UNITED STATES DEPARTMENT OF COMMERCE
Economics and Statistics Administration
U.S. Census Bureau
Washington, DC 20233-0001

November 21, 2012

2012 AMERICAN COMMUNITY SURVEY RESEARCH AND EVALUATION REPORT MEMORANDUM SERIES #ACS12-RER-31

DSSD 2012 AMERICAN COMMUNITY SURVEY MEMORANDUM SERIES #ACS12-MP-06

MEMORANDUM FOR	ACS Research and Evaluation Steering Committee
From:	Anthony G. Tersine***/Signed/*** Assistant Division Chief, Decennial Statistical Studies Division
Prepared by:	Rachel Horwitz ACS Data Collection Methods Branch Decennial Statistical Studies Division
Subject:	Data Quality Assessment of the American Community Survey Internet Response Data

Attached is the final American Community Survey Research and Evaluation report "Data Quality Assessment of the American Community Survey Internet Response Data." The data quality analysis uses data from the April 2011 American Community Survey Internet Test. The objective of the analysis is to determine whether Internet respondents provided comparable data to mail respondents. Specific data quality indicators examined in this report are outliers, rounded values, correlations between household income and property value and personal income and education, and a comparison of response error using a computer assisted telephone reinterview. This report summarizes the results of the data quality analysis.

If you have any questions about this report, please contact Rachel Horwitz at 301-763-2834 or Jennifer Tancreto at 301-763-4250.

Attachment

cc:

ACS Research and Evaluation Team
Debbie Griffin (ACSO)
Todd Hughes
Agnes Kee
Andrew Roberts
Brian Wilson
Barbara O'Hare (DIR)
Nancy Bates
Frank Vitrano
Mary Ann Chapin (20RPO)
Justin McLaughlin
Tony Tersine (DSSD)
Michael Bentley
Mary Davis
Steven Hefter
Joan Hill
Samantha Barron
Jennifer Tancreto
Mary Frances Zelenak
David Johnson (SEHSD)
Bob Kominski
Colleen Hughes (POP)
Anne Ross
John Studds (ASD)
Glenn Eanes
Joe Mistichelli

American Community Survey Research and Evaluation Program
November 21, 2012

Data Quality Assessment of the American Community Survey Internet Response Data

FINAL REPORT

Rachel Horwitz, Jennifer Guarino Tancreto,
Mary Frances Zelenak, Mary C. Davis
Decennial Statistical Studies Division

United States™ Census Bureau

Intentionally Blank

Table of Contents

LIST OF TABLES

LIST OF FIGURES

EXECUTIVE SUMMARY

Objective

This report compares the quality of Internet data collected during the 2011 April American Community Survey (ACS) Internet Test to mail return data from the same test. This analysis is intended to help determine whether the Internet provides comparable data to mail in terms of expected relationships between variables and response error.

Methodology

We used two basic strategies to assess the quality of the data received from the ACS Internet Test respondents: data quality indicators about the types of responses received and measures of response error. Specifically, we used outliers, the percentage of rounded values for numeric income entries, the correlation between household income and property value, and the correlation between education and income as data quality indicators. We also used data from a follow-up reinterview, which re-asked both mail and Internet respondents a pre-determined set of ACS questions to measure response error. We attempted to control for characteristic differences between mail and Internet respondents using propensity weights.

Research Questions and Results

How do data quality indicators compare between Internet and mail?

Mail and Internet responses had about the same percent of outliers for personal and household income. While rounding is common across both modes, Internet responses did have a slightly higher percentage of rounded values when compared with mail responses in most cases. However, the difference in percent of rounded entries is generally quite small and is not cause for concern.

There were no concerning findings in the correlations between related questions as well. The correlation between household income and property value questions for mail and Internet responses is not statistically different when the data are top-coded using ACS rules. The correlation between personal income and education questions is significantly lower for top-coded mail responses than Internet responses. This difference is due to a higher percent of negative mail responses compared to Internet responses. Whether this difference exists because Internet respondents have higher incomes than mail respondents or because the "loss" check box is used differently between the modes is an area for future research.

Is response error comparable between mail and Internet returns?

We measured response error using gross difference rates. We pre-selected 15 questions that represent the various Internet question response formats. There were 87 different comparisons of response error across modes. Only three of the 87 comparisons produced significant differences between mail and Internet, suggesting similar response error properties across modes. The response error was lower for Internet than mail in two of the three significant differences.

In addition to there being few significant differences in response error between the modes, the response errors themselves were very low across most estimates. The highest gross difference rate (GDR) for both modes was for one ancestry region reported in the Ancestry question (12.5 percent for Internet and 15.5 percent for mail). However, across both modes, only 21.3 percent of the GDRs are higher than five percent and 5.7 percent are higher than 10 percent.

1. BACKGROUND

1.1 Description of the American Community Survey and April 2011 Internet Test

The American Community Survey (ACS) is a mixed mode, mandatory survey of housing units.[1] The Census Bureau samples about 3.5 million housing unit addresses in the ACS each year. Since its inception, most sampled housing units receive a questionnaire in the mail. If they do not complete the questionnaire in the first month and we have a phone number for the address, they are contacted for a Computer Assisted Telephone Interview (CATI). A sample of the addresses that have still not completed the survey after the CATI month are then visited by a field representative to conduct a Computer Assisted Personal Interview (CAPI).

The survey consists of a series of demographic questions, a series of questions about the housing unit, and detailed questions that are asked for each person in the household, one person at a time. The survey takes approximately 38 minutes to complete, on average. The actual length is dependent on the household size and the number of questions that are applicable, based on skip patterns.

In 2000, the ACS tested the use of the Internet as an additional response mode. Researchers found that offering the Internet as a response option during the mail phase actually decreased the overall response rate, and that very few respondents completed the questionnaire on the Internet (Griffin *et al.*, 2001). Since 2000, technological advances have been instrumental in the trend towards becoming a paperless society. To that end, Internet use has become more common as people use it for shopping, financial transactions, gathering information, and general communication. In the survey world, declining response rates have inspired survey organizations to investigate the use of these new technologies, specifically the Internet, to collect data.

The April 2011 ACS Internet Test evaluated the feasibility of providing an Internet response option to sampled addresses. The overall purpose of this test was to determine the best methods for informing people about the ACS Internet response option and encouraging them to respond. Results from the April Internet Test suggest that response rates do not suffer due to offering an Internet option, and are actually increased for some notification strategies (Tancreto *et al.* 2012). This positive finding will result in an Internet option in ACS production in 2013[2]. This necessitates the need to look at the response data from the Internet to ensure it compares to the quality of data received from mail respondents, which is the purpose of this report.

[1] Group Quarters are included in the ACS but are not part of this report.

[2] An additional Internet Test was conducted in November of 2011 to determine which notification strategy will be used in 2013 ACS production. Details on that test can be found in Matthews *et al.*, 2012.

1.2 Comparing Data Quality Between Internet and Mail Cases in the April 2011 ACS Internet Test

Many surveys are turning to the Internet because it is a relatively inexpensive mode of data collection. Additionally, returns come back faster than mail surveys, which can (in some surveys) reduce reminder mailings and speed up data processing and analysis. While the Internet offers many benefits, we need to ensure the data obtained from the Internet are comparable to, or better than, those obtained through mail or the overall quality of the survey estimates will suffer.

There are several indicators researchers use to assess the quality of their data. Some of the most common are response rates, item nonresponse, response speed, equivalence of response, and response error (Tuten *et al.* 2000). Tancreto and her colleagues' 2012 report included both response rates and item nonresponse rates for the April 2011 ACS Internet Test. However, there are additional factors we can examine to get a better sense of the quality of the Internet data, as compared to mail.

Rounded values for numeric entries can be one indicator of poor data quality. Specifically, they can signal satisficing (Couper 1997). Couper used respondents' answers to a series of questions to classify them as satisficers or other respondents. He found that the satisficers were significantly more likely to use rounded numbers than other respondents were. Additionally, using the Current Population Survey, Fricker and Tourangeau (2010) used rounded values, among other indicators, to determine whether respondents with a low response propensity across survey rounds were more likely to have lower quality data. As expected, they found that the percentage of rounded values increased as response propensity decreased. Therefore, if the percentage of Internet responses with rounded values is higher than the percentage found on mail returns, it could suggest Internet respondents are not taking the task as seriously as mail respondents.

Another factor we can consider to determine if we are receiving similar quality across modes is whether expected correlations exist in responses received in both modes. Specifically, there are well-documented relationships across ACS responses such as the correlation between education and income and between income and property values (Ashenfelter and Rouse 1999, U.S. Census Bureau). If these relationships do not hold across modes, there may be an issue with the quality of the data and more research may be needed to determine where the differences lie.

Although we did not find any research on extreme values and outliers in Internet surveys as compared to mail surveys, the results discussed in Section 4.1.3 and 4.1.4 from the correlation analysis led us to believe this is another important indicator. If Internet reports have more outliers and extreme values that do not reflect true values, it could suggest an issue with the instrument or respondents are not answering the questions as honestly as mail respondents are.

Finally, it is also important to analyze response error to ensure consistency in reporting between data sources when the quantities being measured are defined in a consistent way across sources (Brackstone 1999). Mode can impact how respondents read a question and what they believe is expected of them. For example, a larger text box can result in longer write-in responses (Couper *et al.* 2001). Additionally, respondents using a mail form do not need to proceed through the questions in order, whereas Internet respondents are forced to by design. This too can affect how different questions are interpreted and how carefully they are read.

While it is necessary to measure data quality for a new mode to ensure that any mode benefits do not come at the cost of quality, it is especially important to check data quality for the ACS because monthly data are accumulated over months and years. Adding a new mode with low data quality could negatively affect trends between the year before the new mode is introduced and the following years. Therefore, we need to ensure the data we are receiving are consistent and differences are a result of actual respondent data, not the mode by which the data are collected. While there are many ways to measure data quality, we think that the indicators described in this section give a preliminary look at the issue and will help identify some potential sources of error.

2. METHODOLOGY

2.1 Sample Design

2.1.1 *April 2011 ACS Internet Test Sample Design*
The April 2011 ACS Internet Test was designed to simulate a typical one-month mail data collection period in the ACS. Therefore, we selected a national sample of 120,000 households for this test.

In the April 2011 ACS Internet Test, we stratified tracts into two groups: Targeted and Not Targeted. The Targeted group consists of tracts containing households that we expected to use the Internet at a higher rate based on past research. The balance of tracts was placed into the Not Targeted group. At the time the sample was selected, we suspected these groups varied by age, education, and computer experience. Additionally, we used four types of notification strategies and two general approaches–choice and push–to inform sample members of the Internet option. For the choice strategies, respondents received the Internet option instructions and paper form at the same time. Alternatively, in the push strategies, sample members received the Internet invitation first and then later received a mail form if they had not responded online. For more information on the Targeted and Not Targeted groups and the notification strategies, please see Tancreto *et al*. (2012). This analysis does not explicitly differentiate between the two strata or between the notification strategies. However, sampling weights are applied to account for the probability of selection.

2.1.2 *Content Reinterview Analysis Sample Design*
The sampling frame for the reinterview analysis was mail and Internet respondents to the April 2011 ACS Internet Test with a small sample selected from the ACS production respondents for the April 2011 panel. Only cases with a phone number (either vendor

supplied or respondent provided) were included in the frame. The sample was selected from two[3] subsets of the 2011 April ACS Internet Test respondents. Eligibility required a valid name and phone number.

(1) **Internet respondents** – Occupied housing units that either submitted an Internet response indicating they had completed the interview or had answered enough questions to be considered a sufficient partial without officially submitting.

(2) **Mail respondents** – Occupied housing units that returned a mail questionnaire, with at least two demographic questions answered (not including the name fields).

In addition, a sample of 140 mail respondents in the ACS April 2011 production panel was selected as the control. People classified as Internet respondents and mail respondents are not directly comparable due to differing definitions of a response. This limitation is discussed in greater detail in Section 3.2.

In total, 5,255 Internet and 4,501 mail respondents were sampled for the reinterview. There were two cut dates for the sample, allowing both early and late respondents to be part of the sample.

2.2 Analysis Methods

2.2.1 *Propensity Weighting Adjustment*

We know that the demographics differ for those who respond using the Internet than respond using mail (Tancreto *et al*, 2012). Therefore, comparisons across mode could reflect differences in the population and not actual differences in the data quality. For example, we know that younger people are more likely to respond using the Internet and are more likely to round than mail respondents (Fricker and Tourangeau, 2010). If the Internet responses have more rounded values than mail responses, this is likely a result of the respondents being younger than an issue with the Internet mode.

Therefore, to help control for differences in mail versus Internet respondents, we identified a set of variables from the ACS that are predictive of whether a respondent will answer by mail or Internet (age, sex, marital status, building type, and urban/rural status[4]). Using these variables, and what mode each respondent used, we created two logistic propensity models: one that predicted the likelihood of being an Internet respondent and the other predicted the likelihood of being a mail respondent. We then used the inverse of these propensities to weight the estimates of the Internet and mail respondents, respectively. In general, the propensity weights were low with a minimum adjustment of 1.0 and a mean of 2.0. However, there was a handful of large adjustments; with a maximum of 19.6. This weighting adjustment was used for all the analyses

[3] The reinterview sample design included a sample of nonrespondents to the 2011 ACS Internet Test, but these cases are not included in this analysis because we do not have a set of measures from the first interview.

[4] These calculations only include respondents who had a valid answer for all five variables.

discussed in this report and should help correct some of the inherent differences between the mail and Internet populations and focus on the differences in the data themselves.

2.2.2 *Data Quality Indicators Analysis*

The universe for the data quality indicators analysis (rounded values and correlations) included any respondent who accessed the April 2011 ACS Internet Test instrument or responded to the survey by mail, across all test panels (excluding production). While these households did not necessarily need to answer every question in the survey or submit the survey to be included in this analysis, they did need to answer the following items: property value, education, and the series of personal income questions for every household member.

We used four basic data quality indicators that provide a preliminary assessment of the quality of Internet and mail responses: outliers, rounded values, the correlation between responses to property value and household income, and the correlation between education and income. We discuss the analysis methods for each separately.

Outliers

The first data quality indicator we examined was outlier values for personal income and household income. We compared the number and percent of mail and Internet entries that were less than -$250,000 and greater than $500,000 for personal and household income.

Rounded Values

We analyzed the percentage of rounded values for wages, self-employment income, and interest income[5]. For each income category, we calculated the percentage of entries that were rounded to the hundreds and thousands, along with the associated standard errors. We then calculated the differences in the percentage of rounded values between mail and Internet responses and tested whether the differences were statistically significant.

Correlations

In order to analyze the specified correlations, we needed to ensure the universes for mail and Internet responses were the same. Specifically, the Internet instrument uses skip logic so respondents do not receive questions that do not apply to them. However, on the mail form, respondents may answer questions that they should have skipped. Therefore, at the household level, we removed households from the analysis that indicated they did not own their residence but provided a property value and households with individuals under age 15 with wages income, because we could not know their true household income. At the person level, we removed individuals that met the criteria to create equivalent universes:

- Individuals under age 15 with wages income, and
- Individuals that had completed 12th grade, but did not indicate whether they received a high school diploma or a G.E.D.

[5] Additional income variables could not be used because of errors on the data file. However, these errors do not affect the total income values used in the outlier and correlation analyses.

Correlations were calculated using Proc Corr in SAS®. We calculated correlations using top-coded values based on ACS rules. This was done because of a fundamental difference between the mail and Internet versions of these questions, which is discussed at greater length in Section 3.1. Finally, we used the propensity weighting adjustment described in Section 2.2.1 to help control for differences between Internet and mail respondents.

2.2.3 *Content Reinterview Analysis*

The reinterview was designed to determine whether there are differences in levels of response error by data collection mode. A subset of 15 ACS questions (shown in Appendix A) were re-asked using ACS production CATI wording (and all CATI skips) to look at consistency of responses and minimize context effects. The questions were selected to represent every Internet question format type, such as radio buttons, check boxes, open-ended text boxes, and drop-down menus. Additionally, questions were chosen that would be less likely to have recall bias and time specific responses (e.g. "during the past week").

The follow-up interviews were only conducted with the original respondent from the ACS Internet Test. We conducted the follow-up interview within 3 weeks of the original response date and asked the original respondent questions about all persons in the household. After the third attempt to reach the original respondent failed, the interview was not conducted.

The overall response rate for the reinterview was 86.4 percent. Specifically, the response rate for Internet respondents was 86.6 percent, whereas the response rate for mail respondents was 86.2 percent. These response rates are not significantly different.

In order to assess the difference in response error across modes, we calculated gross difference rates (GDRs). This measure is a measure of response variation between the response from the original response and the response from the reinterview. We used Table 1 and the following formula to calculate the GDR.

Table 1. Cross-classification of Original Response by Follow-up Response

Follow-up Response (reinterview)	**Original Response**		
	Yes	*No*	*Total*
Yes	a	b	a+b
No	c	d	c+d
Total	a+c	b+d	n = a+b+c+d

$$GDR = \frac{b+c}{n}$$

We calculated the GDRs separately for mail and Internet respondents to compare the GDRs between the two modes. We used both sample weights and replicate weights so

the sampled population represented the population of Internet and mail respondents in the Internet Test universe. Additionally, because the respondents were not randomly assigned to the mail or Internet condition, we again used the propensity weights that were described in Section 2.2.1.

3. LIMITATIONS

3.1 Universe – Data Quality Indicators

Total household income was calculated by summing total individual income over all household members 15 years or older. Since income responses are needed from all household members to calculate a total, households where total income was missing for one or more household members were excluded. This eliminated 1.7 percent of households from the universe of inference.

Additionally, there was a truncation issue with the Internet output due to a processing error for several of the income variables. Specifically, social security, retirement, public assistance, other income and supplemental security income were all truncated so the last digits were missing. While these five variables were eliminated from the analysis, the household members were still included for the rounded value analysis of the other types of income.

Total income was also not used in the rounded value analysis because the instrument calculated most total income automatically by summing previously entered values. Therefore, if a respondent rounded for the components of their total income, they necessarily will have a rounded total income value. Due to this dependency, we did not think total rounded income was a useful measure.

Finally, the numeric mail responses (personal income and property value) went through an editing process where high values are top-coded. There was no similar process for Internet responses. Therefore, during analysis we top-coded all of the Internet responses following the ACS rules to make the responses from the two modes comparable.

3.2 The Definition of a Response

It is important to define what is considered a response. This definition is slightly different if the household replied by mail or Internet. For mail returns, if the case returned a non-blank[6] form or completed an interview through Telephone Questionnaire Assistance (TQA), it is accepted as a response. We assume that mail respondents have completed as much information as they are willing to provide by sending back the form or ending the TQA interview.

It is a little more complicated to interpret the intent of Internet respondents who started the survey but did not complete it (i.e. break-offs). They may have deliberately left the

[6] ACS operations consider a form to be non-blank even if there is only minimal information provided, specifically, a phone number or name of a household member. This definition originates from the Failed Edit Follow-Up operation, where these cases can be contacted by phone to collect the missing data.

survey, timed out, or forgot to or could not (due to login problems) return to complete the survey. We chose to classify Internet returns into three groups—Complete, Sufficient Partial, and Insufficient Partial responses—based on how far the respondent got in the survey, using rules originally developed for the CATI/CAPI modes. Insufficient Partial responses listed at least one household member but did not complete the housing questions while Sufficient Partial responses started answering the detailed person questions after the housing questions.

These definitions differ because mail cases with minimal data are sent to the Failed Edit Follow-Up procedure (FEFU)[7]. This procedure does not exist for Internet responses, so stricter rules were used to classify a return as a response. However, this does limit the comparability between modes.

3.3 Content Reinterview as a Replication

This analysis may be limited by the use of different modes for the interview and reinterview. Using different data collection modes for the original interview and follow-up can result in mode effects, especially because the original interviews were conducted in different modes (mail versus Internet). The effect of the change from mail to CATI and from Internet to CATI may affect respondents differently, resulting in a violation of replications. This can contribute to differences in answers between the original interview and the reinterview, which will in turn affect the response error measurement.

The use of GDRs assumes that the reinterview is a replication of the original interview. However, because the reinterview was conducted in a different mode than the original interview, we expected some violation of this assumption. Therefore, we used net difference rates (NDRs) to test whether the assumption holds for all of the comparisons we analyzed. We found high NDRs for the following variables: whether the household has a mortgage, insurance included in mortgage, Ancestry ("other" and "North American" regions), and Health Insurance purchased directly from an insurance provider and Other Health Insurance. Both modes violate this assumption for all of these variables. This violation needs to be considered when drawing conclusions about the response error for these variables.

3.4 Propensity Weighting

While we attempted to correct for the differences between Internet and mail respondents through propensity weighting, there are more complex techniques, such as propensity score matching, that may do a better job of accounting for respondents self-selecting their response mode. However, we opted for a simpler analysis due to a lack of time.

[7] The April 2011 ACS Internet Test did not have FEFU. However, we used the same respondent definitions that were defined when FEFU was present.

4. RESEARCH QUESTIONS AND RESULTS

4.1 How do Data Quality Indicators Compare between Internet and Mail?

As an assessment of data quality, we compared the number of outliers and percentage of rounded income values between Internet and paper responses. Additionally, we compared the correlation between household income and property value as well as between education and total personal income.

4.1.1 *Numeric outliers*

One indicator of poor data quality is outliers. We focused on outliers for personal income and household income for this analysis. Outliers were considered incomes less than -$250,000 and greater than $500,000. In both mail and Internet, the percent of responses that were outliers were not significantly different for both personal income and household income; 0.35 percent of mail and 0.31 percent of Internet responses were outliers for personal income and 0.76 percent of mail and 0.73 percent of Internet responses were outliers for household income. This result suggests that Internet is performing comparably to mail and that outliers are not a large concern in either mode.

4.1.2 *Rounded Dollar Values*

To assess the percent of rounded income values, we compared the frequency of entries rounded to the hundreds and thousands to the total number of values entered that were greater than 100 and 1,000, respectively, for the following variables: wages income, self-employment income, and interest income. Table 2 shows the percentage of rounded values (to the hundreds and thousands) for Internet and mail, along with the corresponding differences and standard errors.

Table 2. Percent of Rounded Income Values[8]

	Income Type		
Hundreds	**Wages**	**Self-Employment**	**Interest**
Percent Rounded Internet	84.7	80.8	68.1
(SE)	(0.3)	(1.1)	(0.7)
Percent Rounded Mail	82.9	78.2	62.4
(SE)	(0.3)	(0.9)	(0.6)
Internet - Mail (%)	1.8*	2.7*	5.7*
(SE)	(0.4)	(1.2)	(1.0)
Thousands	**Wages**	**Self-Employment**	**Interest**
Percent Rounded Internet	75.2	68.9	55.8
(SE)	(0.4)	(1.3)	(1.0)
Percent Rounded Mail	73.7	67.7	47.9
(SE)	(0.4)	(0.9)	(0.7)
Internet - Mail (%)	1.6*	1.2	7.9*
(SE)	(0.6)	(1.5)	(1.2)

Source: U.S. Census Bureau, American Community Survey Internet Test, April to May 2011
*Statistically different at the $\alpha = 0.10$ level.

The first thing to note is that the majority of respondents rounded their answer, regardless of which mode they used to complete the survey, which suggests this is a problem in both modes. Comparing the two modes, all of the differences in the percent of rounded income values between mail and Internet responses were significant[9] for rounding to thousands, except self-employment income. As Table 2 shows, the percent of rounded values for mail responses is consistently lower than for Internet responses, although the differences were generally small, especially for wages and self-employment.

There are many plausible explanations as to why Internet responses were more likely to be rounded. It is possible, that as the satisficing literature suggests, Internet respondents are not as invested in the task as mail respondents are. Alternatively, there could be a fundamental difference in how people think of Internet surveys as compared to paper forms. Finally, there may be differences in the types of people who respond by Internet. These respondents may be more likely to round regardless of what mode they use[10]. Although the propensity weighting adjustment attempted to correct for this, we used a relatively simple model. These rates were also calculated without the propensity adjustment and the differences were larger. Therefore, it appears this rounding phenomenon may be related to the type of people who respond by Internet, and a more

[8] Other income types were excluded from this analysis due to the limitations described in 3.1. Additionally, we excluded Total Income because we felt this variable was not comparable across modes.
[9] Differences were tested using a two-sided test by weighting the estimate of rounded values by the Internet Test base weights and replicate weights.
[10] We hypothesized that younger respondents would be more likely to round. The data do not support this hypothesis. For wages, younger respondents were more likely to round their responses on the mail form, but older respondents were more likely to round their responses on the Internet instrument.

complex weighting adjustment might be more likely to capture this.

We also thought the notification strategy might have an impact on the percentage of rounded values (the strategies were described in Section 2.1). Therefore, we compared the percent of rounded values from Internet responses from the Push treatments to mail responses from the Push treatments, mail responses from the Choice treatments, and Internet responses from the Choice treatments. We used the Push Internet as a basis for comparison because it is the strategy that will be used in ACS production in 2013. The largest difference in percent of rounded values for each income type was between Internet and mail responses from the Push strategies, which ranged from 4.6 percent for wages to 13.1 percent for interest income. In both cases, Internet responses had more rounded values. On the other hand, there was no difference between Push and Choice Internet responses for interest income and the difference was only 2.3 percent for wages. Differences for self-employment income ranged from 4.7 percent to 8.1 percent. Although we see that Internet respondents in the Push treatments generally round more, rounding is still a common behavior across all respondents.

4.1.3 *Correlations*

Another measure of data quality is whether known relationships exist in both modes. Specifically, we examine the correlation between property value and total household income and the correlation between personal income and education. Higher property values are generally associated with higher household incomes and, in general, higher levels of education result in higher income (U.S. Census Bureau and Ashenfelter and Rouse, 1999). These relationships should hold regardless of which mode respondents use to complete the survey. If the correlations are stronger for one mode than the other, it could suggest a potential data quality issue and more research would be needed to determine if the issue is related to mode.

Due to the limitation discussed in Section 3.1, where mail responses were top-coded but Internet responses were not, we top-coded the Internet responses to create comparable universes. Specifically, property values greater than $9,999,999 were set to $9,999,999 and any personal income value greater than $9,999,999 was set to $9,999,999. Additionally, any personal income value greater than $9,999,999, which was used to calculate household income, was set to $9,999,999. While additional editing is done before releasing final ACS numbers, this analysis provides a look at the top-coded raw data in both modes. Table 3 provides the correlations for both property value and house hold income and personal income and education.

Table 3. Correlation between Property Value and Household Income and Personal Income an Education (with top-codes)[11]

	Internet	Mail	Test: mail and Internet significantly different?
Property Value and Household Income	0.30	0.31	No, Z = 0.56
Personal Income and Education	0.23	0.13	Yes, Z = 16.24

Source: U.S. Census Bureau, American Community Survey Internet Test, April to May 2011

[11] Z-scores of greater than 2.58 were considered significantly different at the α=0.10 level.

Looking at the "Property Value and Household Income" row in Table 3, the correlation for Internet and mail are not statistically different. On the other hand, the personal income and education correlation is significantly different. The mail correlation is lower because 0.58 percent of mail incomes were negative compared to 0.30 percent of Internet incomes[12]. Additional research can inform whether this difference reflects a true difference in incomes across modes or whether negative values are entered differently in the two modes, even though the presentation of the question is the same in both modes.

4.2 Is the response error comparable between mail and Internet returns for a pre-determined set of ACS questions?

To determine whether response error was comparable between mail and Internet returns, a sample of April 2011 ACS Internet Test respondents (some mail, some Internet) were contacted to participate in a content reinterview. We used GDRs to measure the response error for mail and Internet respondents between their original answers and their answers to the CATI reinterview. We compared GDRs between modes for 15 questions. Within each question, we compared the GDRs for each response category, resulting in 87 total comparisons. Overall, GDRs were low for both modes for all of the comparisons tested. The largest mail GDR was 15.5 percent and the largest Internet GDR was 12.5 percent, both for the "other" ancestry region. Additionally, 21.3 percent of the GDRs, across both modes, were greater than five percent and only 5.7 percent were greater than 10 percent. Therefore, response error is low overall for the questions analyzed in this study.

In addition to looking at the overall error rate, we also wanted to compare the response error across mode. For this analysis, we tested the hypothesis that mail and Internet GDRs are the same at the $\alpha = 0.10$ level using a two-sided test[13]. Out of 87 comparisons, only three were significantly different. Specifically, there were significant differences in whether the household has a mortgage (for households that answer 'No' to having a mortgage), in whether insurance is included in the mortgage payment, and ancestry (for Western European ancestries). Each of these questions will be discussed in the subsequent sections.

The three questions with significant differences in the response categories are discussed in the following sections, whereas the GDRs for the rest of the variables, separated by question format, are presented in Appendix B. For the questions for which there were no significant differences, there is no clear trend in terms of which mode produces the nominally lower GDR.

[12] Difference is significant at the α=0.05 level with a Z-score of 5.51.

[13] Variables with three or four response options used a two-sided test in which the error rate has been controlled using the Bonferroni multiple comparison method at the $\alpha = 0.10$ level. Variables with more than four response options used the Bonferroni-Holm multiple comparison method at the $\alpha = 0.10$ level.

4.2.1 *Have a Mortgage*

The mortgage question asks respondents whether they own their home with a mortgage, have a contract to purchase, or do not have a mortgage (Figures 1 and 2).

Figure 1. Have a Mortgage – Mail Version

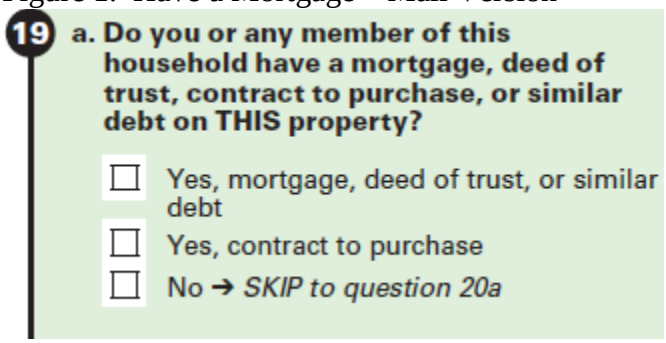
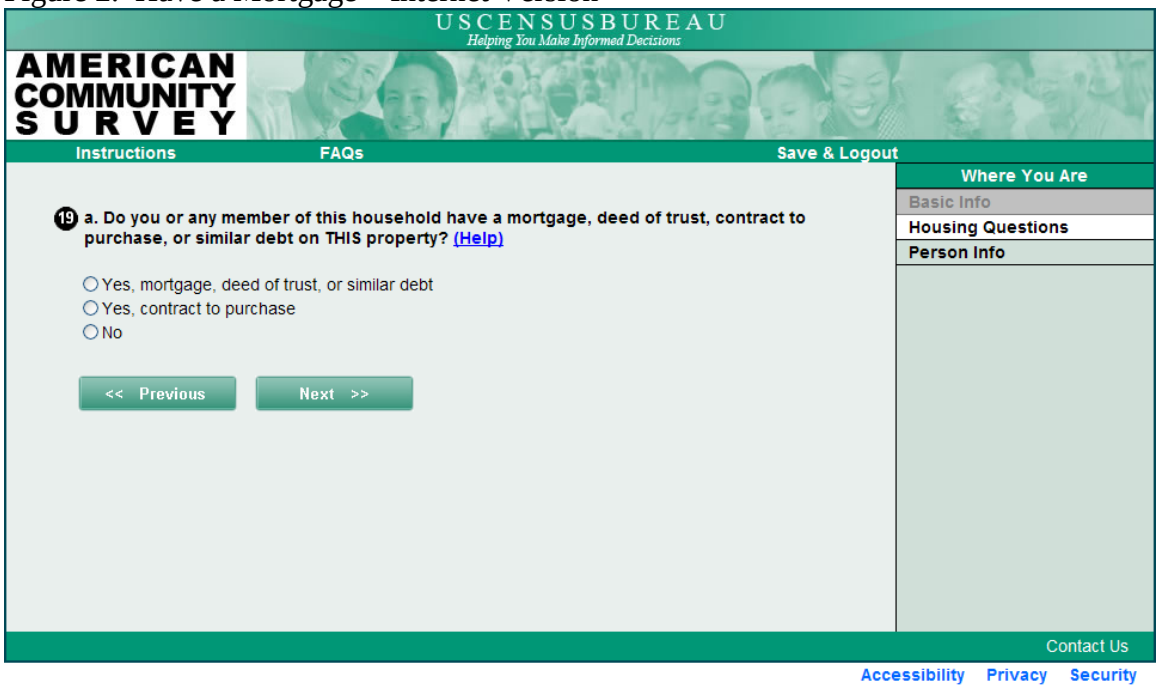


Figure 2. Have a Mortgage – Internet Version



The universe for the question is people who are homeowners or people who are purchasing a home. Therefore, since mail respondents did not receive the automatic filter, we applied the universe used for Internet to the mail cases. Table 4 provides the GDRs for "have a mortgage" by mode.

Table 4. Gross Difference Rates for Have a Mortgage by Mode

	Internet		Mail					
Have a Mortgage	Estimate (%)	Std Error (%)	Estimate (%)	Std Error (%)	Internet - Mail (%)	Std Error (%)	Significant?	Performed better?
Unweighted sample size	(n=1,894)		(n=1,027)					
Yes, mortgage, deed of trust, or similar debt	10.0	(0.9)	7.6	(1.1)	2.4	(1.4)	No	
Yes, contract to purchase	2.1	(0.5)	1.9	(0.6)	0.1	(0.6)	No	
No	8.7	(0.8)	6.1	(0.9)	2.6	(1.2)	Yes	Mail

Source: U.S. Census Bureau, American Community Survey Internet Test, April to May 2011

Mail has a significantly smaller GDR than Internet for the "No" mortgage response option. There are two primary differences in the presentation of this question that could influence responses between modes: the Help option on the Internet version and the skip pattern on the mail form. It seems unlikely that the skip pattern would result in lower response error because respondents looking to skip questions probably would not answer similarly in the follow up when they did not know about the skip. On the other hand, the Help option should affect all response options equally, not just the "No" option. Therefore, it is unclear why these rates are different other than differences in the sample of respondents selected. Additionally, this comparison violated the assumption that the reinterview was a replication of the original interview (as discussed in Section 3.3), so this difference may not be meaningful even though it is significant.

4.2.2 *Insurance Included in Mortgage Payment*

The insurance included in mortgage payment question asks respondents if their monthly mortgage payment includes payments for fire, hazard, or flood insurance (Figures 3 and 4).

Figure 3. Insurance Included in Mortgage Payment – Mail Version

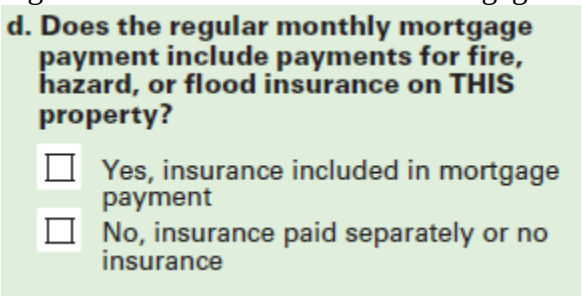
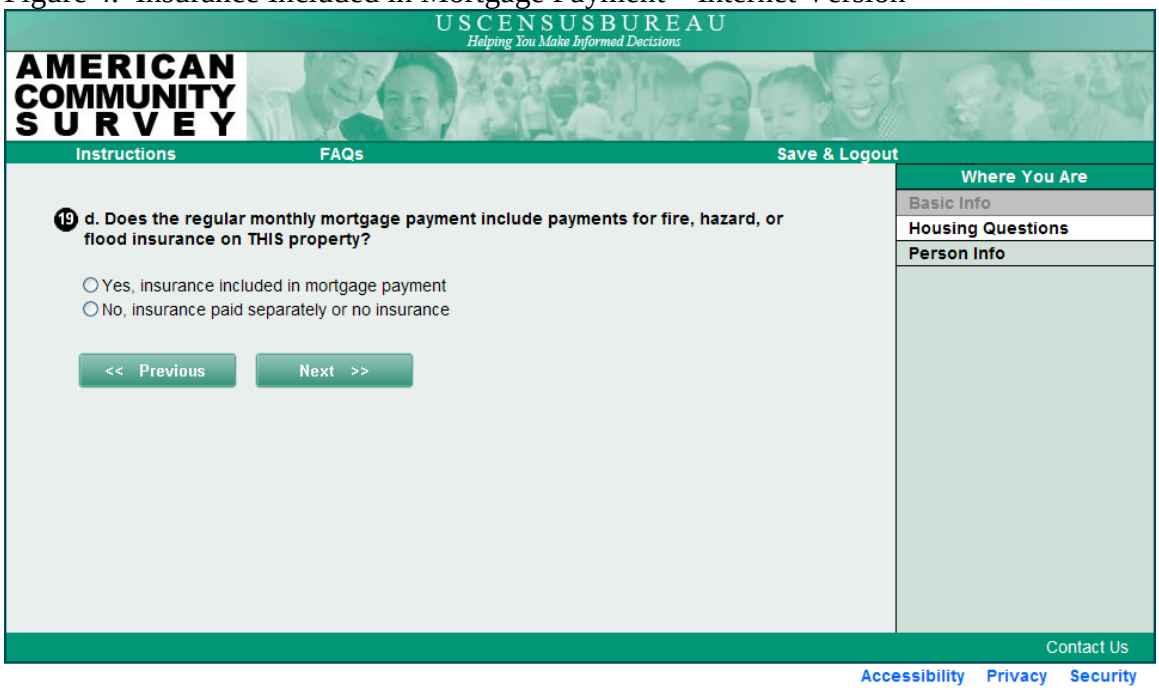


Figure 4. Insurance Included in Mortgage Payment – Internet Version



Respondents answering on the Internet and in the reinterview only received this question if they responded previously that they had a mortgage. Because mail respondents did not receive this filter, we applied the same universe so the results would be comparable (i.e. renters who inadvertently answered this question were removed from the analysis). Table 5 provides the GDRs for the mortgage insurance question by mode.

Table 5. Gross Difference Rates for Insurance Included in Mortgage Payment by Mode

	Internet		Mail					
Insurance included in mortgage	Estimate (%)	Std Error (%)	Estimate (%)	Std Error (%)	Internet - Mail (%)	Std Error (%)	Significant?	Performed better?
Unweighted sample size	(n=1,492)		(n=792)					
Insurance Included?	10.2	(0.9)	14.2	(1.7)	-4.0	(1.7)	Yes	Internet

Source: U.S. Census Bureau, American Community Survey Internet Test, April to May 2011

The GDR for Internet respondents is significantly lower than for mail respondents. It is not clear why Internet would perform better than mail for this question. The question has the same appearance in both mail and Internet. It is possible that because respondents are already at a computer, it is more convenient to look up payment plan information to answer the question. However, this is purely speculation and there is no obvious reason why this difference exists. Additionally, as with the have a mortgage question, this

comparison also violated the assumption that the reinterview was a replication of the original interview.

4.2.3 *Ancestry*

The ancestry question, in both modes, asks respondents to provide their (or a household member's) ancestry or ethnic origin in a text field (Figures 5 and 6).

Figure 5. Ancestry – Mail Version

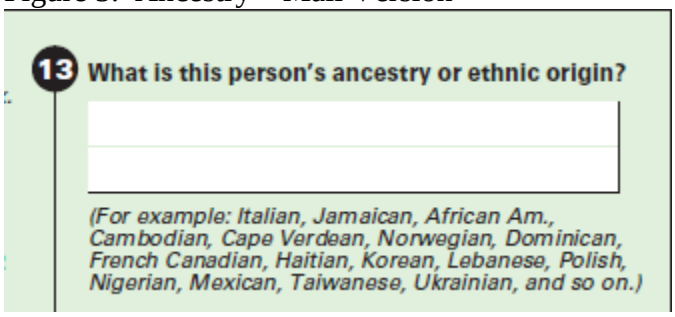


Figure 6. Ancestry – Internet version

The universe for this question was every individual in the household. This analysis only used the first ancestry listed, although respondents could list as many ancestries as they wanted in the write-in field. We limited the analysis to the first ancestry because most respondents only provided a single response, so missing data was very high for the second variable. In addition, the analysis can become complicated if respondents answer in a different order in the follow up than they did in the original interview. Although this can happen when using only the first ancestry, this limitation helps simplify a potentially complicated analysis. In order to calculate the GDRs for ancestry, the ancestry codes

were grouped into 12 regional ancestry groups. There was one region where Internet performed better than mail (Table 6).

Table 6. Gross Difference Rates for Regional Ancestries by Mode

	Internet		Mail					
Ancestry	Estimate (%)	Std Err (%)	Estimate (%)	Std Err (%)	Int - Mail (%)	Std Err (%)	Significant?	Performed better?
Unweighted sample size	(n=3,483)	(X)	(n=2,797)					
African	1.4	(0.4)	0.6	(0.2)	0.8	(0.5)	No	
Asian	1.4	(0.3)	0.7	(0.4)	0.7	(0.4)	No	
Eastern European	2.6	(0.6)	2.7	(0.5)	-0.2	(0.7)	No	
Northern European	9.5	(0.7)	12.6	(1.0)	-3.1	(1.3)	No	
Southern European	1.3	(0.3)	2.0	(0.4)	-0.7	(0.4)	No	
Western European	7.8	(0.8)	11.6	(1.0)	-3.8	(1.4)	Yes	Internet
Caribbean	0.7	(0.3)	0.2	(0.1)	0.5	(0.3)	No	
Central American	1.0	(0.2)	1.0	(0.3)	0.1	(0.4)	No	
South American	0.3	(0.1)	0.1	(0.1)	0.2	(0.1)	No	
North American	9.3	(1.0)	9.1	(0.9)	0.2	(1.3)	No	
Oceanian[14]	0.0	(0.0)	0.0	(0.0)	0.0	(0.0)	No	
Other	12.5	(0.9)	15.5	(1.2)	-3.0	(1.5)	No	

Source: U.S. Census Bureau, American Community Survey Internet Test, April to May 2011

While the response error seems relatively balanced across modes, the types of responses we receive vary by mode. Specifically, 10.6 percent of Internet respondents entered multiple ancestries. On the other hand, 30.3 percent of mail respondents provided multiple entries.[15] This is likely due to the design of the mail questionnaire (Figure 5). The text box following the question has a faint horizontal line through the middle of the box, suggesting two items should be entered. On the other hand, the text box on the Internet instrument is just a standard, open box (Figure 6). By design, it is taller than the other text boxes in the instrument and included arrows at the end to try to mimic the two-line format in mail. However, these features did not seem to have the desired effect of obtaining similar responses across modes or minimizing mode effects. Therefore,

[14] Estimates are less than 0.05 percent.
[15] On the CATI reinterview, 15.5 percent of respondents reported dual ancestries.

respondents actually provide more information on the mail form than the Internet, even though the GDRs suggest they are comparable.

Additionally, there is an inherent issue with GDRs and multiple entries; respondents have to list the ancestries in the same order in the reinterview as they did on the paper or Internet form in order to count as a match. For example, a respondent could have listed "Spanish" and "Norwegian." However, in the reinterview, they told the interviewer "Norwegian" and then "Spanish." This would not count as a match when calculating GDRs. It is possible that this is more of an issue for Western European countries because there are more dual ancestries from this region due to the countries' close proximity to each other[16]. Additionally, it is possible people do not show a preference of country because they are all European, so order effects are more likely to occur.

To test this hypothesis, the GDRs were recalculated after limiting the universe to respondents who only reported one ancestry. This adjustment resulted in no significant differences between mail and Internet. This result suggests that either respondents are reporting their ancestries in a different order in the reinterview than they did on the original survey resulting in a mismatch, even though the data are consistent, or the data reported in the original survey and the reinterview are inconsistent.

5. SUMMARY

We used several indicators to assess whether the quality of data obtained from Internet respondents is similar to that obtained from mail respondents. Specifically, we looked at income outliers, the percent of rounded values for income entries, correlations, and differences in response error across modes. Overall, we found similarities between Internet and mail.

An analysis of numeric entries showed limited differences between Internet and mail responses. The percent of income outliers was similar across modes and while there are more rounded values in Internet, the majority of respondents rounded regardless of their response mode. Additionally, Internet responses did not have higher response error than mail in most cases and looking at the analyses overall, the two modes appear to be equivalent.

Based on the data quality indicators examined in this report, there is no cause for concern about the quality of the data collected via the Internet instrument as compared to the mail questionnaire. While there are some things we can look into further, such as rounded numeric entries, negative income responses, and Internet respondents not providing as much ancestry information, the Internet appears to provide a comparable level of quality as mail.

[16] The percent of individuals with dual ancestries from Western Europe is statistically higher than the percent of individuals with dual ancestries from other regions.

Acknowledgements

We would like to thank the following Census Bureau staff for their valuable contributions and assistance to the development and analysis of this project: Amy Lauger, Brenna Matthews, Michelle Ruiter, Andrew Roberts, Brian Wilson, Cynthia Ramsey, Tony Tersine, John Studds, Chris Butler, Joe Misticelli, Brian Ridgeway, Don Keathley, Gail Denby, Michael Coan, Fred Meier, and Samantha Barron.

REFERENCES

Ashenfelter, O. and Rouse, C., (1999). Schooling, Intelligence, and Income in America: Cracks in the Bell Curve. National Bureau of Economic Research. Cambridge, Ma.

Brackstone, G. (1999). Managing Data Quality in a Statistical Agency. Statistics Canada, Survey Methodology, Catalogue No. 12-001-XPB, 25(2).

Couper, M., Traugott, M., Lamias, M. (2001). Web Survey Design and Administration. *Public Opinion Quarterly* 65(2): 230-253.

Couper, M. (2000). Web Surveys: A Review of Issues and Approaches. *Public Opinion Quarterly* 65(4): 464-494.

Couper, M. (1997). Survey Introductions and Data Quality. *Public Opinion Quarterly* 61(2): 317-338.

Fricker, S., Tourangeau, R. (2010). Examining the Relationship Between Nonresponse Propensity and Data Quality in Two National Household Surveys. *Public Opinion Quarterly* 74(5): 934-955

Griffin, D., Fischer, D.P., Morgan, M.T. (2001). Testing an Internet Response Option for the American Community Survey. Presented at the American Association of Public Opinion Research. Montreal, Quebec, Canada.

Matthews, B., Davis, M.C., Tancreto, J.G, Zelenak, M.F., Ruiter, M. (2012). 2011 American Community Survey Internet Tests: Results from the Second Test in November 2011. http://www.census.gov/acs/www/Downloads/library/2012/2012_Matthews_01.pdf

Tancreto, J.G., Zelenak, M.F., Davis, M., Ruiter, M., Matthews, B. (2012). 2011 American Community Survey Internet Tests: Results from the First Test in April 2011. 2012 American Community Survey Research and Evaluation Report Memorandum Series #ACS12-RER-13-R2, DSSD 2012 American Community Survey Memorandum Series #ACS12-MP-01-R2, June 11, 2012. http://www.census.gov/acs/www/Downloads/library/2012/2012_Tancreto_01.pdf

Tuten, T. Urban, D., Bosnjak, M. (2000). Internet Surveys and Data Quality: A Review. *Online Social Sciences*.

U.S. Census Bureau. Net Worth and Asset Ownership Tables: 1984 – 2010. http://www.census.gov/hhes/www/wealth/publications.html

U.S. Department of Commerce (2011). Digital Nation: Expanding Internet Usage. National Telecommunications and Information Administration. February 2011.

Appendix A

Questions for CATI Reinterview

Table 7. Question Descriptions and Types Used in CATI Reinterview

Question Description	Internet Question Type
Tenure	Mark One
Rent amount	Dollar write-in
Meals included	Yes/No
Have a mortgage	Mark One
Mortgage amount	Dollar write-in
Real Estate tax included	Yes/No
Insurance included	Yes/No
Place of Birth	Mark one plus character write-in
Citizenship/Year of Naturalization	Mark one plus numeric write-in
Year of Entry	Numeric write-in
School Enrollment	Mark One
Grade Enrolled	Mark one plus numeric write-in
Educational Attainment	Mark one plus numeric write-in but also has the headers
Ancestry	Character write-in
Health Insurance	Yes/No - Multiple parts

Appendix B

Gross Difference Rates for Reinterview Questions by Question Type for Questions Without Any Significant Findings

Question Type: ***Yes/No***

Table 8. Gross Difference Rates for Meals Included

	Internet		Mail				
Meals Included	Estimate (%)	Std Error (%)	Estimate (%)	Std Error (%)	Internet - Mail (%)	Std Error (%)	Significant?
Unweighted sample size	(n=665)		(n=430)				
Yes/No	0.7	(0.4)	0.7	(0.5)	0.0	(0.6)	No

Source: U.S. Census Bureau, American Community Survey Internet Test, April to May 2011

Table 9. Gross Difference Rates for Real Estate Tax Included

	Internet		Mail				
Real Estate Tax Included	Estimate (%)	Std Error (%)	Estimate (%)	Std Error (%)	Internet - Mail (%)	Std Error (%)	Significant?
Unweighted sample size	(n=1,504)		(n=797)				
Yes/No	5.5	(0.7)	6.1	(1.3)	-0.6	(1.4)	No

Source: U.S. Census Bureau, American Community Survey Internet Test, April to May 2011

Question Type: ***Yes/No Multiple Parts***

Table 10. Gross Difference Rates for Health Insurance by Mode

	Internet		Mail				
Health Insurance	Estimate (%)	Std Err (%)	Estimate (%)	Std Err (%)	Int - Mail (%)	Std Err (%)	Significant?
Insurance purchased through an employer	4.4 n=6,858	(0.4)	5.5 n=3,948	(0.6)	-1.1	(0.7)	No
Insurance purchased directly	12.4 n=4,719	(0. 9)	14.1 n=2,654	(1.2)	-1.7	(1.3)	No
Medicare	1.7 n=4,822	(0.3)	1.7 n=3,168	(0.3)	0.0	(0.4)	No
Medicaid	4.4 n=4,515	(0.8)	4.1 n=2,400	(0.6)	0.4	(1.0)	No
Military	1.1 n=4,495	(0.3)	0.8 n=2,360	(0.3)	0.2	(0.4)	No
VA	1.7 n=4,454	(0.4)	1.5 n=2,349	(0.3)	0.2	(0.5)	No
Indian	0.7 n=4,404	(0.6)	0.2 n=2,254	(0.1)	0.5	(0.6)	No
Other	8.5 n=4,460	(0.9)	6.9 n=2,281	(0.7)	1.7	(1.1)	No

Source: U.S. Census Bureau, American Community Survey Internet Test, April to May 2011

Question type: ***Mark One***

Table 11. Gross Difference Rates for Tenure

	Internet		Mail				
Tenure	Estimate (%)	Std Error (%)	Estimate (%)	Std Error (%)	Internet - Mail (%)	Std Error (%)	Significant?
Unweighted sample size	(n=3,249)		(n=2,475)				
Owned, Mortgage	3.2	(0.4)	3.0	(0.4)	0.2	(0.6)	No
Owned, Clear	3.5	(0.4)	3.5	(0.5)	0.1	(0.6)	No
Rented	0.9	(0.2)	1.2	(0.3)	-0.2	(0.3)	No
Occupied	1.2	(0.3)	1.4	(0.3)	-0.3	(0.4)	No

Source: U.S. Census Bureau, American Community Survey Internet Test, April to May 2011

Table 12. Gross Difference Rates for School Enrollment by Mode

	Internet		Mail				
School Enrollment	Estimate (%)	Std Err (%)	Estimate (%)	Std Err (%)	Int - Mail (%)	Std Err (%)	Significant?
Unweighted sample size	(n=7,154)		(n=4,669)				
No, has not attended	1.9	(0.2)	2.0	(0.3)	-0.2	(0.4)	No
Yes, public school, public college	2.0	(0.2)	2.1	(0.3)	-0.1	(0.4)	No
Yes, private school, private college	1.3	(0.2)	2.9	(0.3)	0.7	(0.4)	No

Source: U.S. Census Bureau, American Community Survey Internet Test, April to May 2011

Question type: ***Mark one plus write-in***

Table 13. Gross Difference Rates for Place of Birth

	Internet		Mail				
Place of Birth	Estimate (%)	Std Error (%)	Estimate (%)	Std Error (%)	Internet - Mail (%)	Std Error (%)	Significant?
Unweighted sample size	(n=7,463)		(n=4,393)				
US/Outside the US	0.3	(0.1)	0.2	(0.1)	0.1	(0.1)	No

Source: U.S. Census Bureau, American Community Survey Internet Test, April to May 2011

Table 14. Gross Difference Rates for Citizenship

	Internet		Mail				
Citizenship	Estimate (%)	Std Error (%)	Estimate (%)	Std Error (%)	Internet - Mail (%)	Std Error (%)	Significant?
Unweighted sample size	(n=7,040)		(n=4,824)				
Born in US	0.3	(0.1)	0.3	(0.1)	0.1	(0.2)	No
Born in a US territory	0.1	(0.1)	0.0	(0.0)	0.1	(0.1)	No
Born abroad of US Parents	0.3	(0.1)	0.3	(0.1)	0.0	(0.1)	No
Citizen by naturalization	0.5	(0.1)	0.4	(0.1)	0.1	(0.1)	No
Not a US citizen	0.2	(0.1)	0.2	(0.1)	0.0	(0.1)	No

Source: U.S. Census Bureau, American Community Survey Internet Test, April to May 2011

Table 15. Gross Difference Rates for Year of Naturalization

	Internet		Mail				
Year of Naturalization	Estimate (%)	Std Error (%)	Estimate (%)	Std Error (%)	Internet-Mail (%)	Std Error (%)	Significant?
Unweighted sample size	(n=278)		(n=104)				
2005 or later	1.5	(0.8)	1.2	(0.6)	0.3	(1.0)	No
2000 - 2004	1.8	(0.8)	4.4	(2.1)	-2.6	(2.2)	No
1995 - 1999	5.6	(2.3)	11	(5.2)	-5.4	(5.6)	No
1990 - 1994	5.2	(2.6)	7.2	(4.1)	-2.0	(4.7)	No
1985 - 1989	2.6	(1.1)	7.9	(3.1)	-5.3	(3.3)	No
1980 - 1984	2.6	(1.1)	2.6	(1.6)	0.0	(1.9)	No
Before 1980	0.8	(0.5)	2.6	(1.4)	-1.8	(1.5)	No

Source: U.S. Census Bureau, American Community Survey Internet Test, April to May 2011

Table 16. Gross Difference Rates for Grade Attending

	Internet		Mail				
Grade Attending	Estimate (%)	Std Error (%)	Estimate (%)	Std Error (%)	Internet - Mail (%)	Std Error (%)	Significant?
Unweighted sample size	(n=1,567)		(n=612)				
Nursery school, preschool	0.1	(0.1)	0.7	(0.4)	-0.6	(0.4)	No
Kindergarten	0.3	(0.2)	1.0	(0.4)	-0.7	(0.5)	No
Grade 1 through 12	0.8	(0.3)	0.8	(0.4)	-0.1	(0.5)	No
College undergraduate years	2.3	(0.5)	2.7	(0.7)	-0.4	(0.8)	No
College or professional school	1.8	(0.5)	2.2	(0.6)	-0.4	(0.7)	No

Source: U.S. Census Bureau, American Community Survey Internet Test, April to May 2011

Table 17. Gross Difference Rates for Educational Attainment

	Internet		Mail				
Educational Attainment	Estimate (%)	St Error (%)	Estimate (%)	Std Error (%)	Internet - Mail (%)	Sd Error (%)	Significant?
Unweighted sample size	(n=7,124)		(n=4,581)				
No schooling completed	1.3	(0.2)	2.1	(0.3)	-0.7	(0.3)	No
Nursery school	1.0	(0.1)	1.4	(0.3)	-0.4	(0.3)	No
Kindergarten	0.4	(0.1)	0.6	(0.1)	-0.2	(0.2)	No
Grade 1 through 11	1.9	(0.3)	2.0	(0.3)	-0.1	(0.4)	No
12th grade - No diploma	1.2	(0.2)	1.2	(0.2)	0.0	(0.3)	No
Regular high school diploma	5.3	(0.4)	6.0	(0.4)	-0.7	(0.6)	No
GED or alternative credential	1.2	(0.2)	1.3	(0.2)	-0.1	(0.3)	No
Some college credit, but less than 1 year	5.8	(0.4)	5.6	(0.4)	0.2	(0.5)	No
1 or more years of college credit, no degree	7.3	(0.4)	8.0	(0.6)	-0.7	(0.7)	No
Associate's degree	3.3	(0.3)	3.0	(0.3)	0.3	(0.5)	No
Bachelor's degree	2.5	(0.2)	2.3	(0.3)	0.2	(0.4)	No
Master's degree	1.6	(0.2)	1.1	(0.2)	0.5	(0.3)	No
Professional degree beyond a bachelor's degree	1.3	(0.2)	1.6	(0.2)	-0.2	(0.3)	No
Doctorate degree	0.6	(0.1)	0.8	(0.2)	-0.2	(0.2)	No

Source: U.S. Census Bureau, American Community Survey Internet Test, April to May 2011

Question type: ***Numeric write-in***

Table 18. Gross Difference Rates for Year of Entry

Year of Entry	Internet Estimate (%)	Internet Std Error (%)	Mail Estimate (%)	Mail Std Error (%)	Internet - Mail (%)	Std Error (%)	Significant?
Unweighted sample size	(n=609)		(n=216)				
2000 or later	1.6	(1.0)	2.0	(1.0)	-0.4	(1.6)	No
1990 - 1999	0.7	(0.4)	3.5	(1.5)	-2.9	(1.6)	No
1980 - 1989	2.2	(1.1)	2.0	(1.3)	0.2	(1.7)	No
1970 - 1979	1.8	(0.7)	2.0	(1.5)	-0.2	(1.6)	No
Earlier than 1970	0.9	(0.5)	1.5	(1.4)	-0.7	(1.4)	No

Source: U.S. Census Bureau, American Community Survey Internet Test, April to May 2011

Question type: ***Dollar write-in***

Table 19. Gross Difference Rates for Rent Amount

Rent Amount	Internet Estimate (%)	Internet Std Error (%)	Mail Estimate (%)	Mail Std Error (%)	Internet - Mail (%)	Std Error (%)	Significant?
Unweighted sample size	(n=643)		(n=406)				
Less than $250	1.4	(0.6)	1.2	(0.6)	0.2	(0.7)	No
$251 - $501	1.6	(0.6)	2.7	(1.1)	-1.1	(1.3)	No
$502 - $750	2.4	(0.7)	3.1	(1.1)	-0.6	(1.1)	No
$751 - $1,001	2.4	(0.6)	2.2	(1.0)	0.3	(1.1)	No
$1,002 - $1,250	1.8	(0.6)	1.2	(0.7)	0.6	(0.9)	No
$1,251 - $1,501	2.0	(0.7)	1.3	(0.7)	0.6	(1.0)	No
$1,502 - $1,750	1.4	(0.6)	0.6	(0.5)	0.9	(0.7)	No
$1,751 - $2,001	0.4	(0.4)	1.1	(0.7)	-1.7	(0.8)	No
More than $2,002	0.2	(0.1)	1.0	(0.7)	-0.8	(0.7)	No

Source: U.S. Census Bureau, American Community Survey Internet Test, April to May 2011

Table 20. Gross Difference Rates for Mortgage Amount

	Internet		Mail				
Mortgage Amount	Estimate (%)	Std Error (%)	Estimate (%)	Std Error (%)	Internet - Mail (%)	Std Error (%)	Significant?
Unweighted sample size	(n=1,525)		(n=804)				
Less than $501	2.2	(0.5)	2.7	(0.8)	-0.5	(0.9)	No
$502 - $1,001	5.5	(0.8)	4.6	(0.9)	0.9	(1.2)	No
$1,002 - $1,501	7.5	(0.9)	6.8	(1.3)	0.8	(1.5)	No
$1,502 - $2,001	6.5	(0.6)	5.6	(1.0)	1.0	(1.3)	No
$2,002 - $2,501	3.6	(0.5)	3.3	(0.8)	0.3	(0.9)	No
$2,502 - $3,001	2	(0.4)	2.1	(0.6)	-0.1	(0.7)	No
$3,002 - $3,501	0.6	(0.5)	1.3	(0.5)	-0.7	(0.7)	No
$3,502 - $4,001	0.5	(0.2)	1.8	(0.6)	-1.3	(0.6)	No
More than $4,002	1.1	(0.5)	0.8	(0.4)	0.2	(0.6)	No

Source: U.S. Census Bureau, American Community Survey Internet Test, April to May 2011