Do We Hear Different Voices?: Investigating the Differences Between Internet and non-Internet Users On Attitudes and Behaviors

Chan Zhang[1], Mario Callegaro[2], Melanie Thomas[3] & Charles DiSogra[4]

[1] University of Michigan, 426 Thompson Street, Room 4050 Ann Arbor, MI 48104

[2,3,4] Knowledge Networks, 1350 Willow Road, Ste 102, Menlo Park, CA 94025

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. Of the two types of web panels, probability-based panels are generally believed to produce data of higher quality than those from non-probability panels. The latter is prone to sampling coverage error: In only 71% of American households, at least one member uses the internet at any location (including at home), and only 62% of households have an internet connection from home, as of October of 2007. Moreover, some subgroups of the population such as African Americans and Latinos are more likely to be offline than others. This phenomenon is often referred to as the "digital divide." Recent data suggests that this divide may not be closing quickly enough to be dismissed by researchers seeking to study a representative sample of the population. As a solution, Knowledge Networks covers the ISP charges and provided MSN® TV web units in the past and now PC laptop computers 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 are able to participate in surveys online, minimizing the error resulting from the exclusion of non-internet users. In the article, 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 responses to a series of survey questions are analyzed using (multinomial) logistic regressions, to identify whether internet access status still makes a difference to survey results when controlling for relevant demographic variables. The analysis reveals that differences still exist between internet and non-internet households for a series of behaviors and other measures, even after controlling for the relevant demographic variables.

Key words: digital divide, non-internet households, probability-based online panel, KnowledgePanel

1. Background on the Digital Divide Debate

1.1 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 socio-demographic factors. In the same year, the National Telecommunication and Information Administration (NTIA) issued their first report using the Internet and the computer use data supplement of the Current Population Survey (CPS) published with an emblematic title: "Falling through the net: A survey of the 'have nots' in rural and urban America" (Department of Commerce, 1995). Since that time, NTIA and the Pew Center have continued to provide publicly available data on internet access in the U.S.

In hindsight, the year 2002 can now be seen as a turning point in the debate about the digital divide. The NTIA reports, up to that year, focused on inequality of access in terms of demographic factors and the title of almost every document started with "Falling through the net." (See DiMaggio et al. (2004) for a detailed literature review of digital divide studies published through 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 trend that more Americans were going online (Hargittai, 2008, p. 937).

A cursory read of the NTIA report could give the reader the impression that the internet had reached a universal level of access – that is, the impression that the "digital divide" had been successfully bridged. However, if the reader were to peruse the data in the report, it would be evident that the ideal of equal access to the internet had not been attained. The strongest critique came from Steven Martin (2003) who reanalyzed the CPS data used by NTIA, pointing out that computer and internet access had spread more slowly 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, on the one hand, and the richer households, on the other. 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.

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 (U.S. Census Bureau, 2009). 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. Analysis done on this new dataset show again evidence of the persistence of the digital divide among minorities, less educated and lower-income persons, and households with language barriers (e.g. Latinos) (Fairle, 2008).

1.2 Latest Trends
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?" 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 had increased somewhat to 71% and 61.7%, respectively (U.S. Census Bureau, 2009).

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). The question is as follows: 'Do you use the internet, at least occasionally? and "Do you send or receive emails at least occasionally?" The most recent estimates were that 79% of U.S. adults went online in April 2009. Because of the inclusive question wording (occasional and from no specific location) used in their survey, we interpret the Pew data as the upper bound for the measurement of internet usage in the U.S.

Lastly we provide data from the Knowledge Networks HomeTechnology Monitor™ (Knowledge Networks, 2009). Data were collected every year in the spring with a RDD telephone survey. (The study is now conducted on KnowledgePanel®.) 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)?" According to the final 2009 survey, the internet penetration at a household level was 64%.

2. The Digital Divide as Coverage Error For Internet Surveys

Online panels can be classified as probability-based and non-probability-based or volunteer panels. In the first case, every panel member has a known probability of selection based on a sampling frame made of telephone numbers or addresses. By definition, non-probability volunteer 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 MSN® TV web unit or more recently a PC laptop in order to participate in surveys online.

2.1 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 using an area probability sampling frame, for their analysis. They classified internet users as people who used the internet from any location including MSNTV, 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 the "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 be more trusting of others, to be less likely to accept explanations of poverty based on personal factors such as lack of willpower, to be more supportive of women working outside the home, to be more likely to accept more homosexuality, to be more confident in science, and more likely to rate their health as excellent in comparison to non-internet users (see Table 10 on page 24).

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. Overall internet households reported better health and were more health conscious than non-internet households. 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 without any internet access; respondents with access but no use in the past month; low internet users (≤ 5 times a day); and high internet users (≥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 disproportionally 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, non-probability web 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, however, 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 all'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 in our study 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).

2.2 About KnowledgePanel®

Knowledge Networks at its inception and until recently selected households using random digit dialing (RDD) sampling methodology and more recently with an Address Based Sample methodology (DiSogra & Callegaro, 2009). The sampling is done without replacement, to ensure that numbers already fielded by Knowledge Networks do not get sampled again. Once a household is contacted 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 MSNTV set-top box or more recently a laptop computer 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 year with a specific computer usage and technology profile survey.

3. Research questions

1. Whether the differences in responses between internet and non-internet users can be all explained by the demographic differences between those two subpopulations, and how the conclusions would be related to the survey topics.

2. Whether adjustments on more demographic characteristics lead to smaller differences in responses between the internet and non-internet users.

4. Methods

The data was collected from the KnowledgePanel Core Adult demographic profile survey and the Public Affair profile survey from late 2007 to April 2008. The survey invitations were sent to all the panel members during that period. The Core Adult profile survey collected basic demographic information for each panel member, and was a requirement to join KnowledgePanel. The Public Affairs profile survey included various topics such as religion, politics, and communities. The completion rate of the Public Affairs profile survey is 92.5% among all the panel members, based on the method described by Callegaro & DiSogra (2008).

To answer the first research question (i.e. whether the differences in responses between internet and non-internet users can be all explained by the demographic differences), we need to control for as many as demographic characteristics as possible, as measured in the Core Adult profile survey. To answer the second research question (i.e., whether adjustments on an increasing number of demographic characteristics lead to smaller differ-

ences in responses between the two group), we need to vary the control of the demographic variables.

We used logistic regressions to model the survey responses. The differences between internet and non-internet users were measured by the odds ratio associated with the internet access. For most of the questions with more than two response options, the categories are recoded to facilitate analysis. Thus, responses to those questions are recoded into two categories in order to fit the logistic model. One exception is the question asking respondents to choose the most important issue facing the U.S. today. The question has eight options, such as tax, education, health care, and etc. The question was excluded from the analysis, as it is not appropriate to recode data into a dummy variable. Follow-up questions were also dropped. The sample sizes for those questions were usually much smaller than those of the other questions, which made it difficult to compare the results across the questions. Moreover, we also decided to exclude the questions where demographic characteristics could have dominant impacts on the answers. An example is the question on the participations of women's group. For those questions, it may be more appropriate to include interaction terms in the models or restrict analysis to certain subpopulation. 102 response variables are available from the survey. The selection criterion described above gives us a total of 81 response variables for the analysis.

For each survey question, we fitted four regression models with increasing control of the demographic variables.

a) 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.
b) In the basic model we controlled for the most influential variables on internet status: income, age, education, and ethnicity.
c) In the intermediate model, gender, region and metropolitan status were added as predictors.
d) 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 11 explanatory variables including the indicator of home internet access.

5. Results

5.1 Differences between internet and non-internet HH by topics using the full model

The results discussed below are from the full regression model, in which 11 demographic variables are controlled in addition to the home internet access. All the effects reported below are significant if not otherwise stated. The impacts of internet access are discussed by the question topics.

Voting and Attitudes towards Bush: There are two types of questions related to voting. For the questions on voting actions (e.g. whether or not vote in the 2004 presidential election), there are more reports of such behaviors among internet households (hereafter, "HHs") than non-internet HHs. For the questions with regard to voting outcomes (e.g.

which candidate did you vote for in the 2004 presidential election), the answers do not appear to depend on internet access status.

The survey also asked whether respondents think of themselves as republican, democrat, or independent. There was no significant difference between internet and non-internet users in their responses.

Political attitudes and acquaintance with politically influential people: Respondents are asked to rate six statements from strongly agree to strongly disagree, which are designed to measure the attitude towards a series statement on citizen's duty and role in politics (e.g. One statement is "it is a citizen's duty to serve on a jury even if it interferes with his/her private life."). Related to those are questions asking whether respondents are personally acquainted with certain people with different level of political influences, ranging from a current member of the US Congress to a reporter or editor for a local newspaper. As it is very likely that respondents' political attitudes are related to their acquaintance with politically influential people, those two types of questions are discussed together.

For all the six attitudinal questions, there are significant differences between internet HHs and non-internet HHs, when the percentages of "strongly agree/agree" are examined. Compared to the respondents in the non-internet HHs, those in the internet HHs are more likely to believe that ordinary people can influence government decisions and more likely to agree with individual sacrifice for the public interests (e.g. serving on a jury even if it interferes with one's private life), and more likely to believe that most public officials are not really interested in the problems of the average person.

For three of the four acquaintance questions, asking about the respondents' acquaintance with a current member of the US Congress, a local government official, and a reporter/editor for a local newspaper respectively, respondents with internet access at home are more likely to have such acquaintance than those without. There is no significant difference between the two types of households in terms of the acquaintance with a current member of their state legislator.

TV watching of news and political programs: Respondents are first asked about their frequency of obtaining the news through four different channels, including national networks news program, local news programs, cable news networks, and internet. It is not surprising to find that "search for news on the internet" is the outcome variable on which the internet access status has the biggest impact among all the questions analyzed in this study (Odds Ratio=3.019). The analysis also shows that respondents with home internet access tend to watch cable news networks more frequently than those without. The frequency of watching national and local new programs seems to not depend on the internet access status.

Followed are 11 questions on the frequency of watching particular nightly or weekly political TV programs. The significant impact of internet access is only found on the frequency of watching comedy central.

Civic and political activities: The survey includes a series of questions on various civic and political activities. Examples include donating blood, giving money to a charity, contacting a government official, working for a campaign. There are also questions on environment related behaviors, such as taking steps to reduce use of energy, and recycling certain types of materials.

For 18 of the 19 questions on civic and political activities, having internet access at home is found to be significantly associated with higher involvement in the activity. Among those 18 questions where the internet access makes a difference, relatively large impacts are observed as measured as odd rations "OR." Internet HHs were 1.6 times more likely to "having contacted a government official" and 1.8 times more likely to "having written a letter to the editor" in the past 12 months.

With regards to the environment related behaviors, respondents from internet HHs are more likely to report the environmental friendly behaviors in 6 of the 8 questions. In particular, the impacts are significant for all the four recycling questions (papers/cans/glass/plastic).

Miscellaneous factual questions: The survey has questions on whether respondents themselves are gay, lesbian, or bisexual and whether they have friends or relatives that are not heterosexual. For both of the questions, respondents with internet access at home are significantly more likely to say "Yes" than those without. Internet HHs were almost 1.5 times more likely to say there are gay, lesbian, or bisexual, and 1.6 times more likely to report to have friends or relatives that are not heterosexual.

The survey also collects information on how often respondents attend religious services, whether they have any guns, whether they belong to a labor union and whether they have ever served in army. There is no significant impact of home internet access on those questions.

Miscellaneous opinion questions: There are a couple of questions involving respondents' self-assessments, including whether they think of themselves as liberal, whether they would describe themselves as born-again or evangelical Christians, and whether they would describe themselves as an environmentalists. The regression analysis shows that respondents with home internet access are more likely to describe themselves as liberal, and less likely to describe themselves born-again or evangelical Christians than those without. Whether respondents would call themselves environmentalists seem to not depend on the access status.

5.2 Do differences between internet and non-internet HH become smaller with more control of demographic characteristics?

In this study, each survey variable was analyzed using four regression models with increasing control of the demographic characteristics. If the assumption that more control of demographics lead to smaller differences between web users and non-web users is correct, we could expect that:

Pattern A: If the difference between internet and non-internet HH is not significant in the univariate analysis, it will not be significant in the other models where demographic variables are added.

Pattern B: If the difference between internet and non-internet HH is significant in the univariate analysis, it may become insignificant as more demographic variables are added.

Among 81 response variables analyzed in the study, 70 were observed to have the patterns A or B. That is, for majority of the responses, there was no evidence that more con-

trol of demographics could lead to larger differences between internet and non-internet users, although more control did not always help to reduce the differences. Moreover, pattern B was dominant among the two regular patterns (66 out of70). The rest of the 11 questions had irregular patterns. For some outcome variables, the differences between internet and non-internet HH were not significant in the univariate analysis. But the effects became significant when demographics variables were controlled into the model. In some other cases, effects showed up in the univariate analysis, disappeared with the inclusion of the demographic variables in the model, and then came back in the full regression model. Those irregular patterns indicate that increasing the control of demographics may lead to larger differences between internet and non-internet users and thus a larger bias in adjusted estimates.

For most of the survey variables with irregular patterns, the differences between internet and non-internet users did not change very much across the model, as indicated by the small variation of odds ratio. However, dramatic change in the odds ratio is observed for the question asking about respondents' sexual orientation. In this case, the odds ratio are around 1.1 and insignificant in the first three regression models, but in the model with additional control on 4 demographics (i.e. employment status, marital status, presence of kids in the household, and home ownership), the internet users were significantly more likely to report that they consider themselves gay, lesbian, or bisexual than the non-internet users (odds ratio=1.65, p-value<.0001).

For the survey responses where adjustments on demographic characteristics were shown to be helpful in reducing the bias (Pattern B), we examined the effects of additional control of demographics on the bias reduction. To do so, we calculated for each regression model the average odds ratio among all the 66 responses belongs to pattern B. We then examined how the average odds ratio changed across the four models. We found that on average, the differences in responses between the online and the offline respondents were reduced substantially when demographics are controlled in the model. More surprisingly, we found that controlling only 4 basic demographics is on average as effective as controlling 11 demographics characteristics. The average odds ratios associated with internet access are 1.99 (univariate analysis: internet access only), 1.62 (income, age, education, and ethnicity added) 1.61 (gender, region and metropolitan status added) and 1.58 (employment status, marital status, presence of kids in the household, and home ownership added).

For space limitations we cannot report a table with the results of the 81 variables we discussed above. Instead we report some variables to give the reader an idea of the magnitude of change in the odds ratios based on the four models tested.

Table 1. Selected survey questions odds ratios for the four logistic regression models with internet status as explanatory variable.

Questions:	Probabili-ty modeled	Univar.	Basic Model	Intermed. Model	Full Model
In the past 12 months, have you…?					
Given money to an environmental or-ganization	"Yes"	1.669*	1.065	1.054	1.094
Contributed time or expertise to an en-vironmental group	"Yes"	1.521*	1.076	1.076	1.135
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*	1.261*
Recycled plastic	"Yes"	1.897*	1.275*	1.275*	1.258*
Purchased a product specifically be-cause it was made from recycled mate-rials	"Yes"	1.551*	1.245*	1.197*	1.227*
Taken steps to reduce the use of ener-gy(oil, gasoline, electricity)	"Yes"	1.437*	1.188*	1.167*	1.149*
Do you agree or disagree with the following statements about political affairs in this country?					
Most public officials are not really in-terested in the problems of the average person	"Strongly Agree"& "Agree"	1.047	1.242*	1.257*	1.256*
It is a citizen's duty to serve on a jury even if it interferes with his/her private life.	"Strongly Agree""& "Agree"	1.757*	1.248*	1.271*	1.249*
It is a citizen's duty to report a crime even if it might put him or her in some jeopardy.	"Strongly Agree""& "Agree"	1.529*	1.260*	1.272*	1.231*
It is a citizen's duty to pay taxes even if they seem unfair or too high.	"Strongly Agree""& "Agree"	1.676*	1.233*	1.222*	1.190*
Someone like me can't really influence government decisions.	"Strongly Agree""& "Agree"	0.754*	0.895*	0.899*	0.907*
It is a citizen's duty to keep informed about politics even if it is time-consuming	"Strongly Agree""& "Agree"	1.666*	1.307*	1.319*	1.298*

* Significant at the 95% level

Table 1. Selected survey questions odds ratios for the four logistic regression models with internet status as explanatory variable. (Continued)

Questions:	Probabili-ty modeled	Univar.	Basic Model	Intermed. Model	Full Model
How often do you...during a week?					
Watch national network news	"3 times or more" &"Every week or almost every week"	0.832*	1.025	1.012	1.026
Watch local news programs on television		0.823*	1.042	1.027	0.994
Watch cable news networks		1.232*	1.254*	1.283*	1.282*
Search for news on the internet		3.872*	2.774*	2.892*	3.019*
Do you have any friends or relatives who you know to be gay, lesbian, or bisexual?	"Yes"	1.692*	1.387*	1.334*	1.449*
Are you yourself gay, lesbian, or bisexual?	"Yes"	1.089	1.102	1.149	1.645*
Do you actively participate in any of the following types of organizations or groups?					
Hobby, sports team, or youth group	"Yes"	2.091*	1.470*	1.499*	1.426*
Neighborhood association or community group	"Yes"	1.653*	1.342*	1.317*	1.281*
Do you happen to have in your home or garage any guns or revolvers?	"Yes"	1.305*	1.191*	1.253*	1.064

* Significant at the 95% level

5. 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 observing no differences between the two groups. We remind the reader that in our definition, non-internet households have a MSNTV that enables them to answer online surveys. Our results are consisted with previus findings (e.g. Rookey, Hanway, & Dillman, 2008) providing evidence on the need to include non-internet households in order to have a full picture of the U.S. population.

Acknowledgements

References

Callegaro, M. & DiSogra, C. (2008). Computing response metrics for online panels. *Public Opinon Quarterly*, 72, 1008-1032.

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.

Disogra, C. & Callegaro, M. (2009) Computing response rates for probability-based web panels. *Proceedings of the Joint Statistical Meetings, Survey Research Methods section [Cd-Rom]*. Alexandria, VA: American Statistcal Association.

Disogra, C. & Callegaro, M. (2009) Recruiting probability-based web panel members using an address-based sample (ABS) frame: Results from a pilot study conducted by Knowledge Networks. *Proceedings of the Joint Statistical Meetings, Survey Research Methods section [Cd-Rom]*. Alexandria, VA: American Statistcal Association.

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., 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

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. (2009). Home Technology Monitor. Retrieved July 2009, from http://www.knowledgenetworks.com/htm/index.html

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

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/

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). Demographic of internet users. Retrieved July 2009 from http://www.pewinternet.org/Static-Pages/Trend-Data/Whos-Online.aspx

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.

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.

U.S. Census Bureau. (2009). Internet Use Triples in Decade, Census Bureau Reports. Retrieved June 2009, from http://www.census.gov/Press-Release/www/releases/archives/communication_industries/013849.html

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