*. (pp. 235-249) Hoboken, NJ: Wiley.

Valliant, R., & Lee, S. (2005). Economic characteristics of internet users vs. nonusers and implications for web-based surveys. *Webuse&Society*, 1, 34-51.

## APPENDIX A.

**Table A.** Survey questions odds ratios for the four logistic regression models with internet status as explanatory variable. 95% confidence intervals for each model and Max rescaled R-square for the full model.

Questions:	Probability modeled	Univariate		Basic Model		Intermediate Model		Full Model				R <sup>2</sup>	Max-resc. R <sup>2</sup>		
		OR	95% CL	OR	95% CL	OR	95% CL	OR	95% CL	OR	95% CL				
<i>In the past 12 months, have you...</i>															
given money to an environmental organization	"Yes"	1.669*	1.490	1.870	1.065	0.943	1.203	1.054	0.933	1.191	1.094	0.967	1.238	0.056	0.104
contributed time or expertise to an environmental group	"Yes"	1.521*	1.260	1.836	1.076	0.882	1.313	1.076	0.882	1.313	1.138	0.931	1.391	0.014	0.048
recycled newspapers or other papers	"Yes"	1.993*	1.864	2.132	1.339*	1.242	1.443	1.327*	1.229	1.433	1.275*	1.180	1.379	0.124	0.176
recycled cans	"Yes"	1.642*	1.530	1.762	1.203*	1.114	1.299	1.197*	1.106	1.295	1.145*	1.057	1.241	0.080	0.118
recycled glass	"Yes"	1.911*	1.789	2.042	1.258*	1.170	1.353	1.281*	1.187	1.382	1.260*	1.166	1.360	0.149	0.201
recycled plastic	"Yes"	1.897*	1.775	2.027	1.275*	1.185	1.371	1.275*	1.182	1.376	1.256*	1.163	1.356	0.121	0.169
purchased a product specifically because it was made from recycled materials	"Yes"	1.551*	1.440	1.670	1.245*	1.150	1.347	1.197*	1.104	1.296	1.228*	1.132	1.332	0.052	0.072
taken steps to reduce the use of energy(oil, gasoline, electricity)	"Yes"	1.437*	1.343	1.538	1.188*	1.104	1.280	1.167*	1.084	1.257	1.149*	1.066	1.239	0.041	0.057
taken other actions to improve the environment	"Yes"	1.235*	1.070	1.426	1.061	0.911	1.236	1.049	0.900	1.222	1.083	0.928	1.264	0.009	0.024
Any of these actions	"Yes"	1.895*	1.716	2.093	1.330*	1.190	1.486	1.280*	1.143	1.432	1.231*	1.098	1.380	0.049	0.111
<i>In your opinion, what is the most important issue facing the U.S. today?</i>															
Taxes	"Yes"	0.936	0.753	1.165	0.892	0.706	1.127	0.951	0.752	1.203	0.913	0.720	1.158	0.005	0.026
Education	"Yes"	1.584*	1.327	1.892	1.296*	1.073	1.565	1.268*	1.049	1.531	1.272*	1.051	1.539	0.014	0.044
War on Terrorism	"Yes"	0.895*	0.820	0.977	0.915	0.833	1.004	0.941	0.856	1.034	0.919	0.835	1.011	0.014	0.023
Situation in Iraq	"Yes"	0.944	0.873	1.020	0.879*	0.809	0.956	0.871*	0.801	0.947	0.912*	0.838	0.993	0.015	0.023
Economy /Jobs	"Yes"	1.023	0.937	1.115	1.074	0.978	1.179	1.084	0.986	1.191	1.105*	1.005	1.215	0.024	0.040
Moral Issues	"Yes"	1.229*	1.111	1.359	1.233*	1.108	1.371	1.218*	1.095	1.356	1.085	0.974	1.210	0.021	0.039
Healthcare	"Yes"	0.886*	0.813	0.965	0.986	0.899	1.081	0.967	0.882	1.060	0.993	0.904	1.089	0.007	0.012
Other	"Yes"	0.987	0.876	1.112	0.938	0.826	1.066	0.945	0.832	1.074	0.985	0.865	1.120	0.009	0.021

\* Significant at the 95% level

**Table A.** Survey questions odds ratios for the four logistic regression models with internet status as explanatory variable. 95% confidence intervals for each model and Max rescaled R-square for the full model. (Continued)

Questions:	Probability modeled	Univariate			Basic Model			Intermediate Model			Full Model				
		OR	95% CL		OR	95% CL		OR	95% CL		OR	95% CL		R <sup>2</sup>	Max-resc. R <sup>2</sup>
<i>How interested are you in politics and public affairs?</i>	"Very Interested"	1.304*	1.211	1.403	1.105*	1.019	1.198	1.146*	1.057	1.243	1.189*	1.095	1.291	0.083	0.117
<i>Do you agree or disagree with the following statements about political affairs in this country?</i>															
Most public officials are not really interested in the problems of the average person	"Strongly Agree"	0.859*	0.794	0.930	1.097*	1.007	1.194	1.117*	1.026	1.217	1.117*	1.024	1.217	0.020	0.031
It is a citizen's duty to serve on a jury even if it interferes with his/her private life.	"Strongly Agree"	1.309*	1.218	1.407	1.077	0.997	1.164	1.089*	1.008	1.178	1.098*	1.015	1.188	0.045	0.062
It is a citizen's duty to report a crime even if it might put him or her in some jeopardy.	"Strongly Agree"	0.933	0.864	1.007	0.893*	0.823	0.970	0.903*	0.832	0.980	0.887*	0.817	0.965	0.015	0.023
It is a citizen's duty to pay taxes even if they seem unfair or too high.	"Strongly Agree"	1.154*	1.060	1.257	0.918	0.838	1.006	0.914	0.834	1.002	0.906*	0.826	0.994	0.022	0.035
Someone like me can't really influence government decisions.	"Strongly Agree"	0.580*	0.524	0.642	0.795*	0.712	0.888	0.799*	0.715	0.892	0.811*	0.725	0.907	0.017	0.039
It is a citizen's duty to keep informed about politics even if it is time-consuming	"Strongly Agree"	1.039	0.950	1.135	1.026	0.938	1.121	1.039	0.950	1.135	1.055	0.964	1.154	0.032	0.049

\* Significant at the 95% level

**Table A.** Survey questions odds ratios for the four logistic regression models with internet status as explanatory variable. 95% confidence intervals for each model and Max rescaled R-square for the full model. (Continued)

Questions:	Probability modeled	Univariate		Basic Model		Intermediate Model		Full Model							
		OR	95% CL	OR	95% CL	OR	95% CL	OR	95% CL	R <sup>2</sup>	Max-resc. R <sup>2</sup>				
<i>How often do you...during a week?</i>															
Watch national network news	"3 times or more"	0.832*	0.779	0.889	1.023	0.950	1.101	1.013	0.941	1.092	1.029	0.954	1.110	0.122	0.164
Watch local news programs on television	"3 times or more"	0.880*	0.822	0.943	1.102*	1.021	1.188	1.085*	1.006	1.171	1.052	0.974	1.137	0.083	0.113
Watch cable news networks	"3 times or more"	1.166*	1.088	1.250	1.228*	1.138	1.325	1.254*	1.162	1.353	1.256*	1.162	1.357	0.068	0.093
Search for news on the internet	"3 times or more"	3.667*	3.349	4.016	2.643*	2.404	2.907	2.745*	2.495	3.020	2.884*	2.618	3.176	0.079	0.109
<i>Do you have any friends or relatives who you know to be gay, lesbian, or bisexual?</i>															
	"Yes"	1.768*	1.654	1.889	1.416*	1.319	1.521	1.365*	1.270	1.467	1.474*	1.370	1.586	0.051	0.071
<i>Are you yourself gay, lesbian, or bisexual?</i>															
	"Yes"	1.103	0.910	1.338	1.120	0.910	1.377	1.168	0.948	1.438	1.662*	1.347	2.050	0.037	0.151
<i>Do you actively participate in any of the following types of organizations or groups?</i>															
Hobby, sports team, or youth group	"Yes"	2.091*	1.904	2.297	1.470*	1.331	1.622	1.499*	1.358	1.656	1.425*	1.289	1.574	0.042	0.063
Neighborhood association or community group	"Yes"	1.653*	1.498	1.824	1.342*	1.208	1.491	1.317*	1.185	1.463	1.280*	1.151	1.424	0.052	0.085
<i>Do you happen to have in your home or garage any guns or revolvers?</i>															
	"Yes"	0.766*	0.713	0.823	0.840*	0.777	0.907	0.798*	0.736	0.865	0.938	0.863	1.020	0.133	0.182
<i>Does anyone in the household own a pet?</i>															
	"Yes"	1.596*	1.471	1.673	1.383*	1.287	1.486	1.352*	1.258	1.453	1.231*	1.143	1.326	0.112	0.155

\* Significant at the 95% level

## APPENDIX B. QUESTION WORDING

*Many people also work to improve the environment. In the past 12 months, have you*  
[RANDOMIZE ALL EXCEPT "NONE OF THESE"]

Given money to an environmental organization .....	1
Contributed time or expertise to an environmental group .....	2
Recycled your newspapers or other papers .....	3
Recycled your cans.....	4
Recycled your glass.....	5
Recycled your plastic.....	6
Purchased a product specifically because it was made from recycled materials .....	7
Taken steps to reduce your use of energy (oil, gasoline, electricity) .....	8
Other .....	9
None of these [SINGLE SELECT].....	10

*In your opinion, what is the most important issue facing the U.S. today?*  
[RANDOMIZE ALL EXCEPT "OTHER"]

Taxes  
Education  
War on terrorism  
Situation in Iraq  
Economy/Jobs  
Moral issues  
Healthcare  
Other (please specify:\_\_\_\_)

*Are you personally acquainted with any of the following people? That is, if you met or called the person, would he or she recognize you or your name?*

A current member of the US Congress or Senate.....	1
A current member of your state legislature.....	2
A local government official.....	3
A reporter/editor for a local newspaper, radio or TV station.....	4
None of these [SINGLE SELECT].....	5

*In general, how interested are you in politics and public affairs?*

Very interested.....	1
Somewhat interested .....	2
Slightly interested.....	3
Not at all interested .....	4

Do you agree or disagree with the following statements about political affairs in this country? **[RANDOMIZE]**

	Strongly Agree	Agree	Neither Agree nor Disagree	Disagree	Strongly Disagree
--	----------------	-------	----------------------------	----------	-------------------

- Most public officials are not really interested in the problems of the average person
- It is a citizen's duty to serve on a jury even if it interferes with his/her private life
- It is a citizen's duty to report a crime even if it might put him or her in some jeopardy
- It is a citizen's duty to pay taxes even if they seem unfair or too high
- Someone like me can't really influence government decisions
- It is a citizen's duty to keep informed about politics even if it is time-consuming

*How often do you...*

	Three times a week or more	Every week or almost every week	One to three times a month	Less than once a month	Never
--	----------------------------	---------------------------------	----------------------------	------------------------	-------

- Watch national network news programs such as NBC Nightly News, ABC World News Tonight, or The CBS Evening News?
- Watch local news programs on television?
- Watch cable news networks such as CNN, MSNBC, or Fox News?
- Search for news on the Internet?

*Do you have any friends or relatives who you know to be gay, lesbian, or bisexual?*

- Yes, friends ..... 1
- Yes, relatives ..... 2
- Yes, both friends and relatives ..... 3
- No 4
- I would prefer to not answer this question ..... 3

*Are you yourself gay, lesbian, or bisexual?*

- Yes ..... 1
- No ..... 2
- I would prefer to not answer this question ..... 3

*Do you actively participate in any of the following types of organizations or groups?*

- Issue-oriented political organization ..... 1
- Non-partisan civic organization ..... 2
- School club or association ..... 3
- Hobby, sports team, or youth group ..... 4
- Neighborhood association or community group ..... 5
- Group representing racial/ethnic interests ..... 6
- None of these **[SINGLE SELECT]** ..... 7

*Do you happen to have in your home or garage any guns or revolvers?*

Yes .....1  
No .....2

*Does anyone in the household own a pet?*

Yes .....1  
No .....2