



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

August 16, 2012

2012 AMERICAN COMMUNITY SURVEY RESEARCH AND EVALUATION REPORT MEMORANDUM SERIES
#ACS12-RER-26

DSSD 2012 AMERICAN COMMUNITY SURVEY MEMORANDUM SERIES
#ACS12-MP-04

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: Use of Paradata to Assess the Quality and Functionality of the American
Community Survey Internet Instrument

Attached is the final American Community Survey Research and Evaluation report "Use of Paradata to Assess the Quality and Functionality of the American Community Survey Internet Instrument." The paradata analysis used data from the April 2011 American Community Survey Internet Test. The objective of the analysis was to explain how respondents interacted with the Internet instrument and identify potentially problematic questions, features, and procedures in the instrument. This report summarizes the results of the paradata 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

Kathy Ashenfelter (CSM)

Temika Holland

Beth Nichols

Victor Quach

Barbara O'Hare (DIR)

Nancy Bates

Frank Vitrano

Mary Ann Chapin (DMD)

Justin McLaughlin

Alan Berlinger (DSCMO)

Tim Gilbert (DSD)

Tony Tersine (DSSD)

Michael Bentley

Mary Davis

Steven Hefter

Joan Hill

Samantha Stokes

Brenna Matthews

Jennifer Tancreto

Mary Frances Zelenak

David Johnson (SEHSD)

Scott Boggess

Bob Kominski

Enrique Lamas (POP)

Colleen Hughes

Anne Ross

Janice Valdisera

John Studds (ASD)

Glenn Eanes (ASD)

Joe Mistichelli (ASD)

Use of Paradata to Assess the Quality and Functionality of the American Community Survey Internet Instrument

FINAL REPORT

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

Intentionally Blank

TABLE OF CONTENTS

EXECUTIVE SUMMARY	vi
1. BACKGROUND	1
1.1 Motivation for Analyzing Paradata from the April 2011 ACS Internet Test	1
1.2 Research using Internet Paradata	1
2. FUTURE RESEARCH CRITERIA.....	3
3. METHODOLOGY	5
3.1 Description of the ACS.....	5
3.2 Data Collection Methods	5
3.3 Sample Design	5
3.4 Analysis Design	6
3.5 Internet Instrument Design	6
4. LIMITATIONS.....	7
4.1 Timestamps.....	7
4.2 Glitches	7
4.2.1 Logout issues	8
4.2.2 Links	8
4.2.3 Other Javascript issues.....	8
4.2.4 Timeout issues	9
5. RESEARCH QUESTIONS AND RESULTS	10
5.1 Are there any problematic screens or questions?.....	10
5.1.1 Breakoff rates.....	10
5.1.2 Error messages.....	16
5.1.3 Answer changes	29
5.1.4 Use of the Help feature	36
5.2 Do respondents use the Help option and does it appear to be useful?.....	37
5.2.1 Do respondents use the Help option?.....	37
5.2.2 Is the Help option useful?	38
5.3 Are there any problems with the authentication procedures?.....	39
5.3.1 Failed re-entries	39
5.3.2 Lockouts.....	41
5.3.3 TQA entry inquiries	41
5.4 Are there any screens on which respondents are more likely to leave the instrument and how often do they return?	42
5.5 Is there additional information about the use of the instrument or survey features the paradata can help explain?	44
5.5.1 Clicked links	44
5.5.2 Use of Review & Edit feature.....	45
5.5.3 Use of “Print” feature.....	47
5.5.4 Use of “Previous” button	48

5.5.5 Environment.....	50
5.5.6 Average completion time.....	52
5.5.7 Reasons for calling TQA	55
6. SUMMARY.....	56
Acknowledgements.....	59
References.....	59
Appendix A: Paradata Collected.....	A-1
Appendix B: Alternative Breakoff Rate by Section	B-1
Appendix C: Reasons for Calling TQA	C-1

LIST OF TABLES

Table 1. Future Research Criteria	4
Table 2. Breakoff Rate by Household Size.....	11
Table 3. Breakoff Rates by Screen	11
Table 4. Screens on which Respondents Frequently Received an Error Message	17
Table 5. Number of Errors Received by Respondents.....	18
Table 6. Reasons for and Reactions to Error Messages on Place of Birth.....	20
Table 7. Reasons for and Reactions to Error Messages on Date of Birth/Age	23
Table 8. Responding Behavior for Errors on Respondent Name Screen.....	26
Table 9. Responding Behavior for Errors on Year Built Screen	28
Table 10. Changes before “Next” was Selected (Page on which change took place)	30
Table 11. Changes after “Previous” was Selected (Page on which “Previous” was selected)	33
Table 12. Screens on which Help was Frequently Selected	36
Table 13. Number of Invalid Re-entries by Respondent	40
Table 14. Pages on which Respondents Logged Out.....	42
Table 15. Frequency and Percentage of Respondents who Clicked each Link	44
Table 16. Percent of Respondents that Chose to Review Each Section of the Instrument.....	46
Table 17. Five Screens on which Respondents Selected “Previous”.....	48
Table 18. Devices Used to Access the Instrument.....	50
Table 19. Average Screen Resolution for Anyone who Accessed the Instrument	51
Table 20. Operating Systems Used to Access the Instrument	51
Table 21. Browsers Used to Access the Instrument	52
Table 22. Completion Time Details for Basic Demographic Section	53

Table 23. Completion Time Details for Housing Section.....	54
Table 24. Completion Time Details for Detailed Person Section.....	54
Table 25. Completion Time Details for Review & Edit Section	55
Table 26. Future Research Criteria	57

LIST OF FIGURES

Figure 1. Pick Next Person Screenshot.....	12
Figure 2. Finished Person Screenshot.....	13
Figure 3. Respondent Name Screenshot.....	14
Figure 4. PIN Screenshot.....	15
Figure 5. Distribution of Where in the Instrument Respondents Broke off.....	16
Figure 6. Place of Birth Screenshot.....	19
Figure 7. Date of Birth/Age Screenshot.....	22
Figure 8. Wages Amount Screenshot.....	25
Figure 9. Year Built Screenshot.....	28
Figure 10. Educational Attainment Screenshot.....	31
Figure 11. Failed Login Screenshot.....	39
Figure 12. Summary Screenshot.....	45
Figure 13. Detailed Review Screenshot.....	46
Figure 14. Median Completion Time (in minutes) by Household Size.....	53
Figure 15. Furthest Section Reached by Breakoffs (%).....	B-1

EXECUTIVE SUMMARY

Objective

This paper provides results of the paradata analysis from the April 2011 American Community Survey (ACS) Internet Test. This analysis is intended to help explain how people interact with the Internet version of the ACS and to evaluate and identify problematic questions, features, and procedures in the instrument.

Methodology

Paradata were collected as a by-product of the Internet data collection process for the April 2011 ACS Internet Test for all Internet respondents. The specific experimental treatments and strata used in the Internet Test were not considered for the paradata analysis; all data were analyzed together to provide insight into the effectiveness of the Internet survey instrument for all respondents that chose that mode of response.

Research Questions and Results

a. Are there any problematic screens or questions?

At a high level, this research did not find any evidence that there were any major issues with the screens or questions in the instrument. We were able to identify a handful of screens that were problematic for a small proportion of respondents, which affect breakoff rates and item nonresponse. We propose additional testing that may help reduce respondent issues with these screens. Among these are screens on which respondents logged out or broke off the most frequently. These pages include some transitional screens (Pick Next Person and Finished Person) as well as the Respondent Name and Date of Birth/Age screens. The prevalence of exiting on these screens was fairly low, but they made up the largest proportion and help provide a focus for future research.

b. Do respondents use the Help option and does it appear to be useful?

Respondents did not seem to have difficulty locating the Help link, as it was used more frequently than has been seen in past research. Further, in some cases it seemed respondents used the information to help formulate their answers.

c. Are there any problems with the authentication procedures?

From the data we had, we did not find any problems with authentication. However, paradata do not allow us to detect some potential issues associated with authentication, such as trouble with the initial login and respondents losing their PIN and therefore never attempting to re-enter the instrument.

d. Are there any screens on which respondents are more likely to leave the instrument and how often do they return?

This research found no evidence that any specific screen resulted in meaningful numbers of respondents leaving the instrument. As compared to other screens in the instrument, respondents were more likely to log out of the instrument on transitional pages (Pick Next Person and Finished Person) and questions asking for sensitive information, such as the amount received in wages and their work address. While the majority of respondents do return, these pages also have high breakoff rates compared to other questions in the survey.

e. Is there any additional information about the use of the instrument or survey features the paradata can help explain?

Respondents generally interacted with the instrument the way in which we expected. They clicked the privacy and security link on screens with more sensitive questions, a little over 20 percent used the Review & Edit feature, they completed the survey in the expected amount of time, and just over two percent used a mobile device to complete the survey.

1. BACKGROUND

1.1 Motivation for Analyzing Paradata from the April 2011 ACS Internet Test

Paradata have been defined as process data, or all the data collected during the response process that do not include the response itself (Couper, 1998). Paradata exist in both interviewer-administered surveys and computer assisted self-administered surveys. For interviewer-administered surveys, they can include response times, respondent utterances (pauses, hedges, stutters), respondent expressions, and interviewer observations (e.g., toys in a yard). In computer assisted self-administered surveys, such as Internet surveys, location of breakoffs, changed answers, error messages, mouse clicks, location and quantity of help requests, and response times can all be collected. These paradata can be used to identify potential problems with the survey instrument from question wording to design. Additionally, they can help researchers understand the process the respondent uses to complete the survey, which can in turn help researchers identify ways to make the survey task less burdensome for respondents.

To date, the American Community Survey (ACS) has used paradata to help with telephone call center scheduling, as well as in computer-assisted telephone and personal interviews, but has not used paradata to evaluate past instruments' content and design. Rather, the ACS has relied on experimental testing of content changes in the 2006, 2007, and 2010 ACS Content Tests to study item nonresponse rates, benchmarks to other surveys, and response error measurements. Results from the ACS Content Tests can be found at http://www.census.gov/acs/www/library/by_series/content_test_evaluation_reports/.

1.2 Research using Internet Paradata

Paradata from the April 2011 ACS Internet Test can be used to assess the quality of the instrument design and indicate what modifications may make it easier to use and attractive to potential respondents. Many researchers have used paradata from Web surveys to test different aspects of their instruments and questions (Heerwegh, 2003, Couper *et al.*, 2001, Redline *et al.*, 2009, Dillman and Bowker, 2000).

In 2003, Heerwegh used an Internet survey to analyze response times for knowledge questions. He asked respondents three different knowledge questions, and for each question the average response time for respondents who got the answer correct was significantly shorter than the average for respondents who answered incorrectly. These results suggest that when respondents do not know the answer to a question, response times increase as they try to generate an answer instead of simply retrieving from memory. Other research on response times in Web surveys suggest that more clauses in a question, words per clause, answer categories, and factual and attitudinal questions, as compared to demographic questions, all lead to longer response times (Yan and Tourangeau, 2008).

In another paper, Heerwegh (2002) found that over a battery of eight questions, 30 percent of respondents changed their answer to at least one question. He suggests this finding could either indicate these respondents are motivated and want to do a good job or there could be usability problems with a question that has significantly more changed answers than average. In other research, the number of changed answers, backtracks, and confusion due to response option format led to higher rates of abandoned interviews (Jeavons, 1999). Identifying potentially problematic screens can not only lead to higher quality data for those specific items, but better overall response across the survey.

Internet survey paradata also can provide insight into design problems in regard to the layout and format of response options. Redline *et al.* (2009) experimented with including headings for questions with long lists of response options (they tested this format with several questions, one of which was the ACS's educational attainment question). They found longer response times for questions with headings. This information, combined with mouse click paradata, led to the conclusion that respondents were confused by the headings and thought they should provide an answer for each group of response options.

In addition, Couper *et al.* (2001) presented respondents with two different response option formats for a single question. Half of the sample received five entry boxes where they could type in their response; this represented a simple scenario. The other half received a more complex grid with radio buttons. While providing the same information, the radio buttons presented respondents with 65 input locations, compared to five for the entry box version. They found significantly longer response times for the radio button format, as well as more incorrect responses due to its complexity. Other commonly-used features in Internet surveys such as drop-down menus, unclear instructions on how to fill out the questionnaire, and the absence of navigation aids also have led to higher breakoff rates and incorrect responses (Dillman and Bowker, 2000).

This research suggests paradata can be used to identify potential problems with the survey instruments' content, design, and features. The features and design of the 2011 ACS Internet Test instrument are a product of extensive research, laboratory testing, and two field tests, but have not been used in ACS production. This paper analyzes the paradata from the first of the two 2011 field tests, which was conducted in April. These data allowed us to fully evaluate how well the instrument is working and whether there are modifications that need to be made prior to production due to problematic questions or features.

2. FUTURE RESEARCH CRITERIA

Prior to conducting this research, we defined criteria that would determine the need for future research. In this report we explored how respondents used the Internet instrument and were looking for aspects of the survey that respondents may have found challenging or led them to break off. When these problems were found, we indicated ideas for future research to assess exactly what the problem is and how best to remedy it. The criteria table below (Table 1) outlines the potential problematic aspects of the survey that we felt, if met, would indicate the need for future research. Where available, the criteria were defined by benchmarks from prior research. However, some of the analyses described in this report are new, and we were unable to find any research on what is generally considered “high” or “low.” For these items we decided to not set specific criteria, but rather wait to see what the data looked like and see if there were any outliers or things that looked suspicious.

Table 1. Future Research Criteria

Criteria	Possible Future Research Topics
Overall breakoff rate more than 10 percent ¹	Determine where breakoffs are occurring and why
Breakoff rates for individual screens are outliers, as compared to other screens in the survey	Change question wording/format
Help used in less than 10 percent of cases ²	Change location/prominence/ease of use
A considerable percentage of Internet users have multiple unsuccessful re-entries	Alter authentication/PIN structure
A notable percentage of respondents access the survey via a mobile device	Create a mobile-friendly instrument
Less than 5 percent of respondents use Review & Edit feature	Reconsider need or functionality of Review/Edit
Median completion time greater than 45 minutes ³	Test alternative instrument designs like topic-based/dashboard navigation
High percentage of changed answers on a particular screen, compared to other screens	Testing to determine why/change wording/format
Screen renders a high percentage of error messages, compared to other screens	Determine cause/change wording/formatting of question
Respondents finish the survey but fail to submit	Place “Submit” button on Review & Edit screen

¹ Breakoff rates of five percent were found for the person-based design in the 2005 National Census Test (Bentley *et al.*, 2006). Additionally, the 2001 ACS Internet Test saw breakoff rates of 24 percent (Griffin *et al.* 2001); however there were only 320 Internet respondents in this test. Therefore, we are using a selection criterion of ten percent instead because of the length of the ACS as compared to the Census and the increase in Internet use and access over the past decade.

² Research in Web surveys has shown approximately 13 to 25 percent of respondents use “Help” when it is available (Conrad *et al.* 2006; Lind *et al.* 2001). Therefore, if fewer people than average, across other surveys, use the ACS help, it could be due to the design of the feature itself.

³ The average response time for the ACS across all survey modes and household sizes is 38 minutes. This threshold of 45 minutes accounts for respondents multi-tasking while completing the survey.

3. METHODOLOGY

3.1 Description of the ACS

The ACS is a mixed mode, mandatory household survey of all persons living or staying at a sampled address. The Census Bureau samples about 3.54 million housing unit addresses in the ACS each year. Since its inception, sampled units receive a questionnaire in the mail. If they do not complete the questionnaire, they are switched to a Computer Assisted Telephone Interview (CATI) if we have a phone number for the address. A sample of the addresses that have still not completed the survey are then visited by a field representative to conduct a Computer Assisted Personal Interview (CAPI).

The survey consists of a series of demographic questions that are asked one question at a time for each household member, a series of questions about the housing unit, and finally 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.

3.2 Data Collection Methods

In the April 2011 ACS Internet Test, we tested four different strategies for notifying sampled units about the Internet response mode using combinations of the five ACS mailing pieces (pre-notice letter, initial questionnaire mailing, reminder postcard, and for nonrespondents only, replacement questionnaire mailing and additional reminder postcard). The notification strategies were: Prominent Choice, Not Prominent Choice, Push Internet on a Regular Mailing Schedule, and Push Internet on an Accelerated Mailing Schedule. Respondents in the Prominent Choice and Not Prominent Choice treatments received a questionnaire along with their invitation to complete the survey online. Therefore, the respondents who noticed the Internet option had a clear choice between modes for completing the survey. On the other hand, the Push treatments did not receive the paper questionnaire until two or three weeks after the invitation to complete the survey online. Therefore, these respondents were pushed to use the Internet, more so than the other treatments. This likely resulted in people using the Internet who normally prefer to respond by paper. For more information on the notification strategies and results, see “2011 American Community Survey Internet Tests: Results from the First Test in April 2011” (Tancreto *et al.*, 2012).

3.3 Sample Design

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 expect 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 of sample selection, we suspected these groups varied by age, education, and computer experience and we wanted the

opportunity to identify differences in responding behavior across these groups. For more information on the Targeted and Not Targeted groups, see Tancreto *et al.* (2012).

The estimates and proportions in this report are not weighted. The goal was to understand how all users interacted with this instrument and not to make inferences to the population. However, to ensure there were not important differences between strata, the estimates in this report were first analyzed by strata. We did not find any differences across strata, so the estimates provided in this report are for all Internet cases combined.

3.4 Analysis Design

Included in this report is an analysis of the paradata collected during the ACS Internet Test in April 2011, which is intended to help assess the quality and efficiency of the Internet instrument. There were three major categories of paradata collected: survey access/authentication, session information, and features. The full list of paradata that were collected can be found in Appendix A. The goal of this analysis was to study behavior related to the user's interaction with the Internet instrument. Specifically, we attempted to detect potentially problematic questions, formats, and procedures in the ACS Internet instrument. Paradata were collected and analyzed for every screen and for every household that logged into the survey. The analysis focuses on observing response behaviors to identify any systematic outliers (e.g., a large number of breakoffs for one screen) and the functionality of the log-in and PIN authentication.

While we expected some differences across notification strategy, specifically between the Choice and Push panels, the purpose of this analysis was to learn about the users' experience using the instrument. All analyses discussed in this report, except the breakoff rates, are at the unit or case level. The breakoff rates, however, were analyzed at both the unit level and the person level because we hypothesized that households with more people have a higher probability of breaking off.

3.5 Internet Instrument Design

The ACS Internet instrument was designed to be similar to the mail and CATI/CAPI ACS data collection modes. This strategy was used to minimize mode effects, while taking advantage of the technology, as used in CATI and CAPI, to improve data quality. Consistent with the paper and CATI/CAPI versions, the Internet instrument had four sections of questions: the first section creates a roster of household members; the second asked basic demographic questions for all persons on the roster; the third section, the housing section, asked questions about the housing unit; and the fourth section asked detailed questions about each person in the household. Unique to the Internet mode, at the end of the survey, the respondent had the option of reviewing responses or submitting the survey without reviewing. If respondents chose to review, they could select whether they wanted to see the housing data or the person data for each individual. In other words, they did not need to review all the data. Additionally, they could link to specific questions within the review to change their answers. For more information on the instrument design, please see Tancreto *et al.* (2011).

4. LIMITATIONS

All of the limitations discussed in this section have a mild impact on the estimates presented in this report. However, they are not so problematic as to jeopardize the relevance of the findings. These are just caveats to keep in mind while reading the results because the numbers presented are not as exact as they may appear. Unfortunately, most of these issues concern the nature of paradata and cannot be fixed for future tests.

4.1 Timestamps

Timestamp data are collected both on the client side (user) and the server side (Census Bureau). When respondents log in, log out, and submit their surveys, the time is captured on the server side. Timestamps for all other actions, such as when the respondent enters a page and selects an answer, are captured on the client side. For respondents on the East Coast who have their clocks set closely to the server's clock, this is not an issue. However, for all other respondents, their data points are not sequential. For example, because the login is on the server side, it will appear that respondents on the West Coast logged in to the survey three hours after answering the first question. Since these data points do not map to the order in which respondents progressed through the instrument, it can be difficult to identify whether a respondent logged out or left the survey by closing their browser.

This difference between server and client side timestamps was not an issue for calculating durations⁴. Rather than timing when a respondent logged in and logged out, we instead measured from when they entered the first screen to when they entered the Presummary⁵ screen because both of these measures are on the client side.

4.2 Glitches

As this was among the first times Internet paradata have been collected at the Census Bureau, there were glitches found after testing that affect most of the estimates presented in this report. This section provides specific examples of glitches, including logouts, links clicked, Javascript, and timeout issues. These issues are not related to the instrument, but are simply due to the way paradata are captured and cannot be fixed. Additionally, it is not clear how big of a problem these different issues are. The glitches seem to appear randomly throughout the survey, so it is not possible to search for them explicitly. Most were identified while looking at other data anomalies.

⁴ One respondent appeared to change the clock on their computer during the survey, resulting in starting the survey in January 2011 and finishing in April 2011. This individual was removed from the dataset for completion time calculations.

⁵ The Presummary screen appears when the respondent has viewed all questions for every person on the roster. On this screen, the respondent is given the choice of reviewing his/her answers or submitting the survey.

4.2.1 Logout issues

When a respondent selects “Save & Logout” and then selects the option to leave the survey, they receive a “Welcome Back” screen when they log back into the survey. However, these screens are not form pages. Form pages collect data for the survey and track the survey’s flow. Non-form pages, also called static pages, just display information, but do not collect survey data. Due to the nature of these pages, Javascript has to load before the user clicks “Next.” Otherwise, a respondent can go through all the proper logout procedures, but the “Save & Logout” or “Welcome Back” screens do not appear in the paradata log. These screens are the best way to determine who logged out and back into the survey as compared to leaving the survey by closing the browser or never leaving. Therefore, this issue leads to an underestimate of who returned to the survey. Other work-arounds were attempted, but they each have their own drawbacks, which also appear to lead to an undercount. Unfortunately, the magnitude of this issue is unknown as there is no way to explicitly search for the problem.

4.2.2 Links

When a respondent selects a link, the paradata first records a *type* of “link clicked,” the page the link was clicked on, and the specific link that was clicked. The next record that appears is a *type* of “entry” into the link, followed by an “exit” from the link, and then the next action the respondent takes. However, in some cases, respondents entered a linked page without ever displaying the “link clicked” *type*. Specifically, it appears that respondents changed their answers at the end from the Review & Edit screen without ever clicking the link to send them to that page. As with the logout issues discussed in the prior section, this appears to be a Javascript loading time issue with non-form pages. This type of issue results in under estimates of the number of links clicked in the survey.

4.2.3 Other Javascript issues

Using Javascript to collect paradata can cause several issues. First of all, if a respondent does not have Javascript enabled, no data will be collected from the user side. These individuals are included in the response rate calculation, but are otherwise excluded from this analysis. Roughly 140 households (out of 19,406 households that accessed the survey) did not have Javascript enabled for the entire survey. It is possible households who have Javascript disabled are different than other participants.

Another potential issue with Javascript is load times. If load times are long on a particular page, paradata may not be collected for that page. For example, in order to first log into the survey, respondents must first enter a User ID and then click “Login.” Every respondent should then be presented with the “Address” screen asking if he/she lives at a specific address. However, there were several cases identified for which the “Address” screen does not appear in a respondent’s paradata. Some respondents appear to enter the survey at the PIN screen, which appears three screens after the login, while some enter at a random page later in the survey. This could be either the result of a respondent enabling Javascript part way through the survey, or long load times. Due to this anomaly, these cases had to be excluded from all duration calculations because it was not possible to determine exactly how long they had been in different sections of the

instrument. Further, we are likely underestimating the exact number of respondents who accessed different screens, especially at the beginning of the instrument.

Similarly, long load times can result in “holes” in the paradata. Specifically, several respondents skipped from the “Race” question to “Year Built,” which is two questions later in the instrument. Further, we know these are issues with the paradata and not with the instrument because their final data output contains answers for the items missing in the paradata, which means the respondent saw the screen. These types of glitches cannot be fixed on the client side, but are problematic when trying to determine how many people visited each screen or what actions were taken on a screen because they are not included.

The program the Census Bureau uses to output Javascript, XML, collects paradata by nesting actions within headings. These paradata can then be sorted by any one of the variables corresponding to the headings. We opted to have the paradata sorted by time so we could see how respondents move through the instrument. However, the time is recorded in seconds. Therefore, it is possible for two actions to occur within the same second. This mostly occurs with write-in fields. Specifically, the timestamp for the write-in is not recorded until the respondent clicks out of the text box. Often, their next click is on the “Next” button. Therefore, they receive the same timestamp for both actions. When this occurs, Javascript does not know in what order to put these two events since it can only sort on one variable. Therefore, it selects one to be first and one to be second, which does not necessarily reflect the correct order. This makes programming a challenge because the paradata may show a respondent entering one page before ever leaving the prior. We attempted to account for various paradata orders, but some may have been missed, which would lead to an underestimate of statistics such as percentage of changes or percentage of breakoffs/logouts after an error, to name a few.

4.2.4 Timeout issues

Respondents “timeout” of the survey if they are idle for more than 15 minutes. When they return to the survey and try to interact with it, they are informed they have timed out and are asked to log back in. However, in some cases the paradata shows respondents timed out multiple times in a row. This is not possible because the timeout is not activated until a button is selected, so there has to be at least one other action before there can be another time out. Therefore, this type of glitch leads to an overestimate of the number of timeouts because timeouts are calculated using the “timeout” label in the paradata rather than the duration of inactivity or the additional login request.

In addition to the specific glitches previously discussed, there were some minor additional problems that presented themselves while analyzing the data that do not have an explanation as to why they occurred. These problems may impact the estimates presented later in the paper, but due to their low incidence they should not impact the conclusions or results in any substantive way. These issues will be discussed in more detail in the applicable sections.

5. RESEARCH QUESTIONS AND RESULTS

5.1 Are there any problematic screens or questions?

To answer this question, breakoff rates, error messages, and answer changes were analyzed.

5.1.1 Breakoff Rates

A breakoff is often defined as a case where someone begins a survey but does not finish it. For the purposes of this paradata analysis, we defined a breakoff as any respondent that accessed the survey (saw the first screen that appears after logging in), but did not reach the presummary screen, which appears after the respondent has seen all applicable screens for all people in the household. In total, 19,406 people accessed the instrument and 16,253 completed the survey on the Internet, resulting in an overall breakoff rate of 15.8 percent. This rate is consistent with the rate reported in the ACS Research and Evaluation Report for the first 2011 ACS Internet Test⁶ (Tancreto et al., 2012).

Additionally, two different meta-analyses have been conducted on breakoff rates in Web surveys. Specifically, Musch and Reips (2000) obtained an average breakoff rate of 16 percent in their meta-analysis. Their meta-analysis included 29 online, voluntary psychology studies. On the other hand, Lozar-Manfreda and Vehovar (2002) obtained a breakoff rate of 34 percent; their analysis included surveys from private companies and individuals that mostly varied in their contact strategies. While the studies in these meta-analyses differ from a large-scale government survey, it is encouraging that the breakoff rate for the ACS Internet Test is a little lower than the average found in these studies. The breakoff rate may have been lower as compared to other studies because the ACS is a mandatory government survey. On the other hand, the breakoff rate may be higher than it would normally be because we did not send a reminder/replacement mailing to those who did not finish the survey but got far enough through to be considered sufficient partial responses. It also should be noted that 10.9 percent of the respondents categorized as Internet breakoffs ended up completing the survey using a mail form, which reduces the breakoff rate by approximately two percent.

In addition, we found that the breakoff rate increased as household size increased (Table 2). This finding is expected based on prior research (Galesic and Bosnjak, 2009). Because increases in household size lead to longer surveys, there is more opportunity for the respondent to grow frustrated or be distracted from the survey.

⁶ The ACS Research and Evaluation Report did not calculate an overall breakoff rate, but rather includes a rate for each treatment. That report used final outcome codes to compute the breakoff rates, which were found to be consistent with the rate computed using paradata. The 15.8 percent rate presented in this report falls in the middle of the rates reported in the ACS Research and Evaluation Report.

Table 2. Breakoff Rate by Household Size

Household size	Breakoff Rate (%)
1 person	5.4
2-3 people	13.6
4-5 people	20.6
6+ people	31.7

Source: U.S. Census Bureau, ACS Internet Test, April to May 2011

The breakoff rate itself does not provide any information on which questions were potentially difficult for respondents. Therefore, in addition to the overall rate, we calculated a screen breakoff rate.

The screen breakoff rate compares the number of breakoffs on each screen to the total number of visits to that screen, which tells us where in the instrument the greatest proportion of visits are resulting in breakoffs. Additionally, we wanted to know on which screen the majority of breakoffs were occurring. Therefore, we calculated the percent of breakoffs by comparing the total number of breakoffs on each screen to the total number of breakoffs in the instrument.

Table 3 provides screen breakoff rates for eight screens with the highest breakoff rates. These screens account for a third of all breakoffs. A review of the data across all screens found that only four pages had a screen breakoff rate that exceeded one percent. Three of these pages had fewer than 300 total visits (two of the three had fewer than 25 visits). Additional research with a larger sample size will help determine whether these three pages require additional investigation. For the other pages, considering respondents saw over one hundred different screens, the fact that a handful of pages account for a large percentage of the breakoffs suggests they are worth more investigation.

Table 3. Breakoff Rates by Screen

Breakoff Page	Breakoff Frequency	Number of Screen Visits	Screen Breakoff Rate (%)	Percent of Breakoffs
Pick next person	454	42,714	1.1	14.8
Finished person	171	26,106	0.7	5.6
Respondent name	142	19,027	0.7	4.6
Wages amount	128	24,323	0.5	4.2
PIN	118	19,102	0.6	3.9
Contact later (wrong address)	7	22	31.8	0.2
Housing unit status (if vacant)	3	261	1.2	0.1
Contact (if business)	1	23	4.4	0.0 ⁷

Source: U.S. Census Bureau, ACS Internet Test, April to May 2011

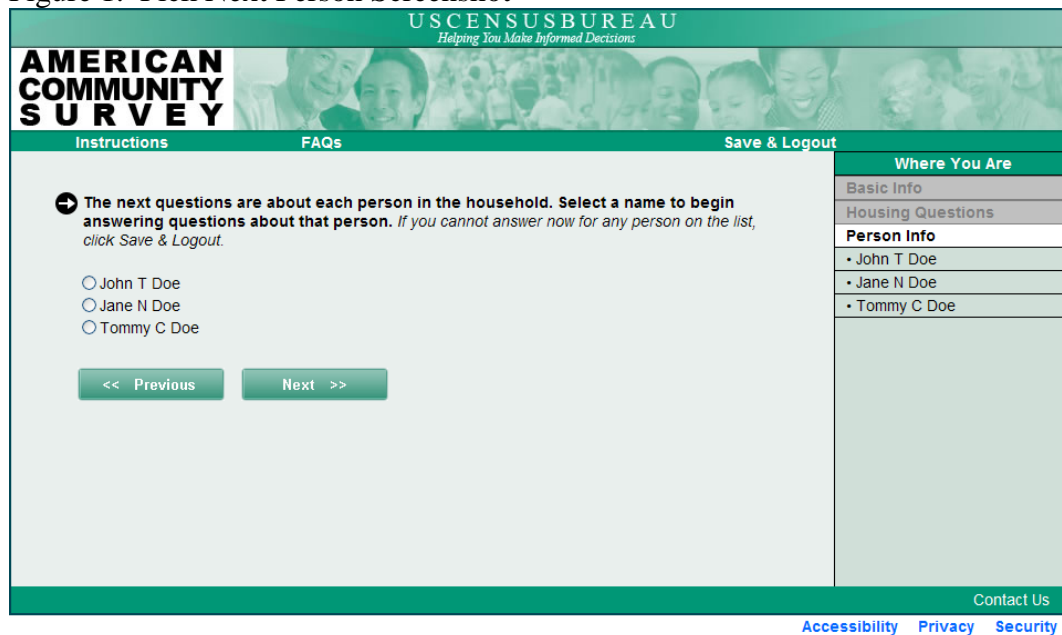
It is clear from Table 3 that the Pick Next Person screen accounts for the largest percent of breakoffs, although compared to the number of visits to the screen, the percentage is quite low. This is a transitional screen that respondents with more than one household

⁷ Less than 0.05 percent of respondents received an error on this page.

member are shown to select which household member they will answer detailed person questions about next (see Figure 1). It was designed to give the respondent flexibility to proceed through the instrument when there was a person on the household roster for which the respondent felt he or she could not answer questions.

There are two main reasons why we suspect there are more breakoffs on this screen as compared to the other screens in the instrument. First, there is explicit language on the screen that tells respondents to leave the instrument and come back later if they cannot answer for another person in the household⁸. It is very possible that respondents simply do not know the responses to the questions for other household members and therefore leave the instrument before starting the next set of questions. This is more common for unrelated households than related households (Matthews, 2012a). The data support this hypothesis for Internet-reporting households, as they are more likely not to provide any data for a nonrelative as compared to a relative or self-report. Specifically, no data were reported for nonrelatives in approximately 24 percent of households, whereas self reports were blank in just over two percent of cases and relatives were not reported in approximately 11 percent. This is particularly important because it suggests that Internet respondents may be less likely than mail respondents to collect data for nonrelatives. This phenomena needs to be investigated for mail respondents to see if it occurs across modes. If mail respondents are more likely to respond for nonrelatives, it may be easier to give the questionnaire to other household members as opposed to providing an Internet household member with the URL and login information, which can be more complicated. Additionally, Internet respondents may not make note of their PIN, so they cannot ask someone else to return.

Figure 1. Pick Next Person Screenshot



⁸ The language on the Pick Next Person screen changes slightly after the first time it is seen.

The second reason we suspect higher breakoffs on the Pick Next Person screen is related to the percent of breakoffs and screen breakoff rate observed for the Finished Person screen as well (Figure 2). Respondents receive the Finished Person screen once they have seen all the detailed person questions for each household member. It informs respondents that their answers have been saved and they will be able to go back and edit their responses at the end of the survey. Both the Pick Next Person screen and the Finished Person screen are transitional pages. The respondent has completed one section of the instrument and has not yet begun the next. Therefore, it is an opportune time to take a break from the survey. Further, 88.5 percent of people who left the instrument on either the Pick Next Person or Finished Person screen used the Save & Logout link. This suggests they may have intended to return, as opposed to just closing the browser, especially because a large percent of people who used the Save & Logout link did return to the survey (see Section 5.4).

Figure 2. Finished Person Screenshot



Increased breakoff rates on transitional pages have been seen in other telephone and Internet surveys (Groves and Kahn, 1979; Peytchev 2009). The increase in breakoffs is explained as an additional participation request. Respondents decided to participate in the survey when they first began, but because transitional pages inform them there are more questions to come, they have to agree to participate again instead of moving smoothly through the survey. Peytchev suggests this is especially true of pages where there is not a question displayed, as in the Finished Person screen. He proposes combining the transitional page with the next screen that contains a question to alleviate this issue. It may be possible to combine the Finished Person and Pick Next Person screens so respondents have a question to answer and do not see two transitional pages in a row.

Although less than one percent of respondents broke off on the Respondent Name screen (Figure 3), these breakoffs are important because the question appears so early in the survey. Specifically, this question appears after respondents have verified their address and received a PIN and asks respondents to provide their full name and phone number. Therefore, if respondents leave on this screen, we have not collected any information about the household.

There are two potential issues with the Respondent Name question. First, Peytchev (2009) found that open-ended questions, which are less simple than closed questions, can increase breakoffs, especially at the beginning of a survey. He specifies that numeric write-ins pose less of a risk than character write-ins, but there is still an effect. Finally, he also found that sensitive questions can lead to respondents abandoning the survey. However, he did not test a name and phone number question. This is an important distinction because name and phone number can be considered sensitive. It is possible that respondents do not want to provide their names due to privacy concerns. One possible way of reducing this rate may be to let respondents know they can use initials if they do not feel comfortable providing their full names⁹. Additional text explaining why name and phone number are needed, along with security information as part of the question, may also help respondents be more comfortable sharing both of these pieces of information.

Figure 3. Respondent Name Screenshot

The screenshot shows the 'Respondent Name' screen of the American Community Survey. At the top, there is a green header with the text 'USCENSUSBUREAU Helping You Make Informed Decisions' and 'AMERICAN COMMUNITY SURVEY'. Below the header, there are three navigation links: 'Instructions', 'FAQs', and 'Save & Logout'. The main content area contains a question: 'What is your name and your telephone number? We may contact you if there is a question.' Below the question, there are input fields for 'First Name', 'MI', and 'Last Name'. Underneath these, there are fields for 'Area Code' and 'Number'. A green 'Next >>' button is located below the input fields. At the bottom of the page, there are links for 'Contact Us', 'Accessibility', 'Privacy', and 'Security'.

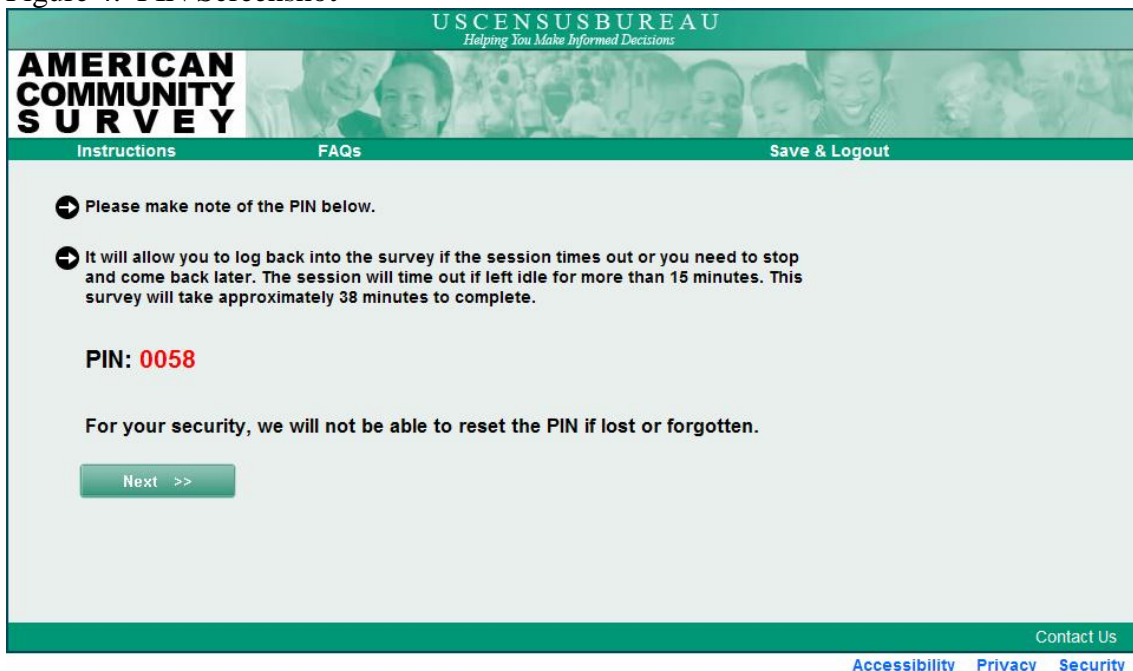
It is expected that breakoffs on the Wages Amount screen (Figure 8) are among the highest for a similar reason as the Respondent Name screen. Wages are seen by many

⁹ We request full names and not initials so if we call back we know who to ask for. It is possible multiple people in the same household have the same initials, so we cannot be sure we are talking to the original respondent.

people as sensitive information that they do not want to provide to outside parties. Peytchev (2009) argued that respondents are more likely to break off on questions requiring numeric write-ins. Further, breakoffs on the utilities and mortgage/rent questions were lower (between less than 0.05 percent to 0.2 percent) than Wages Amount, suggesting there might be an element of privacy concerns for respondents. There was also higher item nonresponse for this item in the mail production sample for the April 2011 panel (Matthews, 2012b). However, other hypotheses are that this question could be difficult to answer for other household members, especially if they are unrelated, or that the income question series comes late in the survey so respondents are simply tired and do not want to continue.

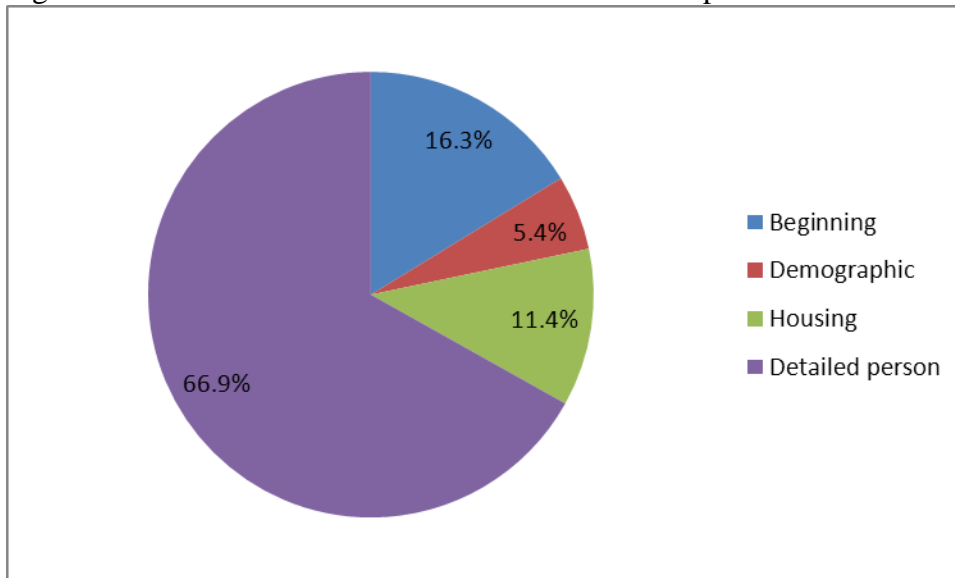
Finally, the PIN screen appears after respondents have verified their address (Figure 4). This screen states that the survey will take an average of 38 minutes, which was added to justify the need for a PIN. Some respondents may have just entered the instrument to see what it looked like, especially in the Push treatments, where there was no paper questionnaire. However, upon seeing the time commitment, these people may have lost interest in completing the survey. Specifically, in an experimental survey, Yan and his colleagues (2011) measured the impact of providing expected durations. This study used two different questionnaires, one long and one short. Additionally, there were two conditions for each survey, one that listed a short duration (five minutes for the short questionnaire and 25 minutes for the long questionnaire) and one that listed long durations (10 minutes for the short questionnaire and 40 minutes for the long questionnaire). They found that the longer duration for each survey resulted in a higher breakoff rate on that screen. Additionally, this screen is also transitional, as it does not ask a new question, so there could be an effect of an additional request for participation.

Figure 4. PIN Screenshot



While the breakoffs by screen identify which pages may be problematic and require additional research, we also looked at where in the instrument breakoffs occurred to make sure the findings are consistent with our expectations. Figure 5 provides a distribution of where in the instrument respondents broke off (beginning, demographic questions, housing questions, detailed person questions), shown in sequential order. Appendix B provides a similar distribution showing the furthest point in the instrument respondents reached before breaking off.

Figure 5. Distribution of Where in the Instrument Respondents Broke off



Source: U.S. Census Bureau, ACS Internet Test, April to May 2011

The distribution of breakoffs seems logical. Respondents who breakoff during the beginning section (which includes address verification, respondent name, the roster section and PIN screens) might just be seeing what the survey is all about. We believe that this could be due to the respondents needing to provide names or realizing the survey will be time consuming. The demographic section is very short and the questions are fairly basic and what one would expect on a survey, so it follows there would be fewer breakoffs; although some respondents may consider age and date of birth to be sensitive since they are personal identifiers. Similarly, the housing section is relatively short and the questions are straight-forward. However, it includes a series of questions on the amount paid in utilities and rent or mortgage, and the value of their property, on which respondents may need to look up their answer. Finally, the detailed person section has the most questions, contains the Finished Person and Pick Next Person screens, and may be difficult for one person to answer for everyone in the household. Additionally, the time commitment to answer for everyone in the household may be more than the respondent is willing to give. Therefore, it is not surprising that almost 70 percent of all breakoffs occur in this section.

5.1.2 Error messages

All of the screens containing questions deemed critical for the ACS, questions critical for skip patterns, and all questions asking for a dollar amount have error messages associated

with them¹⁰ (56 screens). For example, if the respondent left one of these questions completely blank, they received a message saying: “Please answer this important question.” On the other hand, if a respondent enters an invalid value, they receive a message saying: “Please enter only numbers.” In addition, there are more specific errors that instruct respondents on how to fix invalid entries which vary by question.

In total, there were 24,001¹¹ error messages rendered throughout the instrument for 8,715 respondents. This means that about 45 percent of Internet respondents received at least one error message and respondents that received an error message received an average of three. Overall, there were 18 screens (out of 56) on which more than one percent of respondents received an error. For this analysis we identified six of those 18 screens, shown in Tables 4 and 5, on which at least four percent of respondents received an error message and then analyzed each of them in more detail. Table 4 shows where the highest proportion of errors are occurring while Table 5 focuses on how many errors respondents are receiving on these pages. The “Percent of Errors” column in Table 4 shows the total number of errors rendered on each page as a percent of the total number of errors rendered throughout the instrument (24,001). The “Percent of Screen Visits” column represents the number of respondents who received an error on each page as a percent of the total number of times each screen was visited. Table 4 shows that about 2.4 percent of all the screens with a possible error message that were visited resulted in the display of an error message. That rate varied greatly across pages, with visits to Place of Birth and Date of Birth/Age resulting in a higher percentage of errors than the other screens. While the Contact (if business) screen had a high percent of screen visits, only two error messages were rendered and only 23 respondents saw this screen. Therefore, additional research using a larger sample size is needed to determine whether this screen is problematic for respondents. Additionally, it asks the same information as the Respondent Name screen, so the reason for the errors is likely similar.

Table 4. Screens on which Respondents Frequently Received an Error Message

Page Error Rendered	Number of Errors Rendered	Number of Screen Visits	Percent of Errors (%)	Percent of Screen Visits (%)
Total	24,001	990,102	-	2.4
Place of Birth	7,091	43,526	29.5	16.3
Date of Birth/Age	6,007	48,772	25.0	12.3
Wages amount	1,648	24,274	6.9	6.8
Respondent Name	962	19,027	4.0	5.1
Year built	751	18,611	3.1	4.0
Contact (if business)	2	23	0.0 ¹²	8.7

Source: U.S. Census Bureau, ACS Internet Test, April to May 2011

¹⁰ For dollar amounts, a minimum message appears if a respondent enters a non-numeric value in the write-in field, but more critical items have additional messages.

¹¹ Some screens had more than one error message. For example, on the Respondent Name screen, there was an error message for leaving the name field blank and another message for leaving the phone number blank. This number reflects the total duplicated number of messages rendered.

¹² Less than 0.05 percent of errors were on the Contact (if business) screen.

Changing focus to how many error messages respondents received on the pages discussed in Table 4¹³, we see that respondents who received an error message received almost three error messages on average. On the other hand, all respondents who accessed the instrument received just over one error message per screen on average. Therefore, it appears that most respondents are not being inundated with error messages and those that are inundated either leave multiple questions blank (or multiple parts of the same question) or enter multiple invalid responses.

Table 5. Number of Errors Received by Respondents

Page Error Rendered	Number of Errors Rendered	Number of Respondents	Avg Number of Errors per Respondent (for respondents that received an error message)	Avg Number of Errors per Respondent that Accessed the Survey
Total	24,001	8,715	2.8	1.2
Place of Birth	7,091	3,005	2.4	0.2
Date of Birth/Age	6,007	1,363	4.4	0.1
Wages amount	1,648	1,407	1.2	0.1
Respondent Name	962	912	1.1	0.1
Year built	751	745	1.0	0.0

Source: U.S. Census Bureau, ACS Internet Test, April to May 2011

Place of Birth

The Place of Birth screen accounts for almost 30 percent (7,091 total errors) of all the error messages rendered across the survey. Respondents received different error messages depending on whether the entire question was left blank or the radio button was selected without providing a corresponding state or country (Figure 6). If the question was blank, the error message read: “Please answer this important question.” On the other hand, if the state or country was missing, the error message read: “Please select this person’s state of birth.” or “Please enter this person’s country of birth.”

¹³ The Contact (if business) screen was not included in the respondent-level analysis because there were only two error messages rendered.

Figure 6. Place of Birth Screenshot

This question structure is different than the normal radio button format seen in the survey up until this point. In most questions, respondents only need to select a radio button and then they can move on to the next question. However, for this screen, respondents have to select a radio button and then the drop down or text box changes from grey to white to indicate it has been activated (see the “In the United States” response option in Figure 6). At this point, respondents need to select a state from the drop down menu or write in a country.¹⁴ It seems likely the unfamiliarity with this format was problematic for respondents. Of all the errors on this screen, 92.7 percent were rendered due to a missing state or country. Further, 94.1 percent of those missing a state or country were corrected, suggesting respondents did not notice the additional fields or think they needed to provide additional information. The other 5.6 percent of errors were due to leaving the question completely blank. Of these errors, 52.0 percent were corrected. The lower correction rate as compared to leaving the state or country blank suggests people may have intentionally tried to skip the question because they did not want to answer or did not know the answer.

Table 6 provides a further breakdown of the types of errors by person in the household. The final column in the table shows the immediate impact the error message had on breakoffs. This column provides the percentage of respondents who immediately broke off after receiving an error as compared to the total number of people who broke off on the screen. The screen as a whole accounted for 2.5 percent of all breakoffs.

¹⁴ This same structure appears in the Year Built question, but only for respondents whose house or apartment was built in 2000 or later. Therefore, many respondents had not experienced this type of question prior to seeing Place of Birth.

Table 6. Reasons for and Reactions to Error Message on Place of Birth

Screen/Person number	Error Frequency	Number of Screen Visits	Number of Screens Left Blank	Percent of Screens Left Blank	Percent of Blanks Corrected	Number of Screens with Missing Values	Percent of Screens with Missing Values	Percent Missing Values Corrected	Percent Breakoffs after Error Message
Total	3,647	43,526	202	0.5	52.0	3,436	7.9	94.1	4.6
1	2,134	17,971	92	0.5	58.7	2,042	11.4	95.0	18.8
2	980	13,013	69	0.5	52.2	911	7.0	93.2	12.0
3	361	6,690	24	0.4	45.8	337	5.0	91.7	20.0
4	111	3,805	13	0.3	30.8	98	2.6	93.9	0.0
5	33	1,370	4	0.3	0.0	29	2.1	93.1	0.0
6+	28	624	0	0.0	-	19	3.0	89.5	0.0

Source: U.S. Census Bureau, ACS Internet Test, April to May 2011

It is interesting that the percent of corrected errors for people who left the question blank decreased as respondents got deeper in the roster. This finding is likely because respondents may not know other household member's places of birth. However, we do not see a similar reduction in the missing value corrections, lending more support that the format of this question is difficult for respondents. Since the majority of respondents corrected the missing value errors, we suspect that the respondents who left the question completely blank would be more likely to breakoff. However, only 15.8 percent of the breakoffs after an error occurred by respondents who left the question blank. Therefore, although respondents usually corrected their answers, the error message did seem to put some respondents off, even if they had already provided some information.

Date of Birth/Age

The Date of Birth/Age screen also seemed to be problematic for some respondents. This question was topic-based, in that respondents answered the question for each person in the household in succession, as opposed to the Place of Birth question which was asked with all the detailed questions for one household member at a time. The Date of Birth/Age question reads "What is <person's> date of birth and what is <person's> age?" (Figure 7). When a respondent enters a valid date, the instrument automatically calculates the person's age. Because it asks for four separate pieces of information (month of birth, day of birth, year of birth, and age), and each can have multiple errors (blank or invalid), it is possible for respondents to receive up to four individual error messages on this screen, which are displayed together at the top of the screen. For example, if a respondent leaves the question entirely blank, he or she will receive four error messages. There are many ways to receive multiple error messages at the same time on this screen and also to adequately answer the question and still receive an error. This structure inflates the number of errors and does not necessarily identify actual problems with the question. The total number of errors rendered on Date of Birth/Age is 6,007. However, once the duplication is taken into account, there are only 1,673 errors. This also helps explain why respondents who received an error message on this screen received 4.4 messages on average, whereas most other screens averaged closer to one message.

Figure 7. Date of Birth/Age Screenshot

The screenshot shows the USCENSUSBUREAU logo at the top with the tagline 'Helping You Make Informed Decisions'. Below this is the 'AMERICAN COMMUNITY SURVEY' header. A navigation bar contains 'Instructions', 'FAQs', and 'Save & Logout'. The main content area displays question 4: 'What is John T Doe's date of birth and what is John T Doe's age? Please report babies as age 0 when the child is less than 1 year old. (Help)'. The date of birth is entered in MM/DD/YYYY format with three input boxes. Below it, there is a prompt to 'Verify or enter correct age.' and an 'Age (in years)' input box. Navigation buttons for '<< Previous' and 'Next >>' are at the bottom. On the right side, a sidebar titled 'Where You Are' lists 'Basic Info', 'Housing Questions', and 'Person Info'. At the bottom right, there are links for 'Contact Us', 'Accessibility', 'Privacy', and 'Security'.

There were several combinations of errors respondents generally made:

- leaving both date of birth and age completely blank;
- providing an age but leaving date of birth blank;
- leaving age blank (this is usually a result of entering an invalid date of birth so age could not be autocalculated); and
- entering an invalid value (e.g., out of range, character value, incorrect number of digits) for age or date of birth.

Just over half of the error messages rendered were due to some part of the question being blank while the other half were due to invalid entries. Overall, 52.6 percent of the errors were due to leaving all or part of the question blank (27.2 percent of the errors involved leaving the both age and date of birth blank, whereas 25.4 percent of the errors occurred when an age was provided without a date of birth). An additional 8.7 percent of the errors were invalid dates of birth, so the age could not be calculated and they did not fill it in on their own accord.

In addition to being interested in why respondents were receiving error messages, we also wanted to know how they reacted to receiving them. After receiving an error message due to leaving all or part of the question blank, respondents corrected all or some part of the question in 70 percent of cases. Table 7 provides detailed information on what errors respondents received and how they reacted. The columns represent each household member for which the question was asked of. The rows describe the specific type of error. Within each subgroup, we provide the reason for the error followed by the percent that were corrected. The distribution of reasons for error messages varies by household size. For person 1, about 22 percent of the errors were due to leaving both age and date of birth blank. This rose to over 55 percent for person 5, suggesting respondents may not

know this information for other household members or do not want to provide children's information. Additionally, the correction rates for persons 1 through 5 (for leaving all or part of the question blank) decreases as respondents get deeper into the household. This provides even more support for the hypothesis that this information is not known for every household member. On the other hand, the percentage of invalid responses dropped as respondents went deeper into the roster (from about 39 percent for person 1 to 13 percent for person 5). For respondents who provided an age but left the date of birth field blank, most of the corrections were for the first person and the proportion of corrections decreased as respondents got deeper into the roster, possibly because they did not know the exact date of birth for the other household members.

Table 7. Reasons for and Reactions to Error Messages on Date of Birth/Age

	Total	Person 1	Person 2	Person 3	Person 4	Person 5
Errors due to Leaving the Question Completely Blank (% of all errors)	27.2	21.7	30.0	30.9	29.7	55.6
Percent Completely Corrected	45.5	60.0	53.6	53.3	42.9	52.0
Percent Partially Corrected	15.8	11.5	20.0	17.3	11.4	16.0
Errors due to DOB Blank (age provided) (% of all errors)	25.4	29.2	20.3	21.8	29.7	24.4
Percent Corrected	64.2	82.4	49.5	45.3	34.3	18.2
Errors due to Age Blank (invalid DOB) (% of all errors)	8.7	5.0	10.9	12.3	15.3	6.7
Percent Corrected	85.6	84.2	86.3	90.0	88.9	66.7
Errors Due to Other Invalid Entries (% of all errors)	38.7	44.1	38.8	35.0	25.4	13.3
Percent Corrected, Now Valid	73.3	88.7	60.8	52.9	46.7	50.0
Percent Corrected, Still Invalid	5.2	4.0	6.5	7.8	4.2	0.0
Total Corrections Across all Blank	70.0	73.9	58.0	57.6	48.9	43.6
Total Corrections Across All Invalid	80.7	92.2	72.8	70.4	66.7	55.6

Source: U.S. Census Bureau, ACS Internet Test, April to May 2011

For invalid responses, it is not clear why some respondents did not correct their errors. Some may have been annoyed by all the messages they received, they may not have been able to identify the error, or perhaps they simply did not want to provide the information. More research is needed on this question. From browsing through the data, the most common reason for an invalid error was entering a 2-digit year instead of a 4-digit year, but the frequency of this occurrence is not currently known due to programming challenges.

Finally, we also looked at breakoffs as a result of seeing the error messages. There were 75 breakoffs on the Date of Birth/Age screen (2.5 percent of all breakoffs in the instrument) and 24 percent of these breakoffs (18 respondents) occurred immediately after receiving an error message.

There are several issues that might make this question problematic. First, the structure of the error messages results in a lot of text for the respondent to read through, but the text does not necessarily pinpoint the actual mistake. Additionally, many respondents may simply not know the date of birth of everyone in their household. By providing an age, respondents are doing the best they can, yet are still receiving an error, which could be frustrating. Finally, with both age and date of birth on the same page, it is possible respondents think they only need to answer one of the two. Therefore, based on research conducted by Couper and his colleagues (2011) and usability testing at the Census Bureau, we will be using drop down menus when this instrument is introduced into ACS production. This change will almost eliminate invalid errors. Additionally, we should look into more specific wording in the error messages so respondents know exactly what error they made and what we expect.

Wages Amount

If respondents answer that they have received wages in the past 12 months, they see the Wages Amount screen (Figure 8). The question contains a write-in field in which respondents are asked to provide the amount they received. There are two possible error messages respondents can get on this page: “Please enter only numbers” and “Please give your best estimate.” The first error appears if a respondent puts anything other than numbers (dollar signs, commas, and decimal points are all allowed) in the write-in field (invalid). The second error appears if the respondent leaves the write-in field blank.

Figure 8. Wages Amount Screenshot

US CENSUS BUREAU
Helping You Make Informed Decisions

AMERICAN COMMUNITY SURVEY

Instructions FAQs Save & Logout

17 The next few questions are about John T. Doe's income during the PAST 12 MONTHS.

For each type of income John T. Doe received, give your best estimate of the TOTAL AMOUNT during the PAST 12 MONTHS. (NOTE: The "past 12 months" is the period from today's date one year ago up through today.)

For income received jointly, report the appropriate share for each person - or, if that's not possible, report the whole amount for only one person and select "No" for the other person.

a. Did John T. Doe receive any wages, salary, commissions, bonuses, or tips during the PAST 12 MONTHS? [\(Help\)](#)

Yes
 No

➔ What was the amount? Report amount from all jobs before any deductions for taxes, bonds, dues, or other items. [\(Help\)](#)

TOTAL AMOUNT for
past 12 months
\$.00

<< Previous Next >>

Where You Are

Basic Info
Housing Questions
Person Info

- John T. Doe
- Jane N. Doe
- Tommy C. Doe

Contact Us
[Accessibility](#) [Privacy](#) [Security](#)

In total, the Wages Amount screen accounted for 6.7 percent of all errors and 6.8 percent of respondents who viewed the page received an error. The vast majority (91 percent) of these errors occurred because respondents did not attempt to answer the question. However, almost 70 percent of respondents who left the question blank provided a non-zero entry after receiving the error message. An additional 8.5 percent corrected the error by writing zeros in the write-in field. Most of the invalid entries were a result of putting a character in the write-in field. Common entries were “R,” “UNK,” and “NOYB,” (none of your business). Fewer respondents corrected errors due to invalid entries.

Although it seems like many respondents are providing an answer after receiving the error message, we should examine the quality of the data. There were cases in which the respondent would type in a dollar value, then delete it, receive the error message, and then enter a completely different value. In other cases people would enter “55555” or similar patterns. Although the magnitude of these behaviors is not known for certain, the corrected results should be viewed with caution.

In addition to this screen rendering many errors, 4.4 percent of respondents selected Help (see Section 5.2) and 3.2 percent of respondents broke off; 28 percent of breakoffs on this screen occurred immediately after receiving an error message. Additionally, 77 percent of the breakoffs occurred on person one, which means we did not collect any detailed data for the rest of those households. It is expected that many of the errors can be attributed to privacy concerns because there were about 122 write-in messages (e.g., “R,” “UNK,” etc.), as well as clicks on the privacy and security links, which will be discussed in Section 5.5.1. The good news is that even if respondents break off on this screen, it appears at the end of the detailed person section, so they have attempted almost all the questions for at least one person, which can help with imputing values for the remaining

household members. However, we may not have any income data for anyone in the household, so household income can be difficult to impute.

We also compared the percent of errors on Wages Amount to the other income amount screens. Across the other income screens, between 0.7 percent and 1.5 percent of respondents who saw each screen received at least one error message. However, there was no error for leaving these screens blank; an error was rendered only in the case of an invalid entry. Therefore, the percentage of respondents receiving errors on the wages screen is much higher than the other income screens, even when taking the number of visits into account.

Respondent Name

As previously mentioned, each respondent sees the Respondent Name screen (Figure 3) after receiving their PIN. They are asked to provide both a name and phone number. Errors are rendered if a respondent leaves all or part of the question blank or if they enter an invalid phone number (incorrect number of digits or entering character values). Overall, the Respondent Name screen accounted for 4.6 percent of breakoffs and 4.0 percent of error messages. Additionally, 5.1 percent of respondents who saw the Respondent Name screen received this error. This screen also saw the highest proportion of clicks on the accessibility, security, and privacy links (see Section 5.5.1), so it stands to reason that some of the errors and subsequent breakoffs are a result of privacy concerns. However, this screen is also at the beginning of the survey, so respondents could just be trying to figure out what types of questions will be asked and how their information will be used. They may not want to actually answer the questions, but rather see what is being asked of them.

To help understand how respondents reacted to this question, we analyzed their responding behavior and what they did after receiving the error messages. Respondents could answer this question in several ways: they could provide no information at all, provide a name and no phone number, provide a name and an invalid phone number, or answer the question completely. Table 8 shows the behavior patterns for respondents who left all or a part of the question blank and therefore received an error message.

Table 8. Responding Behavior for Errors on Respondent Name Screen

Category or Error	Errors	Percent of Respondents	Percent of Total Errors	Percent of Errors Fully or Partially Corrected
Total errors	912	4.7	100.0	79.7
Errors due to leaving name and phone number blank	376	1.9	41.2	74.3
Errors due to leaving only Phone blank	373	1.9	40.9	82.6
Errors due to invalid phone number	163	0.8	17.9	85.6

Source: U.S. Census Bureau, ACS Internet Test, April to May 2011

Of everyone who received an error on the respondent name screen, about 82 percent of the errors were due to respondents leaving part or the entire question blank. However, of the people who left the question completely blank, 74.3 percent provided some additional information (58.9 percent provided complete information after receiving the error, 13 percent provided only their name, and 2.4 percent provided only a phone number). Additionally, for respondents who originally did not provide a phone number, 82.6 percent provided a number after seeing the error message. It should be noted that the quality of the corrected phone numbers is not clear; some respondents entered numbers with a lot of repeated values (e.g., 4444) or sequential values (e.g., 5678). However, we do not know how prevalent this is or which numbers are actually valid.

Respondents who did completely answer the question could still receive an error message if their answer was invalid. Invalid answers occur for telephone numbers if a respondent enters a character in the numeric field, or if they do not enter the correct number of digits in the fields. Of all the errors, 17.9 percent were due to an invalid entry. Of those entries, 85.6 percent were corrected. Some of the invalid telephone numbers may have been a result of the autotab feature, which moved respondents from one set of digits to the next automatically. As we saw in Table 8, a large proportion of respondents corrected invalid and missing phone numbers, but this finding should be examined critically as some of the phone numbers may be made up.

Finally, we looked at breakoffs as a result of receiving an error message. Of the respondents that received an error message, 28.8 percent broke off immediately after receiving the message. Breakoffs after error messages account for 28.2 percent of all the breakoffs on the respondent name page. Additionally, 17 respondents received an error on the respondent name page because they clicked previous through the entire instrument, deleting data where they could (any data not attached to a radio button). These errors, and related breakoffs, do not seem to be related to this screen in particular, but rather the respondent deciding they do not want to provide any data and this is the final page they can use “Previous” to return to.

Year Apartment Building or Home Built

The Year Built screen rendered 3.1 percent of all error messages (745 error messages rendered). This question is similar to the Place of Birth question in that if the respondent selects the radio button associated with “2000 or later,” then he or she is should provide a specific year in a write-in field. However, there are seven other response options that are not associated with a write-in field (Figure 9).

Figure 9. Year Built Screenshot

Error messages can be rendered on this screen for the following reasons: the question is left completely blank, a respondent clicks the “2000 or later” radio button but fails to enter a year into the write-in field, or an invalid number is entered in the write-in field (character value, a year not between 2000 and 2011, a year that does not contain four digits). Table 9 shows the reasons error messages were rendered, along with respondents’ behavior after receiving the error message.

Table 9. Responding Behavior for Errors on Year Built Screen

Error Category	Errors	Percent of Respondents	Percent of Total Errors	Percent of Errors Corrected
Total errors	745	3.8	100.0	91.5
Errors due to leaving the question completely blank	183	0.9	24.2	82.3
Errors due to leaving the write-in for Year blank	537	2.8	72.6	96.0
Errors due to invalid entries	25	0.1	3.2	83.3

Source: U.S. Census Bureau, ACS Internet Test, April to May 2011

We can see that the majority of errors (72.6 percent) were due to leaving the write-in field for year blank¹⁵. Since 96.0 percent of these errors were corrected, we can assume that many of the respondents did not notice the write-in field, similar to what we saw with the Place of Birth screen. However, respondents saw the Year Built screen prior to Place of Birth. Additionally, while 745 respondents received this error on Year Built,

¹⁵ In total, 3,828 respondents were eligible to receive the “year missing” error, as these respondents indicated their house was built in 2000 or later.

only 112 of the same respondents received the error on Place of Birth, so it is possible we are training them on the format of this type of question. There was also a high rate of corrections (over 80 percent) for the errors due to blank and invalid entries. It is probable that many of the respondents who did not correct their response did not know when their building was built, especially if they live in apartment buildings or are renters.

The largest concern for this screen is that error messages this early in the instrument could have a negative effect later on. Specifically, respondents could grow increasingly frustrated with the error messages and be more likely to breakoff later in the instrument. We may be able to assist respondents who do not know the answer to this question in the help text by suggesting different features of the building they can use to approximate the time period. Concerning the missed write-in field, it may help to experiment with different ways of highlighting the field to draw attention to it or making a separate follow up question.

Conclusions

Overall, error messages occur infrequently, although 45 percent of respondents received at least one. There seems to be a trade off between reduced item nonresponse and properly formed responses on one hand, and breakoffs on the other. For the screens described in Table 4, the majority of respondents who received an error made an attempt to correct it. We also saw a slight negative impact of error messages – 3.6 percent of all breakoffs occurred immediately following an error message. Additionally, respondents who broke off received an average of 1.5 error messages, compared to 1.2 error messages for respondents who completed the survey, which is significant at the $\alpha = 0.10$ level with a p-value of less than 0.0001. However, we cannot say with certainty that these breakoffs are directly related to the error messages. Mooney, Rogers, and Trunzo (2003) also found that error prompts, while decreasing item nonresponse, may be increasing breakoff rates.

Our general conclusion is that many more respondents are correcting their data than are leaving the survey, so the error messages are meeting our goal of improving data quality. However, since there is some relationship between error messages and breakoff rates, it may be worth looking into alternate question formats that are less confusing for respondents in order to reduce the number of errors.

5.1.3 Answer changes

Respondents commonly change their answers to survey questions because they misread or misunderstood the question, they accidentally clicked on the wrong response option, or for another reason. This analysis focuses on answer changes before leaving a screen, changes made after selecting the “Previous” button, and changes through the Review & Edit screens at the end of the instrument.

Answer changes before leaving a screen

Answer changes before leaving a screen occur when a respondent selects multiple responses before clicking the “Next” button¹⁶. Unfortunately, due to the way in which

¹⁶ Changed answers include every time a respondent visited a screen, not just their first visit.

the paradata were captured, we are unable to calculate answer changes for screens that use check boxes instead of radio buttons. Therefore, this analysis focuses on all non-check box screens on which respondents changed their answer without clicking the help link or receiving an error message. Overall, 76.2 percent of respondents made at least one answer change before leaving a page. Table 10 shows eight pages where changes were frequently made before “Next” was selected. The “Percent of Changes before ‘Next’” column represents the changes on the specified page compared to the total number of changes made before clicking “Next” across all questions and all respondents. Table 10 also includes the “Changed Answer Rate,” which is calculated as:

$$\text{Changed Answer Rate} = \frac{\text{Number of changed answers for a question when Help was not selected}}{\text{Total number of cases presented with a question that did not select Help}} * 100$$

The last column is the average number of changes per respondent, which compares the total number of changes for a specific question to the number of respondents that made a change to that question.

The changed answer rate follows a fairly even distribution; with almost 45 percent of screens having a rate of one or greater and 55 percent having a rate lower than one. Table 10 displays the screens with a changed answer rate greater than four, but this is not a clear breaking point in the data and the rate continues to decrease evenly.

Table 10. Changes before “Next” was Selected (Page on which change took place)¹⁷

Page	Percent of Changes before “Next”	Changed Answer Rate
Total changes	-	2.3
Educational attainment	13.7	10.6
Date of birth/Age	11.4	7.6
Heating fuel	4.0	7.1
Verify income	4.7	4.8
Employee type	3.9	4.5
Weeks worked	1.5	6.4
Gas included in rent	0.4	4.2
Housing unit status (if vacant)	0.1	7.7

Source: U.S. Census Bureau, ACS Internet Test, April to May 2011

We see from the table that of the total changes we found, most occur on the Educational Attainment question. This finding is not at all surprising because the screen uses labels to separate the different levels of attainment (Figure 10). Research has shown that in both

¹⁷ It should be noted that the estimates presented in this table are likely underestimates due to limitations noted in Section 4.2 above.

paper surveys and Internet surveys, using labels makes respondents think they need to provide a response to each section (Redline *et al.* 2009). This results in selecting multiple radio buttons as respondents work their way down the list.

Figure 10. Educational Attainment Screenshot

The screenshot shows a survey question about educational attainment. The question is: "11 What is the highest degree or level of school John T Doe has COMPLETED? If currently enrolled, select the previous grade or highest degree received. (Help)". The options are:

- NO SCHOOLING COMPLETED**
 - No schooling completed
- NURSERY OR PRESCHOOL THROUGH GRADE 12**
 - Nursery school
 - Kindergarten
 - Grade 1 through 11- Specify grade 1-11
 - 12th grade - **NO DIPLOMA**
- HIGH SCHOOL GRADUATE**
 - Regular high school diploma
 - GED or alternative credential
- COLLEGE OR SOME COLLEGE**
 - Some college credit, but less than 1 year of college credit
 - 1 or more years of college credit, no degree
 - Associate's degree (for example: AA, AS)
 - Bachelor's degree (for example: BA, BS)
- AFTER BACHELOR'S DEGREE**
 - Master's degree (for example: MA, MS, MEng, MEd, MSW, MBA)
 - Professional degree beyond a bachelor's degree (for example: MD, DDS, DVM, LLB, JD)
 - Doctorate degree (for example: PhD, EdD)

Navigation buttons: << Previous, Next >>

Right sidebar: Where You Are

- Basic Info
- Housing Questions
- Person Info
 - John T Doe
 - Jane N Doe
 - Tommy C Doe

Footer: Contact Us, Accessibility, Privacy, Security

Date of Birth/Age, in addition to rendering more error messages than most other screens, also has a relatively high rate of changed answers¹⁸. About 11 percent of all respondents who saw the Date of Birth/Age question changed their responses prior to receiving an error message or seeking help. There are several reasons respondents changed their answer on this page. Some changes were simply to correct entry errors. In other cases, respondents originally entered a 2-digit year, but then, before clicking “Next” and rendering an error, they realized the question required a 4-digit year. There also were some cases where it appeared as though the autotab feature¹⁹ caused respondents to enter data in the wrong field.

¹⁸ For write-in fields, the respondent had to click out of the text box and then re-enter in order to be included in the changed answer calculations.

¹⁹ The autotab feature works in the following way: once respondents entered two digits into the month write-in box, the cursor automatically tabbed over to the day write-in box. Similarly, after two digits were entered in the day write-in box, the cursor was automatically tabbed to the year write-in.

The other questions identified in Table 10 can be difficult for respondents to answer. For example, many respondents, especially in apartment buildings, do not know what kind of heating fuel they use, or they use multiple types. Although the question asks for the type MOST used, 4.1 percent of paper respondents made multiple selections (Fish 2011). The radio button format forces respondents to make one selection, which may have made deciding on one type more difficult, so they went back and forth between response options. Further, some respondents may not have realized they could only make one selection, so they tried to provide multiple responses.

Similarly, the employee type question asks the respondent what type of business he or she worked for. Not everyone knows the answer to this question, for themselves or for other householders, which also was seen in usability testing (Ashenfelter, forthcoming). Additionally, it is possible that respondents click the first answer they think applies, but then realize there is a better answer further down the list²⁰. Finally, most of the questions in Table 10 (employee type, heating fuel, educational attainment, weeks worked, and housing unit status) have relatively long lists of response options and it is possible people are just accidentally making the selection above or below the one they intended. From the paradata we cannot know why respondents are changing their answers, but these questions do appear to be difficult or cumbersome for some people. Additional research is needed to determine exactly why respondents are changing their answers (other than correcting entry errors) and if there are any changes that can make the questions clearer and easier to answer.

The one anomaly in Table 10 is the verify income screen. This question has a yes/no format, as opposed to the longer lists of response options for most of the other questions in the table. While we cannot be sure of the reason, it is possible that a respondent's initial reaction is to agree with the calculation, but then realize it is not correct. On the other hand, respondents could expect to see a salary total, but adding the other income made the total larger than they expected. More research is needed to know why respondents are changing their answer on this page. A simple starting place would be to see if they are changing their answers from "yes" to "no" or "no" to "yes". If there is consistency in the change, it may help inform the motive behind it.

Answer changes after clicking "Previous"

In addition to changing an answer while on a screen, respondents also were able to return to a prior screen to change their answer. Of the 19,406 households accessing the instrument, the "Previous" button was clicked a total of 89,630 times. The selection resulted in a changed answer 20.0 percent of the time. Each respondent may click the "Previous" button as many times as they desire. In total, 13,325 respondents selected "Previous" at least once. Of these respondents, 69.0 percent made at least one correction. While changes on the current screen seem to suggest an entry error or uncertainty, changed answers after using the "Previous" button to return to the prior screen might suggest a different kind of mistake. Respondents may not think they are confused or

²⁰ We do not know if respondents are using the mouse or a keyboard to complete the survey, so it is possible they are using the arrow keys to move down to the correct response option, which would inflate the number of changes.

have a problem with a question until they see the next screen, which may trigger additional knowledge about what the prior question was actually asking. In this case, the question wording may be confusing to respondents as they do not know exactly what is meant.

Table 11 provides pages on which respondents made the most changes after clicking “Previous”. The “Changed Answer Rate” in the last column uses the same formula that was calculated in Table 10. Overall, 0.6 percent of screen visits resulted in a change using the “Previous” button. This rate varied quite a bit across screens, with 53 screens (out of 139 screens on which respondents selected “Previous”) having a changed answer rate greater than one percent and 10 having a changed answer rate greater than five percent. However, the number of people visiting screens varied as well. Of the 10 screens with a changed answer rate greater than five percent, only two had more than 1,000 visits. Therefore, additional research using a larger sample size would help determine whether there were actually a larger percent of changes on these pages or if it was just due to the small number of respondents seeing these screens.

Table 11. Changes after “Previous” was Selected (Page on which “Previous” was selected)²¹

Page on which “Previous” Clicked	Prior Page	Percentage of all Changes after “Previous”	Changed Answer Rate
Total	N/A	100	0.6
Total income estimate	Verify income	4.2	11.2
Anyone else live or stay	Who lives or stays	3.6	2.9
Electricity included	Pay for electricity	3.2	36.4
Add to roster	Anyone else live or stay	2.4	29.5
Gas amount	Pay for gas	2.0	2.9
Electricity amount	Pay for electricity	1.9	1.6
Public assistance amount	Receive public assistance	1.0	25.0
Add to roster_2	Anyone else live or stay (short time)	0.5	12.3
Who stays at another home	Anyone stay at another home	0.4	6.9
SSI amount	Receive SSI	0.3	6.8
Anyone live or stay	You live or stay	0.1	7.3
Contact later (wrong address)	Address	0.1	36.4
Contact (if business)	Address a business	0.0 ²²	13.0

Source: U.S. Census Bureau, ACS Internet Test, April to May 2011

²¹ It should be noted that the estimates presented in this table are likely underestimates due to limitations noted in Section 4.2.

²² Less than 0.05 percent of respondents made a change after clicking “Previous” on the Contact (if business) screen.

The first thing to note is that overall answer changes after using the “Previous” button are less common (0.6 percent) than changes before leaving the question (selecting “Next”) (2.3 percent). However, for individual questions, these rates are higher. It also is interesting that the majority of these questions come from three sections: roster, utilities, and income. It appears as though respondents were alerted by the next page that they may have made a mistake on the prior page. For example, three of these questions ask how much is paid in different utilities. The prior screen, where the change was made, asks whether anyone in the household pays for these utilities. When respondents who answered affirmatively to “Pay for gas²³” and “Pay for electricity” were asked to provide an amount spent, they may have realized they did not actually pay for the utility since they do not receive a bill. Other possibilities are they do not know the amount so they go back to the prior screen to bypass the question or they did not notice the change in utility from “gas” to “electricity” so they thought they were receiving the same question twice. This would have a negative impact on data quality. More research is needed to know which reason represents most respondents.

Three screens in Table 11 appear in the roster section: “Anyone else live or stay,” “Add to roster,” and “Add to roster_2.” There is a similar explanation as to why these three screens appear in the table. Asking if anyone else lives or stays at the address may remind the respondent to go back and add someone they had left off. Similarly, the “Add to roster” screen appears if a respondent says that there are additional persons living in the household. If the respondent does not want to add additional people, either correctly or incorrectly, they may return to the prior screen and change their answer to say there are no additional people.

Finally, there are three screens from the income section: “Total income estimate,” “Public assistance amount,” and “SSI amount.” For each of these screens, the prior question asks respondents whether they received a certain type of income (or to verify the amount received). When asked to provide a number, respondents may not know the answer or want to provide personal information, so they go back and say they did not receive that type of income. Alternatively, they may have realized they made a mistake on the prior screen and really did not receive that type of income. Again, additional research can help determine why respondents are going back to change their answers on these screens.

The changed answer rates are relatively high for the “electricity included” and “add to roster” screens. Of the respondents that visited these screens, about a third returned to the prior screen to change their response. The most likely explanation for this behavior is the wording of the electricity question. Specifically, the instructions say, “*Select “Yes” if electricity and gas are billed together.*” Respondents who pay for electricity but do not pay for gas on the same bill may be confused by this instruction, leading them to select “No.” When they saw the next question asking if electricity is included in rent or if there is no charge, they realized they answered the prior question incorrectly. Another

²³ There are actually three questions that relate to the amount paid for gas, as opposed to two for the other utility questions. In the gas battery, respondents also backed up from the “Do you pay for gas” question to change the question that asks whether they use gas. This accounted for 1.3 percent of changes after clicking “Previous.”

possibility is that respondents did not want to provide a value for the amount they pay in electricity, so they tried to bypass the question by answering negatively. However, they then realized the path they were going down was not applicable so they went back.

Because none of the screens that appear in Table 11 appear in Table 10, there is more support that the next question is informing the need for change. However, more research is needed to determine if respondents are going back in order to bypass certain questions or if they are actually making corrections. If respondents are not including everyone on the roster or are following the wrong route for the utilities questions, it could have a negative impact on data quality.

Changes from Review & Edit Screen

Once a respondent has seen all the applicable survey questions, they are given the option to review their answers. We were interested in how frequently respondents used this feature to either change an existing answer or provide an answer that had previously been missing. The only question that stood out as having noticeably more changes in Review & Edit was the question asking what type of health insurance each household members uses (319 changes or about 20 percent of all changes). With the exception of this question, the use of Review & Edit to change an answer is distributed evenly across screens.

While it appears as though the health insurance question led to a lot of field changes, this estimate inflates the true number of changes or updates on the page. The insurance screen has eight separate fields in which a respondent can enter a value and each one is counted separately. Many respondents only select the insurance they have and leave the rest blank. If they see blank entries in the Review & Edit and go back to fill in all the missing items, the screen could appear to have eight times as many changes than it actually does. After conservatively accounting for this inflation by dividing the number of changes by eight, we see that at least 2.1 percent of the field changes occur on this screen, which is consistent with the rate of changes for other questions.

The information asked in many of the ACS questions may not be known by every member of the household. Therefore, further research is needed to know if respondents left items blank and then performed a record check to complete the survey, if they guessed and then went back after they could double check, or if they realized their answers were incorrect and went back to correct the reported information. Compared to the number of changes for other reasons, however, these numbers are very small and should not be of large concern.

It should be noted, however, a handful of respondents went back and zeroed out their income data. For example, several respondents entered all of their income information on the first time through the survey. However, using Review & Edit, they changed all the income responses to zeros. Therefore, just because a respondent made a change in Review & Edit, it does not mean we are getting more accurate data.

5.1.4 Use of the Help feature

A Help link was available for 107 questions (out of 145 total questions) throughout the instrument²⁴. Table 12 shows screens where Help was frequently selected. The “Percent of all Help Clicks” column compares the total number of times Help was selected on each screen to the total number of times Help was selected in the instrument. The Question Help Rate compares the number of times Help was clicked on each screen to the total number of visits to the screen (e.g., 6.2 percent of householders who saw the Ancestry screen requested help). In total, only one percent of all screen visits resulted in a Help request (for screens on which Help was available). Additionally, the question help rate had a fairly even distribution across screens with 58.9 percent of the screens with Help available have a question help rate of less than one percent²⁵.

Table 12. Screens on which Help was Frequently Selected

Page	Percent of all Help clicks	Question Help Rate
Total	-	1.0
Ancestry	13.7	6.2
Amount paid in Real Estate Taxes	4.3	5.9
Pay for Electricity	4.7	5.0
Property insurance	2.8	3.9
Property value	2.6	3.6
Amount paid in Water	2.4	3.4
Year built	3.2	3.4
Mortgage	2.3	3.3
Wages	4.0	3.2
Condo	2.2	2.4
Mortgage amt	1.1	2.1
Amount paid in Electricity	1.9	2.1
Employee type	3.0	2.1
Mobile home tax	0.2	10.9
Housing unit status (if vacant)	0.2	11.9

Source: U.S. Census Bureau, ACS Internet Test, April to May 2011

The ancestry question received the most requests for help, both in number of clicks and relative to the number of people that saw the screen. While 2,192 respondents requested help on ancestry for the first person, in total it was clicked 2,703 times across all household members, suggesting that some respondents continued to struggle throughout the roster. It could be that ancestry is simply a difficult concept for respondents.

²⁴ There was not a Help option for every question. Help was provided for the same questions as on the mail questionnaire, which are provided in a booklet that accompanies the questionnaire.

²⁵ The “Mobile home tax” and “Housing unit status (if vacant)” screens had only 329 and 261 visits, respectively, as compared to more than 10,000 visits for the other screens in the table. Therefore, additional research with a larger sample size is needed to determine whether a high percent of people actually needed help on these screens.

Additional research is needed to determine with which aspect of the question respondents struggle.

Similarly, for the other questions in Table 12, without further research we cannot say what information respondents were seeking in the Help or what part of the question they had trouble with. For many of these questions, it is possible respondents simply do not know the answer, so they are looking in the Help for clues as to how to answer. In other cases, such as pay for electricity and amount paid in real estate taxes respondents may not be sure what to include and exclude. Additional research can uncover if there is something inherent in the question that is problematic for respondents and if it is something that can be altered.

Conclusions

In order to identify potentially problematic screens, breakoffs, errors, answer changes, and the use of Help were analyzed. Several screens appeared in multiple tables, suggesting they may be the most difficult for respondents to answer. Specifically, Respondent Name, Wages Amount, and Date of Birth/Age seem to pose more problems for respondents.

Respondent Name had a higher proportion of breakoffs and error messages when compared to other screens in the instrument. This is particularly important because it appears so early in the instrument. Although the breakoff and error rates were low overall, losing or frustrating respondents this early in the survey can lead to essentially collecting no data for these households.

The format of the Date of Birth/Age question seems to confuse some respondents. It is easy for respondents to make errors on this screen due to the date of birth fields and age field appearing on the same page, the autotab feature, and the dd/mm/yyyy format. Additionally, it is likely seen as private information by some respondents.

Privacy is likely the reason for the comparatively high number of breakoffs and errors for the Wages Amount and Respondent Name screens. However, it is unclear if anything can be done to make respondents more comfortable answering these questions and to keep them in the instrument, because these issues occur in all surveys and are not unique to the ACS Internet instrument.

5.2 Do respondents use the Help option and does it appear to be useful?

5.2.1 Do respondents use the Help option?

Past research has shown respondents do not often use available help features in Internet surveys (Conrad *et al.* 2003, Conrad *et al.* 2006, Lind *et al.* 2001). Both of the studies conducted by Conrad and his colleagues covered “lifestyle” topics, such as health, food, travel, and cars. Lind and her colleagues asked five questions from the Consumer Price Index-Housing Survey and five questions from the Current Point of Purchase Survey. The research has provided many variations to the location, the format, and how to activate the Help link. For the April Internet Test, we placed the link immediately

following the question, where we felt it would be easy to find. For reference, Figure 10 on page 31 shows the location and form of the “Help” link.

To determine how respondents used the Help feature, a Case Help Rate was calculated as:

$$\text{Case Help Rate} = \frac{\text{Number of respondents that select Help on at least one screen}}{\text{Total Number of Respondents}} * 100$$

Contrary to much of the research, this respondent-level Case Help Rate was relatively high. The Help feature was accessed by 7,753 respondents, meaning that 40 percent of all Internet respondents selected Help at least once. A total of 19,698 Help feature accesses means that some respondents were more likely to use the Help, since they selected it multiple times during the survey process. Additionally, the use of the Help feature was widespread across respondents and questions. Respondents selected Help on 97 percent of the questions for which it was available and when Help was offered on a screen, more than 100 people selected it in over 50 percent of cases.

Overall, the relatively high case help rate suggests that the location, font and size of the link seem to be adequate in drawing respondents’ attention. While it is encouraging that so many respondents are finding the help link and using it, this also means that respondents are having difficulty understanding the survey questions, which may not be unique to the Internet instrument, but rather related to the questions themselves. For a discussion on the specific questions respondents requested Help most frequently, see Section 5.1.4.

5.2.2 Is the Help option useful?

To determine how useful the help was to respondents, we examined the percentage of cases in which a respondent selected an answer, clicked Help, and then changed their answer. Additionally, we looked at the percentage of people who clicked Help before selecting an answer and then provided a response after reading the help text.

Unfortunately, due to the structure of the paradata, these percentages only include cases where the very next action was a field change. Of all the instances that Help was selected, 38.8 percent of the clicks were prior to selecting a response option. On the other hand, only 3.8 percent of all the help selections resulted in an answer change (an initial answer had already been selected). The low percentage of answer changes after selecting help suggests respondents were either using the information to verify the selection they already made or the help text did not provide the information they were looking for, so there was no reason to change their response. Unfortunately, the paradata cannot differentiate between these two options so additional cognitive testing is needed to uncover respondents’ intentions.

5.3 Are there any problems with the authentication procedures?

To assess the authentication procedures, we analyzed the number of failed re-entries, the number of lockouts, and the number of times telephone questionnaire assistance (TQA) was called with access questions²⁶.

5.3.1 Failed re-entries

Once respondents verify their addresses they are provided with a four-digit PIN (Personal Identification Number). This PIN, along with the corresponding User ID, is used to log back into the instrument if the respondent leaves at any time. If the User ID and PIN do not match, the respondent will receive an error message (Figure 11). Further, if a respondent enters an incorrect User ID/PIN combination more than five times, he or she is locked out of the instrument and cannot attempt another login for 15 minutes.

Figure 11. Failed Login Screenshot

The screenshot shows the American Community Survey login interface. At the top, it says "USCENSUSBUREAU Helping You Make Informed Decisions" and "AMERICAN COMMUNITY SURVEY". Below this, a welcome message reads: "Welcome to the American Community Survey. You will need the materials we mailed to you to start the survey. All the information that you provide will remain completely confidential. Para completar en español, oprima aquí." The main section is titled "Please Log In" and contains a barcode with the following text: "ACS 999 999 999 99 999 9999 99 SE Q999-99999", "TO THE RESIDENT OF:", "101 Main St", "Anytown, MD 00043". A green arrow points to a 10-digit User ID field in the barcode area. To the right, there is a red error message box that says "Login failed. Please try again." Below this, there are instructions: "Enter the 10-digit User ID found below the barcode on the materials we mailed to you." with a "User ID:" label and two input boxes, and "Enter the 4-digit PIN we gave you the first time you entered the survey." with a "PIN:" label and one input box. A green "Login" button is at the bottom. At the very bottom, there is a small disclaimer: "The Census Bureau estimates that, for the average household, this survey will take 38 minutes to complete, including the time for reviewing the instructions and answers. You may email comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing burden, to: Paperwork@census.gov. Use 'Paperwork Project 0607-0936' as the subject. Or you may send comments to:"

Due to security regulations, we could not explicitly tell respondents to write down their PINs, but did tell them to “make note” of it. Therefore, there was initially concern that a large proportion of respondents would lose or forget their PIN and have difficulty re-entering the survey after a logout, timeout, or close-out. However, there were only 412

²⁶ We had planned to investigate instances of entering an incorrect User ID. However, only information collected after a household’s initial login is collected in the paradata. Therefore, this information was not available.

failed re-entries across all respondents. As a percentage of all logins, only 1.8 percent were invalid. Further, Table 13 shows the frequency of invalid login attempts. While 11 respondents had more than five invalid logins, this does not mean they were necessarily locked out if they did not occur one after the other.

Table 13. Number of Invalid Re-entries by Respondent

Number Invalid	Frequency	Percent (%)
1	319	77.4
2	53	12.9
3	17	4.1
4	12	2.9
5+	11	2.7
Total	412	100.0

Source: U.S. Census Bureau, ACS Internet Test, April to May 2011

The percent of respondents with more than two invalid logins is less than 10 percent, again suggesting that users easily accessed the survey. For the respondents who had difficulty accessing the survey, 28.0 percent eventually accessed the instrument and 72 percent gave up.

It is possible, however, that respondents who knew they did not have their PIN did not even try to log back in. This issue appears in the TQA data (Roberts, 2012). Roughly 110 people (0.6 percent of people who accessed the survey and 3.5 percent of Internet-related TQA calls) called TQA with re-entry issues, such as a lost PIN, difficulty finding the PIN, or needing the case reset. Therefore, it seems that authentication is not problematic for respondents who remembered their PIN, but could be an issue for other respondents, especially because only highly motivated respondents are likely to call an assistance line.

To further understand respondent’s behavior, we examined re-entries by method of leaving the survey. Respondents who log out see a warning screen that reminds them they will need to use their PIN to log back into the survey. However, respondents who are timed out or close the browser window do not see this warning. Therefore, we believe people who timed out or closed out of the browser might be more likely to have difficulty logging back in. However, this did not seem to be an issue, possibly because some respondents do not actually “note” their PIN. Specifically, we compared the number of invalid logins after logging out to the total number that attempted to log back in after timing out²⁷. Of the respondents that intentionally logged out, 4.9 percent of the 1,179 re-entries were invalid. However, it should be noted that the paradata surrounding logouts is imperfect. Due to complications with timestamps and missing data values, it is possible some successful logins were missed, which would reduce the invalid re-entry percentage.

²⁷ Invalid logins after closing the browser window could not be analyzed because they did not appear clearly in the paradata output.

After timing out, there were only two total invalid re-entries out of 730 attempted re-entries. By definition, in order for a respondent to time out, they must return to the survey and click on a link. However, once they are re-directed to the login screen, they may break off instead of logging back into the survey²⁸. Unfortunately, we cannot say how large of an issue authentication after a timeout is because it is possible that respondents who timed out did not even try to re-enter the survey because they did not have their PIN. However, the same phenomena can occur with logouts, so it is not clear why logouts result in a higher percentage of invalid re-entries. One possibility is respondents stay out of the instrument longer when they log out as opposed to time out, so they have more time to forget or lose their PIN.

5.3.2. Lockouts

Lockouts occur when a respondent has more than five invalid logins in a row. When a lockout occurs, the respondent needs to wait 15 minutes before he or she can attempt to re-enter the instrument. Only seven respondents locked themselves out of the survey. Further, only one person locked themselves out more than once (two times). This statistic lends more support that respondents are not having a lot of difficulty re-entering the instrument. However, as previously mentioned, it is possible that there are not many lockouts because respondents gave up logging back in before they hit the fifth unsuccessful attempt.

5.3.3. TQA entry inquiries

In addition to looking at re-entry issues, there are also initial access issues to take into account. Respondents can call TQA for any difficulty they are having with the ACS, whether related to the Internet or general questions. For the Internet test, interactive voice response (IVR²⁹) was not set up to take Internet-related calls. Therefore, when a respondent called TQA, the representative spoke with them and then coded the reason for the call (Roberts, 2012). Reasons related to initial access authentication include: problems entering URL, need help finding survey website, and help needed finding User ID.

In total, there were 2,288 Internet-related calls out of a mail out universe of 120,000 households that had the option to respond on the Internet. Of these calls, 33.0 percent were concerning initial access problems. Respondents had the most trouble entering the URL and finding the survey website. In usability testing for both the ACS and Census Quality Survey (CQS) (Ashenfelter, 2011) we saw participants entering the URL in the search box instead of the URL box, so this may have been one of the issues callers were having (Stokes, 2011). Although the total number of calls for these two issues only totaled 672, only highly motivated responders generally call an assistance line. Therefore, it can be assumed more respondents had these same problems and either did

²⁸ It is difficult to calculate the number of people who did not attempt re-entry after timing out due to how timeouts are logged in the paradata. They are not directly linked to a page, and the timestamp is on the server side, so it is not necessarily clear when or where in the instrument it occurred.

²⁹ IVR is a voice recognition and response system that allows callers the ability to interact with the host (US Census Bureau) using a pre-programmed system instead of a live representative. The IVR system can answer basic questions for callers or it can guide them to the proper representative who can answer their questions.

not respond or waited for a paper questionnaire. Therefore, it may be useful to experiment with different ways of presenting the URL and directing respondents to the survey website.

5.4 Are there any screens on which respondents are more likely to leave the instrument and how often do they return?

Although respondents are able to leave the instrument in multiple ways, we want as many as possible to complete the survey in one session. If respondents take more than one session to complete the survey, the chances of them losing their PIN, losing interest, or forgetting about the survey increase. Therefore, we are interested in how many sessions it takes respondents to complete the survey, where they leave, and how likely they are to return.

Of the 19,406 respondents who accessed the survey, 81.6 percent only logged in one time. However, not all of these respondents completed the survey; some were breakoffs. Of the 16,253 respondents that completed the survey, 79.9 percent did so in one session. On average, respondents took 1.3 sessions to complete the survey. These results are promising in that the majority of respondents are moving through the instrument in one sitting, so there is less concern of them logging out and never returning.

Of the respondents who left the instrument by timing out or logging out, it is important to know on which page they left so we can see if there is a problem with the survey content or a way we can keep them engaged. Table 14 shows the four pages where the greatest proportion of log outs occurred. The “Percent of All Logouts” column compares the number of logouts on each screen to the total number of logouts, whereas the “Percent of Respondents” column compares the number of logouts on each screen to the total number of respondents who accessed the instrument (19,406). Other than these four screens, no more than one percent of respondents logged out on any one screen, and the next highest was 0.7 percent on the Finished Person screen. Table 14 includes anyone who logged out of the instrument, regardless of whether he or she returned to the survey. Therefore, breakoffs are included in these percentages.

Table 14. Pages on which Respondents Logged Out

Logout Page	Total Number of Logouts	Percent of All Logouts (%)	Percent of Respondents (%)
Pick Next Person	1,250	27.3	6.4
Wages Amount	257	5.6	1.3
Work Address	255	5.6	1.3
Summary	233	5.1	1.2

Source: U.S. Census Bureau, ACS Internet Test, April to May 2011

While we found a handful of pages on which a larger proportion of respondents logged out, we did not find a similar pattern with timeouts. In order for a respondent to time out, he or she must at least attempt to return to the survey, so timeouts are not a subset of breakoffs. There was great variation on the page on which respondents timed out.

Specifically, out of 753 total timeouts, no single page accounted for more than 0.3 percent of all the timeouts (Pick Next Person accounted for the highest proportion at 0.3). This result suggests respondents likely get distracted or interrupted from the survey and do not intentionally leave.

There does seem to be a pattern regarding the pages on which people pause the survey, either by logging out or timing out. To begin with, three of the four pages for logging out are the same pages with the highest breakoffs (Pick Next Person, Wages Amount, and Finished Person). This could suggest that the respondents did not leave out of frustration, but rather intended to return and forgot. For example, as discussed in Section 5.1, the Pick Next Person screen explicitly tells respondents to log out if they do not have information for any remaining people. The large percent of logouts could suggest the people who broke off intended to gather the information they needed on the next person or were waiting for the other person to complete their section, but ended up not doing so.

A very similar explanation can be used for the Finished Person screen; respondents who broke off intended to return but forgot. It is not clear why the Wages Amount screen appears on both lists. It is possible the people who broke off found the question to be invasive while the people who logged out were looking up records.

The Work Address page also appears in Table 14. This question asks the respondent to provide the address of the place the person worked last week. It is possible respondents do not know the answer to this question for themselves or other members of their household. Therefore, logging out or timing out could be a result of the respondent leaving the computer to find the answer.

Since we were worried about people logging out and not having their PIN to re-enter, when respondents select “Save & Logout,” they were reminded they need their PIN to re-enter and were given the option to log out or continue with the survey. Faced with this option, 86.3 percent of respondents that selected “Save & Logout” opted to log out and 13.7 percent chose to continue with the survey. Of the respondents that logged out of the survey, 82.2 percent eventually returned. While the majority of respondents chose to logout, the 13.7 percent continue rate³⁰ suggests the warning message is successful in deterring at least some people from leaving, likely because they are told they need their PIN to return, which they may not have³¹. The most common pages on which respondents opted to continue were Pick Next Person, Summary, Respondent Name, and Wages Amount.

³⁰ These respondents did not necessarily complete the survey.

³¹ The logout and leave rates are calculated by counting the number of people who saw the welcome back screen. However, some respondents did logout and return but did not receive the welcome back screen because it is not a form page. Therefore, the rates are slightly underestimated as they do not count these respondents.

5.5 Is there additional information about the use of the instrument or survey features the paradata can help explain?

5.5.1 Clicked links

On each page of the Internet instrument, the following links were available to respondents: Instructions, FAQs, Contact Us, Security, Accessibility, and Privacy. The presence of the accessibility, privacy, and security links is a Census Bureau Standard. The other links were provided to help respondents obtain the information they may need to complete the survey. The information in these links was provided in a new window. However, clicking on the accessibility, security and privacy links took respondents to a separate tab in the browser. There was some concern that if redirected to another page, respondents would not go back to the survey. Additionally, especially for the privacy and security links, it is important for the information provided to make respondents feel comfortable answering the more sensitive questions. Table 15 shows the number of respondents who clicked each of the links, the number who broke off after clicking a link, the number of respondents who clicked each link as a percent of individuals who accessed the instrument (Click Rate), and the breakoff rate after clicking a link.

Table 15. Frequency and Percentage of Respondents who Clicked each Link

Link Clicked	Number of Respondents Clicking Link	Number of Breakoffs After Clicking Link	Click Rate (%)	Breakoff Rate (%)
FAQs	461	27	2.4	5.9
Instructions	339	13	1.7	3.8
Contact us	99	9	0.5	9.1
Privacy	127	7	0.7	5.5
Security	61	6	0.3	9.8
Accessibility	27	5	0.1	18.5

Source: U.S. Census Bureau, ACS Internet Test, April to May 2011

As previously mentioned, one of the major concerns about some of the links was respondents not returning to the survey after selecting them³². However, this did not appear to be a large problem. In total, 18.5 percent of respondents who selected Accessibility broke off after selecting the link. However, this only totals five respondents, or 0.03 percent of the individuals that accessed the instrument. Further, 9.8 percent of respondents who clicked Security broke off, but again this only accounts for six respondents (0.03 percent of individuals that accessed the instrument). Finally, 5.5 percent of respondents who clicked the Privacy link broke off, accounting for only seven respondents (0.04 percent of individuals that accessed the instrument). It also is not surprising that FAQs and Instructions were clicked much more frequently than the other links since they were located in a more prominent location, just below the banner, as compared to the bottom corner of the page (Figure 12 in Section 5.5.2 can be used for reference).

³² In this case, breakoffs count if they clicked the specified link and did not visit at least one other page before leaving and did not return to the instrument.

It is interesting to note on which pages most respondents requested the Security, Accessibility, and Privacy links. While a handful of respondents clicked all the links on the page looking for information, this behavior is the exception rather than the rule. The majority of respondents only clicked one of the links on each page. Respondents clicked both the Security and Privacy links most frequently on the Date of Birth/Age, Roster screen, Wages Amount, and Respondent Name screens. For Accessibility, the majority of requests were on Date of Birth/Age, Pick Next Person, Review & Edit (review person link), and Wages Amount. We also have seen these screens as the top pages on which respondents broke off and logged out. As mentioned, some of these questions, especially those related to Security and Privacy clicks, are quite sensitive, so it makes sense respondents would want to review the Census Bureau’s policies for handling data.

5.5.2 Use of Review & Edit feature

Once respondents complete the ACS questions, they are given the option to submit the survey or to review their responses. When they choose to review, they are first presented with a Summary screen, which lets them select the sections they would like to review (Figure 12). They can either select the housing data or they can click on a household member’s name and be taken to the detailed review screen containing that person’s data (Figure 13). On the detailed review screens, respondents can click on an answer they want to change and they are taken to that screen with their current answer marked. After making a correction, they are either redirected back to the detailed review screen or they can use the “Next” button to proceed through the instrument again.

Figure 12. Summary Screenshot

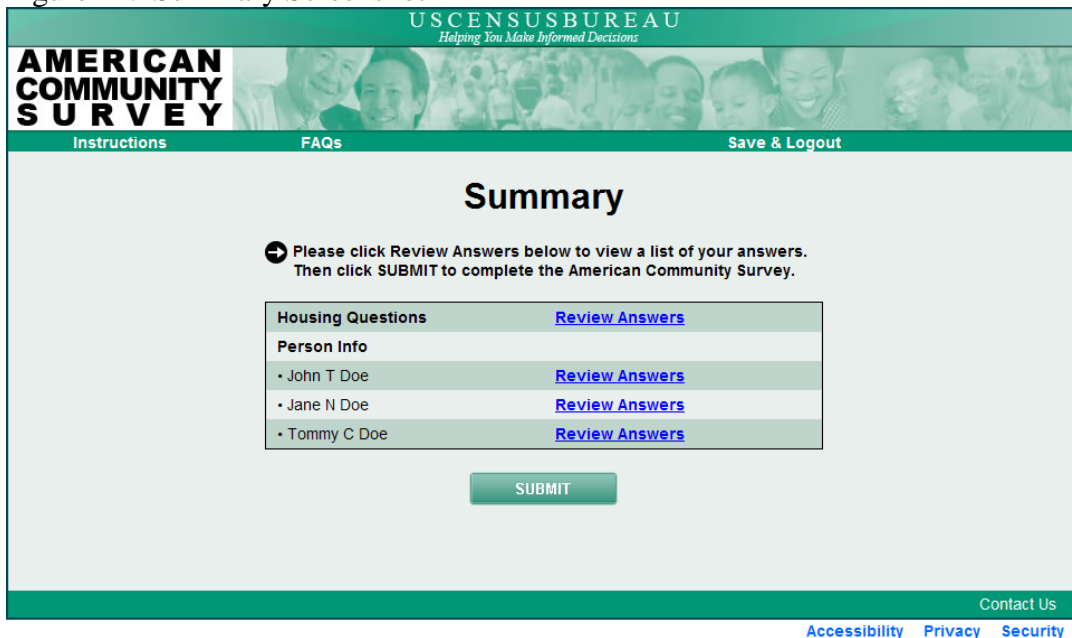
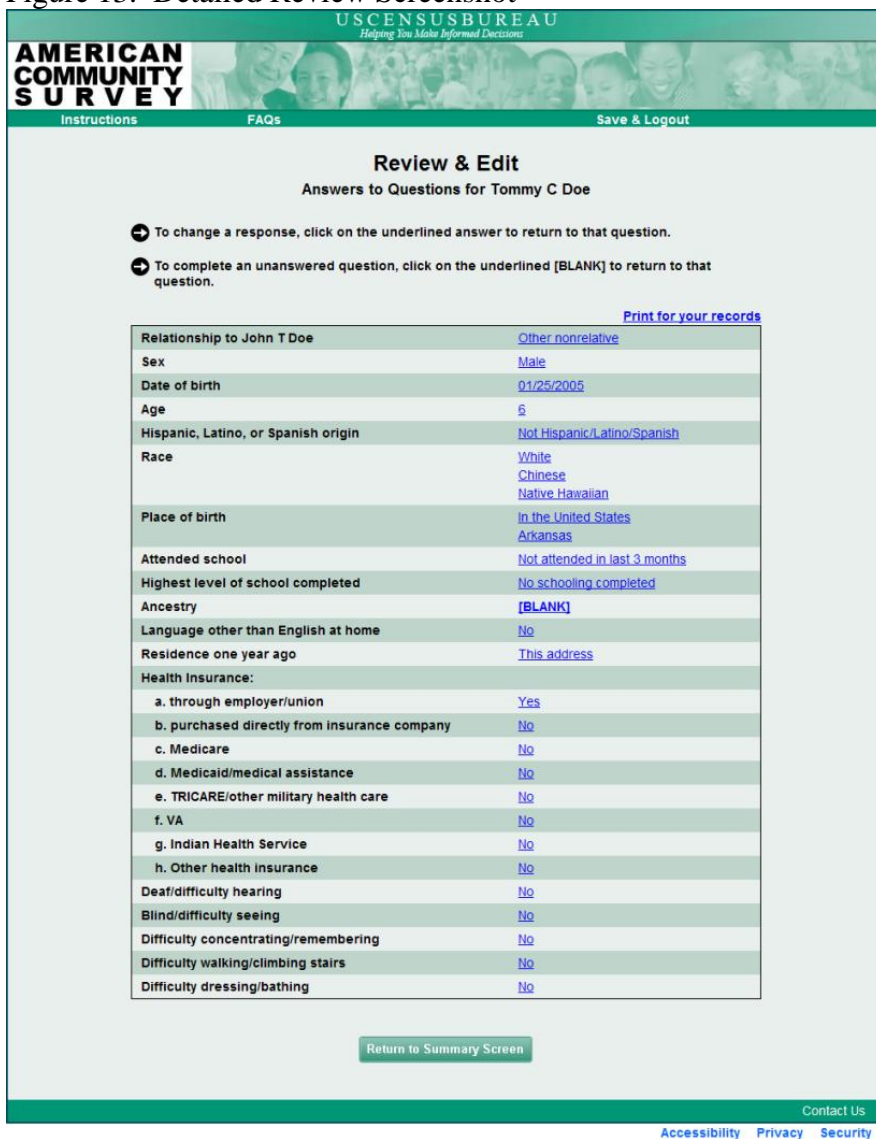


Figure 13. Detailed Review Screenshot



In total, 21.5 percent of the respondents who completed the survey (reached the “Presummary” screen) chose to review their answers (clicked “Review” instead of “Submit”). While not everyone that reached the Summary screen chose to review the detailed data, the vast majority did. Table 16 shows what they chose to review and whether they made changes. The “Percent of Reviews” column is the proportion of people who reviewed the housing and person sections of all the respondents who selected “Review”.

Table 16. Percent of Respondents that Chose to Review Each Section of the Instrument

Data Reviewed	Percent of Reviews (%)	% Changed
Housing	45.1	14.2
At least one person	92.2	50.6

Source: U.S. Census Bureau, ACS Internet Test, April to May 2011

Table 16 shows many more respondents reviewed the person data as compared to the housing data. This is not surprising as the person section contains more information than the housing section, so there is more room for error or needing to follow up with other household members. These results are very encouraging in that respondents that visit the Review & Edit section are using it to not only review, but also change their responses. Unfortunately, at this point in the data analysis we cannot say if respondents are changing answers or answering questions they had previously skipped. However, this could be an area of future research to help understand how people are going through the instrument and using the feature.

Although respondents are using Review & Edit, they are not necessarily improving data quality. Specifically, a handful of cases were identified where respondents used the feature to delete prior responses, such as Wages Amount and Date of Birth/Age.

Another concern was whether to provide a “Submit” link on the detailed review pages or force respondents to return to the Summary screen to submit the survey. The “Submit” link lets us know the respondent is completely finished with the survey. Once respondents select this link, they cannot return to the survey. It is possible for respondents to make it through the entire survey, but not formally submit their answers. Since we collected answers as respondents moved through the instrument, this was not an issue because any household who saw all of the screens was considered a completed case. However, if we did not track answers throughout the survey, we wanted to know how many cases would be lost. We thought that because the “Submit” link was not available on the Review & Edit pages respondents might just close out of the survey either because they could not find it or because it was not easily available to them. However, 97.7 percent of the people who reviewed their answers returned to the Summary screen and submitted the survey. Therefore, making people return to the Summary screen does not appear to be an issue.

5.5.3 Use of “Print” feature

From the detailed review screens, respondents have the option to print their responses (see Figure 13 for the location of the Print link). This feature was introduced because respondents occasionally ask ACS interviewers for copies of their answers, likely because it is a mandatory government survey and they want to have a record of completing it. Of the 3,464 respondents who reached the “Review & Edit” screens, 9.6 percent chose to print out their responses. Respondents had to print the housing and person data using the links on each respective page. Of the 332 individuals who printed their data, 80.1 percent printed the housing data, 88.3 percent printed the person data for at least one person, and 68.4 percent printed both the housing and person data for at least one person.

The print option appears to be a good feature to continue to use in the instrument. Further, only six respondents called TQA to request a copy of their answers, so it seems that the respondents who were interested in printing their results were able to find the link.

5.5.4 Use of “Previous” button

A “Previous” button appears on almost all of the pages in the instrument (other than the Live at this Address, PIN, Welcome Back, and Review/Summary screens). This function provides respondents with the ability to move backwards through the instrument to check or correct their answers³³. It appears that respondents did in fact take advantage of this opportunity. In total, the “Previous” button was clicked 89,630 times by 13,325 respondents. Additionally, 6,470 respondents clicked previous on at least two consecutive screens. This means respondents were not just going back to the previous question, but instead using the button to navigate back through the instrument.

Table 17 lists five screens on which the highest proportion of respondents used the “Previous” button. While these screens represent the highest proportion of clicks, the overall distribution of where respondents clicked “Previous” is evenly distributed across all of the pages in the instrument.

There is a little overlap with the answer changes after clicking “Previous” (Table 11). Specifically, Gas Amount appears in both tables. Additionally, a lot of the utility questions are in both tables, although the specific questions differ (whether they pay for water, how much they pay for water, if they pay for gas, and how much they pay for gas are all in Table 11). Part of the confusion with the Gas Amount question could be because the instructions for the previous question, whether they pay for gas, specifies: “Select “No” if gas and electricity are billed together and you included the gas payment in the electricity question.” Respondents may not read this part the first time through but then realize they do not know how to answer the amount portion of the question, so they go back. One solution to this problem, especially since we saw problems with the Electricity question as well (see Section 5.1.3) would be to put both questions on the same screen.

However, for the other questions listed in Table 17, it is not clear why these pages were the most common without additional research. In some cases respondents changed their answer, but in others they just returned to the next screen or continued to select “Previous”. Cognitive testing is needed to understand what information they were looking for or what their intentions were.

Table 17. Five Screens on which Respondents Selected “Previous”

Page	Frequency of Clicks on “Previous”	Percent of all Clicks on “Previous” (%)
Pay for water	2,546	2.8
Water amount	2,453	2.7
Pay for gas	2,080	2.3
Gas amount	1,516	1.7
Mortgage tax	1,491	1.7

Source: U.S. Census Bureau, ACS Internet Test, April to May 2011

³³ Respondents could also move backwards through the instrument using the back button on the browser. We were not able to collect this action using paradata.

Respondents seemed to exhibit unpredictable behavior when using the “Previous” button. The following are observations based on browsing through the data; they are not quantitative.

For respondents that went back at least three screens, the majority of backups occurred in the battery of income questions, the employment section, or the utilities section. Some respondents backed up to make a change and others simply returned to the original screen with no data edits along the way. Since most of these respondents do not break off after clicking “Previous” multiple times, it suggests they were just double checking an answer, possibly because they did not think a question applied to them, so they wanted to make sure the previous questions were answered correctly. Additionally, they may have thought the questions seemed repetitive, so they wanted to make sure they were requesting new information. More research is needed to understand why people backed up but made no changes.

There were a handful of respondents who returned to the roster section from all different sections of the survey. Some returned to edit the roster information (there was not a pattern between updating the roster and household size), others deleted all their information and broke off, while others returned and did not make any changes. This occurrence was the result of an error in the development of the instrument; the roster section should lock once completed.

The only potential pattern that was uncovered in this analysis is there did seem to be a fair number of respondents that used the “Previous” button after logging out and returning to the survey. This behavior did not seem to occur more frequently for any specific questions. It is possible these respondents looked up a piece of information and wanted to check/correct what they had originally answered.

From browsing the data, it appears respondents use the “Previous” button for three reasons: to change an answer, to check an answer, or to clear out previously entered data. A single respondent could engage in all of these behaviors. There were many respondents who backed up more than one or two consecutive screens in the instrument. This results in longer completion times and could increase the chances of the respondent breaking off before completing the survey. Therefore, it might be worth testing alerting respondents early in the instrument that they will be able to review their responses and make changes before submitting. Although similar language appears on the Finished Person screen, not every respondent reached that screen before backing through the instrument. The benefit of this would hopefully be fewer people backing all the way up in the instrument. On the other hand, if one of the changes results in different skips through the instrument, it is more likely the respondent will answer the new questions by using the “Previous” button than by changing things in the Review & Edit section. Testing could show which has a greater impact on data quality and user satisfaction.

5.5.5 Environment

The survey was accessed via a variety of devices, screen resolutions, operating systems, and browsers. Each will be discussed separately.

Devices

Table 18 provides a list of the devices that were used to access the survey and how frequently each was used. We did not create a version of the survey specifically for mobile devices, so these users had to work with a survey designed to fit on a monitor.

Table 18. Devices Used to Access the Instrument

Devices	Frequency	Percentage (%)	Breakoff Rate (%)
Personal Computer	18,874	97.3	15.8
Other Mobile Device	210	1.1	14.8
iPad	150	0.8	14.7
iPhone	36	0.2	13.9
Android	26	0.1	15.4
iPod Touch	6	0.0	33.3
BlackBerry	5	0.0	20.0

Source: U.S. Census Bureau, ACS Internet Test, April to May 2011

It is possible that additional devices were used that we did not know to look for. For example, there were several tablets on the market other than the iPad. However, we did not know how they were identified in the paradata and each one individually was not common enough to stand out. Therefore, all tablets and phones not mentioned specifically are accounted for in the “Other Mobile Device” category. While almost all respondents used a computer to complete the survey, just less than one percent did use the iPad tablet. With more tablets coming onto the market, this number should be increasing over time. While preliminary testing by the Decennial Statistical Studies Division staff showed that the screens appeared completely on an iPad, tablets come in different sizes, so developing an instrument that works well on the majority of these devices should be investigated.

The final column in Table 18 displays the breakoff rates for each type of device. Other than the iPod Touch and BlackBerry devices (which only account for 11 respondents), the breakoff rates are consistent with the personal computer breakoff rate. Past research has shown an average breakoff rate of 24 percent for mobile devices (Callegaro, 2010), so these findings are encouraging, especially since there was no mobile device version.

Screen Resolution

The most common screen resolution across all computer users is 1024x768 pixels, which is the standard size the Internet instrument was designed to fit. However, none of the Internet respondents had their computers set at 1024x768. Table 19 shows the resolution breakdown across all Internet respondents.

Table 19. Average Screen Resolution for Anyone who Accessed the Instrument

Width	Height	
	Less than 768 (%)	Greater or equal to 768 (%)
Less than 1024	29.2	1.0
Greater or equal to 1024	49.1	20.8

Source: U.S. Census Bureau, ACS Internet Test, April to May 2011

Unfortunately, all this table really tells us is that respondents use all types of screen resolutions, so the same instrument will look different on different computers. It is not possible to program the instrument for every height/width combination, but knowing that half of the respondents have their width set at 1,024 or greater and their height at less than 768 can help programmers determine how to develop the look of the instrument. Respondents who have smaller screen settings may experience more scrolling to see all the response options and links. Additionally, the question text and response option text may break in unexpected locations.

Operating Systems

Table 20 provides a list of the operating systems running on computers that accessed the Internet instrument. Not surprisingly, the vast majority (98 percent) of operating systems were either Windows or Macintosh. Therefore, there does not appear to be much need at this point of testing other systems.

Table 20. Operating Systems Used to Access the Instrument³⁴

Operating System	Frequency	Percentage (%)
Windows	23,547	85.9
Macintosh	3,223	11.8
Linux	155	0.6
Ubuntu	51	0.2
Presto	29	0.1

Source: U.S. Census Bureau, ACS Internet Test, April to May 2011

Browsers

Internet respondents used a variety of browsers to access the survey. However, the top browsers are what would be expected. Table 21 provides a list of browsers used, as well as their frequency and percentage. The “Other” category includes browsers such as Opera, RockMelt, Flock, and other browsers we did not know to search for. This table suggests that we need to continue to test the instrument on browsers other than Internet Explorer.

³⁴ The total number of operating systems and browsers used is greater than the 19,406 respondents that accessed the instrument for several reasons. First, respondents could access the instrument multiple times using different operating systems and/or browsers. Additionally, some sampled addresses may have logged into the instrument but never reached any of the questions either due to a system glitch or because they immediately closed the browser.

Table 21. Browsers Used to Access the Instrument

Browser	Frequency	Percentage (%)
Internet Explorer	17,322	63.2
Firefox	5,286	19.3
Safari	2,644	9.7
Chrome	1,955	7.1
Other	196	0.7

Source: U.S. Census Bureau, ACS Internet Test, April to May 2011

Note that it is possible the Internet Explorer (IE) frequency is an overestimate. The way useragent data are formatted, IE appears on every record, regardless of whether the user actually used IE. This occurs because Microsoft came up with the agent process. Therefore, to calculate the actual IE frequency, we delete the other browsers, such as Firefox, Safari, and other known browsers from the total. What we are left with is assumed to be the total IE users, but this cannot be guaranteed.

5.5.6 Average completion time

The time it takes a respondent to complete the Internet version of the ACS is dependent on household size and how many sessions it took them to complete the survey. Some individuals who accessed the survey logged out and then waited several days before returning to the instrument, likely due to receiving a reminder postcard. Therefore, response times that include individuals that completed in multiple sessions have large outliers and high standard deviations. To account for this, the analysis was divided into two groups, one with everyone together and one with just the people who completed in one session. The tables included in this section for the completion times by household size are only for the respondents who completed in one session since the results are more meaningful. The comparable tables for all respondents are available upon request.

Completion times were calculated for the total survey and separately for the demographic question section, housing section, detailed person section, and Review & Edit³⁵. Only people that saw every relevant question in each section are included in the calculations.

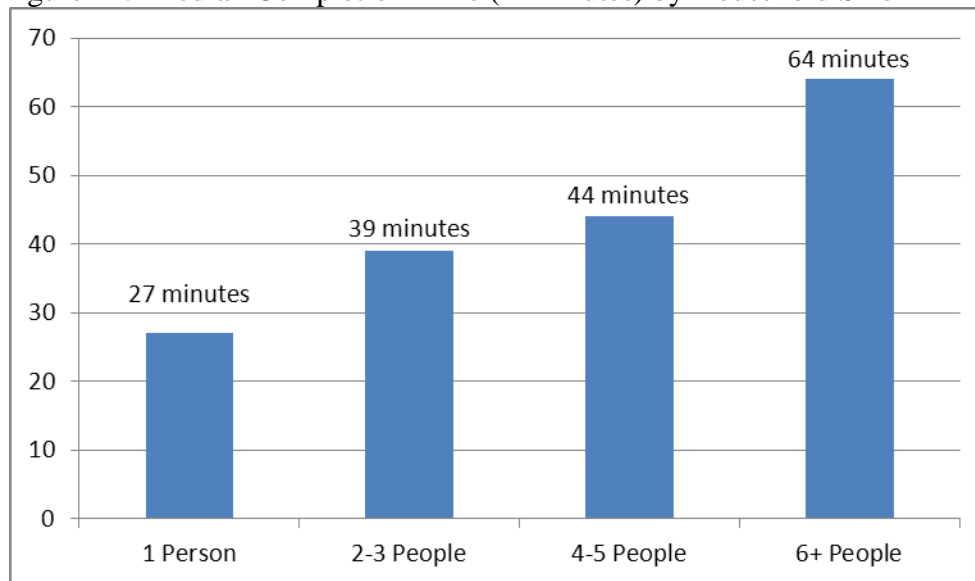
The median completion time was calculated by subtracting the time a respondent entered the “Address Verification” screen from the time he or she entered the “Presummary” screen. Included in this calculation are the individuals who viewed all of the applicable ACS questions in one session. On the PIN screen respondents are told the survey takes approximately 38.8 minutes to complete. Figure 14 provides the median completion time for the total survey by household size. For all respondents who completed the survey in one session, the median completion time was 39 minutes³⁶, which is in line with what is expected from prior analyses of completion times from all modes and all household sizes. However, if someone is using the Internet to complete the form, it is possible they also

³⁵ The Roster section was not included in any of the individual sections, so the sum of the section totals is not equal to the total completion time.

³⁶ The means were distorted by extreme values, which occurred because we allowed 15 minutes of inactivity before the system timed the user out. Therefore, the median completion time was used. The mean response time was 43 minutes.

have an email window open or additional websites, and they are moving back and forth between them. Although these individuals completed in one session, we cannot guarantee they were engaged with the survey the entire time. Therefore, actual completion times may be shorter than those displayed in Figure 14.

Figure 14. Median Completion Time (in minutes) by Household Size



Source: U.S. Census Bureau, ACS Internet Test, April to May 2011

In addition to the median completion time, we also looked at the completion times for the different sections of the instrument. Tables 22-24 provide the minimum, maximum, mean, and median completion times for the basic demographic, housing, and detailed person sections, respectively.

Table 22. Completion Time Details for Basic Demographic Section³⁷

Completion Time- Demographic Section	Family Size			
	1 Person	2-3 People	4-5 People	6+ People
Minimum	18 sec	29 sec	55 sec	1 min 50 sec
Maximum	29 min 55 sec	81 min 30 sec	49 min 30 sec	30 min 22 sec
Mean	1 min 4 sec	2 min	3 min 5 sec	4 min 56 sec
Median	48 sec	1.6 min	2 min 30 sec	3 min 56 sec

Source: U.S. Census Bureau, ACS Internet Test, April to May 2011

³⁷ The completion times for the demographic section have two starting points: the relationship question if the first person is the reference person and the household has more than one person or the sex question if the first person is not the reference person or if the household has one person. There is one less question for single-person households than the rest of the households.

Table 23. Completion Time Details for Housing Section

Completion Time- Housing Section	One Session
Minimum	51 seconds
Maximum	2 hours 13 minutes
Mean	9 minutes 6 seconds
Median	7 minutes 6 seconds

Source: U.S. Census Bureau, ACS Internet Test, April to May 2011

Table 24. Completion Time Details for Detailed Person Section

Completion Time- Detailed Person Section	Family Size			
	1 Person	2-3 People	4-5 People	6+ People
Minimum	2 min 12 sec	1 min 30 sec	4 min 52 min	13 min 30 sec
Maximum	1 hour 19 min	1 hour 50 min	2 hours 43 min	1 hour 58 min
Mean	15 min 35 sec	25 min 9 sec	30 min 22 sec	41 min 35 sec
Median	13 min 22 sec	22 min 41 sec	27 min 8 sec	35 min 44 sec

Source: U.S. Census Bureau, ACS Internet Test, April to May 2011

The proportion of time spent on each section meets expectations based on section length. The demographic section is the shortest section, in that it contains the fewest questions, and also has the least average time spent across all households (median completion time of 1.6 minutes across all household sizes). The housing section is longer than the demographic section but shorter than the detailed person section (median completion time of 22.8 minutes across all household sizes), and the average completion time for the housing section falls between the other two. The long maximum values most likely reflect individuals who were engaged in other activities while completing the survey or were looking up information to answer the survey questions, but were not idle for 15 consecutive minutes. Additionally, the survey length for the detailed person section depends on the age of the household members, so households with longer response times may consist of several individuals over 15 years of age, as these individuals receive all of the income, education, military, and disability questions, which younger groups bypass.

Time spent in the Review & Edit section also was measured. We previously looked at how many people visited this section and how often changes were made, but the time spent will help explain how thoroughly respondents reviewed their answers. The overall median time spent in Review & Edit was only 2.2 minutes. However, since over half the respondents that went to Review & Edit changed at least one answer in the detailed person section, and therefore most likely spent more time in that section than someone who did not review any of the detailed data (only reviewed housing), it makes sense to look at the groups separately (Table 25). The “Clicked a link in Review & Edit” column describes respondents who actually clicked a link to return to a page in the survey. The final column represents the respondents who selected to review their answers and went to the summary screen, but did not review any detailed data. They may or may not have clicked on a link to look at the detailed housing or person responses, but they did not select a link to return them to a question in the survey.

Table 25. Completion Time Details for Review & Edit Section

Completion time	Clicked a link in Review & Edit	No Detailed Links Clicked
Minimum	23 seconds	9 seconds
Maximum	66.1 minutes	50.7 minutes
Mean	4.5 minutes	2.5 minutes
Median	2.6 minutes	1.4 minutes

Source: U.S. Census Bureau, ACS Internet Test, April to May 2011

As expected, respondents who clicked a link to return to a page in the survey spent more time in the section. However, both of the durations are relatively short. Considering respondents were able to scan through their responses and make changes in a relatively short amount of time suggests the layout and navigation of the Review & Edit section is simple for respondents to understand and use. On the other hand, it may have been too hard to follow so they did not spend much time trying. Cognitive testing would be able to determine the usability of this feature.

5.5.7 Reasons for calling TQA

In total, there were 22 Internet-related reasons for calling TQA. Those discussed in this section and in Section 5.3 account for 90 percent of the calls (the rest of the reasons can be found in Appendix C). None of the additional reasons seem to be problematic for the survey moving forward.

TQA received 6,664 calls from sampled Internet Test households. Of those calls, 35.4 percent were related to the Internet (see Appendix C). The most common reason respondents called TQA was because they did not have Internet access (14 percent of Internet calls). An additional 9.2 percent called to request a paper questionnaire and 2.5 percent said they were unwilling to complete the survey using the Internet (Roberts 2012). The bulk of these calls were from the Push panels, although there were still a handful of calls from the other panels. Although there is language in the Push letter telling individuals they will receive a paper questionnaire in a future mailing, it would seem that these individuals did not read the letter or notice this sentence. Therefore, it may be worth experimenting with re-wording the post card to add similar language as is in the Push letter which informs sampled households that they will receive a paper questionnaire or make the language more prominent in the first mailing.

6. SUMMARY

This paradata analysis was intended to help us understand how respondents interact with the Internet instrument and to identify any potential problems with specific questions or with the instrument as a whole. To do this, we looked at completion times, authentication, breakoffs, error messages, clicked links, answer changes, and completion behavior. The overall finding from this research is that the Internet instrument is working as we expected it would. Respondents do not seem to be having too much difficulty accessing the instrument; they are able to navigate through the instrument and use the available features, the majority complete the survey in one session, and they are taking an expected amount of time to do so.

While the instrument is not problematic for respondents, we did uncover several specific screens that could be modified to help respondents and reduce breakoffs. Specifically, the transitional screens have higher logout and breakoff rates than the other screens in the survey. Altering the wording on these screens or combining them with a screen where a response is required could help keep respondents engaged so they do not see the pages as a break or resting place. Other problematic screens included Date of Birth/Age, Wages Amount, and Respondent Name. Some of the issues may stem from privacy concerns on the part of the respondent. In the case of Date of Birth/Age and Respondent Name, they might not think all parts of the question are required or even notice they have two tasks to complete. Therefore, future research could involve testing different strategies to reduce respondent error on these screens.

Table 26 provides an updated look at the Future Research Criteria table presented in Section 2. This table has an additional column that explains the results of each criteria and our future testing plans.

Table 26. Future Research Criteria

Criteria	Result	Recommendations for Future Research
Overall breakoff rate more than 10 percent	Breakoff rate was 15.8 percent. Breakoff rates increased with household size (Table 2 page 11).	Determine the impact of reminder mailing for partial respondents.
Breakoff rates for individual screens are outliers, as compared to other screens in the survey (note screens where progress indicators change)	The breakoff rate on the Pick Next Person screen accounted for the highest proportion of all breakoffs and was also the highest relative to how many people saw the screen (Table 3 page 11).	Test the impact of combining the Pick Next Person and Finished Person screens on breakoffs; consider new wording on the combined screen to encourage continuation in the survey.
Help used in less than 10 percent of cases	40 percent of respondents selected Help at least once and Help was selected on 97 percent of the screens on which it was available (page 38).	No additional testing is needed, however we should look at the question wording and comprehension for questions on which Help was most frequently selected.
A considerable percentage of Internet users have multiple unsuccessful re-entries	Only 1.8 percent of the individuals that accessed the instrument had an invalid entry and 77.4 percent of those people had only one. There were 672 calls to TQA regarding initial access problems (pages 40 and 41).	While the data did not suggest authentication was a serious issue, we are unable to know whether there were more issues with initial access and people not attempting to re-enter because they did not know their PIN (as opposed to not wanting to return or forgetting). Therefore, we think it is valuable to continue testing this structure.
A notable percentage of respondents access the survey via a mobile device	2.7 percent of respondents used a mobile device (Table 18 page 50).	No imminent research is needed, however, this is part of future plans given the expected rise of mobile device usage.

Criteria	Result	Recommendations for Future Research
Less than five percent of respondents use Review & Edit feature	21.5 percent of respondents who completed the survey reviewed their answers (page 46).	No additional testing is needed
Median completion time greater than 45 minutes	The median completion time across all households was 38.8 minutes (page 52).	No additional testing is needed
High percentage of changed answers on a particular screen, compared to other screens	There were a handful of screens that had high changed answer rates, including Educational Attainment and several questions from the roster and utilities sections (Tables 10 and 11 pages 30 and 33).	Additional research will be conducted.
Screen renders a high percentage of error messages, compared to other screens	Five screens were identified as having a relatively high percentage of errors. In particular, Place of Birth and Date of Birth/Age had the highest by a factor of two (Tables 4 and 5 pages 17-18).	Additional research will be conducted to change the way these questions are asked and how the errors are displayed to help respondents answer the questions more accurately and completely.
Respondents finish the survey but fail to submit	97.7 of respondents who reviewed their answers formally submitted the survey (page 47).	No changes needed.

While the Internet instrument worked well, especially for a first test, there is still room for improvement. Further research regarding breakoffs, error messages, and changed answers should help make the instrument easier for respondents to use as well as provide the ACS with superior data quality.

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: Brenna Matthews, Michelle Ruitter, Andrew Roberts, Brian Wilson, Tony Tersine, John Studds, Chris Butler, Joe Mistichelli, Brian Ridgeway, Michael Coan, Fred Meier, and Samantha Barron.

References

Ashenfelter, K., Holland, T., Quach, V., and Nichols, E., (Forthcoming). Final Report for the Usability Evaluation of the ACS Online Instrument Rounds 4a and 4b. Center for Survey Measurement, Research and Methodology Directorate. U.S. Census Bureau.

Ashenfelter, K., Holland, T., Quach, V., and Nichols, E., (2011). Final Report for Rounds 1 and 2 of Usability Testing of the 2010 Census Quality Survey Online Instrument in Conjunction with the 2010 Decennial Census. Center for Survey Measurement, Research and Methodology Directorate (Survey Methodology #2011-08). U.S. Census Bureau. Available online at <http://www.census.gov/srd/papers/pdf/ssm2011-08.pdf>.

Bentley, M., and Tancreto, J. (2006) "Analysis of Self-Response Options and Respondent-Friendly Design from the 2005 National Census Test," *2006 Proceedings of the American Statistical Association*, Survey Research Methods Section [CD-ROM], Alexandria, VA.

Callegaro, M. (2010). Do you Know Which Device your Respondent has Used to take your Online Survey? *Survey Practice*, December.

Conrad, F., Couper, M., and Tourangeau, R. (2003). "Interactive Features in Web Surveys." Paper presented at the Joint Statistical Meetings of the American Statistical Association, San Francisco, CA.

Conrad, C., Couper, M., Tourangeau, R., & Peytchev, A. (2006). "Use and Non-use of Clarification Features in Web Surveys." *Journal of Official Statistics* 22(2):245-269.

Couper, M., Kennedy, C., Conrad, F., Tourangeau, R. (2011). "Designing Input Fields for Non-Narrative Open-Ended Responses in Web Surveys." *Journal of Official Statistics* 27(1):65-85.

Couper, Mick P., Michael W. Traugott, and Mark J. Lamias. (2001). "Web Survey Design and Administration". *The Public Opinion Quarterly* 65:230-53.

Couper, Mick P. (1998). "Measuring Survey Quality in a CASIC Environment." *Proceedings of the Section on Survey Research Methods of the American Statistical Association*. Dallas, TX.

- Dillman, D. A., and Bowker, D. K. (2000). The Web questionnaire challenge to survey methodologists [Online]. Available: <http://survey.sesrc.wsu.edu/dillman/papers.htm>
- Fish, S., (2011). Evaluation of Multiple Responses. American Community Survey Research and Evaluation Program. Available: http://www.census.gov/acs/www/Downloads/library/2011/2011_Fish_01.pdf
- Galesic, M. and Bosnjak, M. (2009). Effects of Questionnaire Length on Participation and Indicators of Response Quality in a Web Survey. *Public Opinion Quarterly* 73(2): 349-360.
- 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.
- Groves, R.M., Kahn, R.L. (1979). *Surveys by Telephone*. New York: Academic Press.
- Heerwegh, Dirk. (2003). Explaining Response Latencies and Changing Answers Using Client-Side Paradata From a Web Survey. *Social Science Computer Review*. 21: 360-373.
- Heerwegh, Dirk (2002). "Describing Response Behavior in Web Surveys using Client Side Paradata." Paper presented at the International Workshop on Web surveys. Mannheim, Germany.
- Jeavons, A. (1999). Ethology and the Web. Observing Respondent Behaviour in Web Surveys. Proceedings of the ESOMAR Worldwide Internet Conference Net Effects 2 London 1999.
- Keathley, Don (forthcoming). Verification and Validation Checks for the 2011 Internet Test Follow-Up Sample Selections. DSSD 2011 American Community Survey Memorandum Series #ACS11-S-30.
- Lind, L., Schober, M., & Conrad, F. (2001). "Clarifying Question Meaning in a Web-Based Survey." Proceedings of the Annual Meeting of the American Statistical Association, August 5-9.
- Lozar-Manfreda, K., Vehovar, V. (2002). Do Web Surveys Really Gain Lower Response Rates than other Survey Modes? A Meta-Analysis. Web survey Workshop and Symposium.
- Matthews, B. (2012a). "Item Nonresponse Rates". Email message dated 6/11/12.
- Matthews, B. (2012b). "Form Completeness/Breakoffs". Email message dated 3/21/12.

Mooney, G.M., Rogers, B., & Trunzo, D. (2003). Examining the Effect of Error Prompting on Item Nonresponse and Survey Nonresponse in Web Surveys. Proceedings of the American Association for Public Opinion Research Conference, Nashville, TN, May 15-18.

Musch, J., Reips, U. (2000). A Brief History of Web Experimenting. Psychological Experiments on the Internet. Birnbaum, M.H. (ed.), Academic Press: San Diego.

Peytchev, A. (2009). Survey Breakoff. *Public Opinion Quarterly* 73(1): 74-97.

Redline, Cleo, Tourangeau, Roger, Couper, Mick, Conrad, Fred, and Ye, Cong. (2009). The Effects of Grouping Response Options in Factual Questions with Many Options. Federal Committee on Statistical Methodology (FCSM).

Roberts, A. (2012). "TQA". Email message dated 6/11/2012.

Stokes, S.L. (2011), "Lessons Learned from the Design and Implementation of the 2010 Census Quality Survey (CQS)," DSSD 2010 Decennial Census Memorandum Series #O-A-40, U.S. Census Bureau.

Tancreto, J.G., Zelenak, M.F., Davis, M., Ruitter, M., Matthews, B. (2012). 2011 American Community Survey Internet Tests: Results from the First Test in April 2011. http://www.census.gov/acs/www/Downloads/library/2012/2012_Tancreto_01.pdf

Tancreto, J.G., Davis, M., Zelenak, M.F. (2011). Developing an Internet Response Mode for the American Community Survey. Poster presented at the annual meeting of the American Association for Public Opinion Research. Phoenix, AZ.

U.S. Census Bureau (2009), "Design and Methodology, American Community Survey," 7-5, April 2009. http://www.census.gov/acs/www/Downloads/survey_methodology/acs_design_methodology.pdf

Yan, T., and Tourangeau, Roger. (2008). Fast times and easy questions: The effects of age, experience and question complexity on web survey response times. *Applied Cognitive Psychology* 22(1): 51-68.

Yan, T., Conrad, F., Tourangeau, R., Couper, M. (2011). Should I Stay or Should I go: The Effects of Progress Feedback, Promised Task Duration, and Length of Questionnaire on Completing Web Surveys. *International Journal of Public Opinion Research* 23(2).

Appendix A: Paradata Collected

Category	Paradata
Survey Access	- Number of users that access the survey via a mobile browser
Login/Authentication	- Number of times an incorrect User ID is entered
	- Number of times an incorrect User ID/PIN combination is entered per User ID
	- Number of times a respondent is locked out of the survey because of an incorrect User ID/PIN combination
	- Number of times a respondent tried to come back into a survey that was already completed and submitted
Spanish Link on ACS	- Number of Times the Spanish Link was clicked
English Link on PRCS	- Number of Times the English Link was clicked
Section Date & Timestamps (additional timestamps are in sections below)	<i>(Used to determine length of time a respondent spends on sections of the instrument)</i>
	- Start of Session; Start of Login Process
	- If Login is successful: End of Login Process for Returning Users - (May help to determine how long it took the respondent to find and enter the correct User ID and/or PIN. However, this is confounded by Internet connection, speed, traffic, etc.)
	- Start of Roster Section
	- End of Roster Section; Beginning of 100% Demographic Items
	- End of 100% Demographic Items; Start of Housing Questions
	- If Person 1: End of Housing Questions, Start of Detailed Questions for Person 1; If Person 2 through Last Person: End of Previous Person's Questions, Start of Detailed Questions for Next Person
	- End of Last Person's Questions when any answers were left blank; Start of Review/Submit Process

	<ul style="list-style-type: none"> - End of Housing Questions when no one lives at the address or no one is on the roster; End of Last Person's Questions when no answers were left blank; End of Skipped Question process; Start of Review Process
Save and Logout	On which screen did respondent click "Save & Logout"
	<ul style="list-style-type: none"> - Number of respondents that initiated a breakoff by clicking "Save & Logout"
Session Timed Out	<ul style="list-style-type: none"> - The screen at which point the session timed out.
	<ul style="list-style-type: none"> - When did the session time out
	<ul style="list-style-type: none"> - Number of Sessions timed out per User ID
End of Survey	<ul style="list-style-type: none"> - When did Respondent click "Exit Survey"
	<ul style="list-style-type: none"> - On which screen did respondent click "Exit Survey"
	<ul style="list-style-type: none"> - Number of respondents who clicked "Exit Survey"
	<ul style="list-style-type: none"> - On which screen did respondent click "Submit"
	<ul style="list-style-type: none"> - When did Respondent click "Submit"
	<ul style="list-style-type: none"> - Number of respondents who clicked "Submit"
Number of Sessions	<ul style="list-style-type: none"> - Number of Sessions per User ID
	<ul style="list-style-type: none"> - Number of respondents that completed the survey in multiple sessions
	<ul style="list-style-type: none"> - Average number of sessions needed to complete the survey
	<ul style="list-style-type: none"> - Time between sessions for a respondent
"Previous" Button	<ul style="list-style-type: none"> - On what screens does a respondent clicks "Previous"
	<ul style="list-style-type: none"> - Number of times a respondent clicks "Previous"
Answer Changes	<ul style="list-style-type: none"> - On which screen(s) does a respondent change an answer (radio button)
	<ul style="list-style-type: none"> - Number of times a respondent changes an answer on a particular screen
	<ul style="list-style-type: none"> - What screens a respondent visits more than once

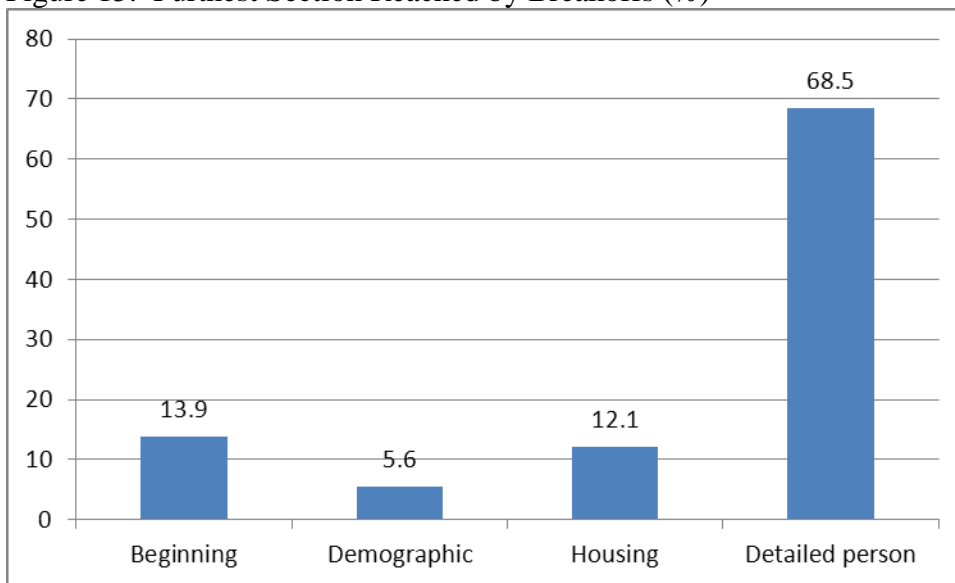
Error Messages	- On which screen(s) does a respondent get an error message
	- Number of times a respondent gets an error message
	- Number of times a respondent selects an answer after receiving an error message
	- Aggregate number of times an error message is rendered on a particular question screen
Help	- For what questions does a respondent click "Help"
	- Number of questions for which a respondent clicks Help but does not give an answer
	- Number of times a respondent clicks "Help"
	- Number of times "Help" is clicked for a particular question
Review/Edit	- Demographic characteristics of respondents who clicked "Review" on the Presummary Screen
	- Number of Respondents that clicked "Review" on Presummary Screen
	- Demographic characteristics of respondents who clicked a "Review Answers" Link on the Summary Screen
	- Number of Respondents that clicked a "Review Answers" Link on Summary Screen
	- For which question did a respondent click an answer to review/edit from the Review & Edit screen
	- Number of times a respondent clicks an answer to review/edit from the Review & Edit screen
	- Number of times a respondent goes to a question from the Review & Edit screen and selects "Help"
	- For which question(s) does the respondent change an answer after seeing the Review & Edit screen?
	- Number of times respondent clicks an answer to review/edit from the Review & Edit screen and changes an answer
Other Links	- On which screen did respondent click "Contact Us" button
	- Number of respondents that click "Contact Us"
	- On which screen did respondent click "Instructions"
	- Number of respondents that click "Instructions"
	- On which screen did respondent click "FAQs"
	- Number of respondents that click "FAQs"

	- On which screen did respondent click "Privacy"
	- Number of respondents that click "Privacy"
	- On which screen did respondent click "Accessibility"
	- Number of respondents that click "Accessibility"
	- On which screen did respondent click "Security"
	- Number of respondents that click "Security"
Write-in	- Number of characters in a write-in (<i>We would compare this to responses on paper before truncation.</i>)
	- Number of times a respondent does not specify an answer
	- Number of questions left blank by respondent (before and after receiving an error message)
Multiple Responses	- Compare questions with multiple responses on paper and Internet (check boxes and write-in responses)

Appendix B: Alternative Breakoff Rate by Section

In Section 5.1.1 we provided a distribution of breakoffs by section of the instrument. This breakoff was based on the last page the respondent saw before exiting the survey. Another way of looking at breakoffs is by measuring the furthest point in the survey the respondent accessed. For example, a respondent could answer several questions in the housing section before clicking “Previous” and returning to the roster section, where they then left the survey. In this case, they would have broken off in the roster section, because that is the last screen they saw. Alternatively, the furthest location is in the housing section. Therefore, Figure 15 provides an alternate view of Figure 5 and shows the furthest location respondents reached before breaking off.

Figure 15. Furthest Section Reached by Breakoffs (%)



Source: U.S. Census Bureau, ACS Internet Test, April to May 2011

This distribution is similar to that seen in Section 5.1.1. The majority of respondents reached the detailed person section prior to breaking off. The largest difference between the two figures is that more respondents reached the detailed person section than actually broke off there. Similarly, more respondents broke off in the beginning section of the survey than reached only that point. This difference means respondents are backing up in the instrument before breaking off. This phenomenon was seen and discussed in the use of the “Previous” button (Section 5.5.4).

Appendix C: Reasons for Calling TQA

Reason	Frequency
Difficulty with Questionnaire	428
Did not receive Questionnaire	150
Claims Filed	505
Lost Questionnaire/Envelope	86
Requested Envelope Mailed	21
General Question on Mailing	260
General Information on Survey Meaning	556
Question on Participant Selection	415
Address Wrong	222
Address Vacant	291
Address May be Group Quarters	71
No Living Quarters - Commercial Only	34
Question on Mandatory/Title 13	210
Request copy of Title 13	28
Confidentiality Concerns	228
Privacy Issue	220
Survey Unnecessary	103
Refuses to Participate	165
Congressman Mentioned	13
Residency/Household Membership Mentioned	46
Question on Specific Questionnaire Item(s)	479
Interview Conducted	1,604
Negative reaction to language brochure	12
Positive reaction to language brochure	0
Language Problem	103
Call from Congressman	2
See if Government	133
Old Age, Illness or Death	193
Requested Personal Visit	12
Refuse to Answer Certain Question	75
Will Comply	1,027
Pen or Pencil	28
Decennial Conflict	137
IVR Complaint or Suggestion	4
Temporarily Occupied	17
Alternative Questionnaire Requested	52
Field rep Complaint or Issue	2
Other	589
*Lost PIN	31
*Problems entering URL	436

Reason	Freq
*Need help finding survey website	404
*Need help finding User ID	179
*Need help finding PIN	37
*Need case reset	39
*Never accessed, no PIN	15
*Technical questions - system requirements	40
*Do not have Internet access	932
*Request paper questionnaire	611
*Unwilling to do on Internet	165
*Website slow	22
*Internet security concerns	66
*Navigation issues -- application presented inappropriate questions	21
*Request copy of answers entered on-line	6
*Request confirmation of receipt of data	15
*Interview too long on Internet	15
*Internet not user friendly/accessibility	75
*Incorrect roster/data	5
Time frame for completing survey	22
Tried to access survey after closeout/time limit	4
*What languages are offered on the Internet	9

Source: Roberts 2012

*Reasons that are related to the Internet