

## Use of Paradata to Assess the Quality and Functionality of the American Community Survey<sup>1</sup>

Rachel Horwitz, Jennifer Guarino Tancreto, Mary C. Davis, Mary Frances Zelenak  
U.S. Census Bureau, 4600 Silver Hill Road, Washington, D.C. 20233

### Abstract

Many of the recent discussions of paradata have focused on survey operations and nonresponse weighting adjustments. In this capacity, paradata have been used to optimize call back schedules, and interviewer observations have been used to supplement information from non-responding households. However, using paradata generated by the respondents has not been given as much attention. Collecting this type of information can lead to a better understanding of how a respondent interacts with and understands a survey, as well as provide researchers with tools to reduce measurement error.

As part of the April 2011 American Community Survey (ACS) Internet Test, the Census Bureau inserted JavaScript code into every page of the online instrument to capture a vast amount of paradata, along with respondent answers. Specifically, all clicked links were captured (navigational buttons as well as radio buttons, check boxes, help, etc.), along with timestamps, field values, errors rendered, invalid logins, timeouts, and logouts. These data were collected to help evaluate the quality of the ACS Internet instrument by making sure respondents were using the instrument as expected. Additionally, these paradata can help identify problematic screens or other issues with the instrument that might impact data quality.

This paper analyzes the paradata collected during the 2011 ACS Internet Test. Specifically, it discusses the authentication procedures, problematic screens (as determined by errors, use of help, and breakoffs), response times, and data quality indicators. It will discuss the potential problems with the instrument, things that worked well, and plans for future research based on the findings from this test.

**Keywords.** Paradata, Internet, Quality, Data Collection, Content

### 1. Background

Paradata have been defined as process data, or all the data collected during the response process that does not include the response itself (Couper, 2000). 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 (toys in a yard). In computer-assisted self-administered surveys, such as Internet surveys, they include location of breakoffs, changed answers, error messages, mouse clicks, and location and quantity of links clicked, including help. 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 respondents use to complete the survey, which can help identify ways to make the survey task less burdensome.

---

<sup>1</sup> This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress. Any views expressed on statistical, methodological, technical, or operational issues are those of the authors and not necessarily those of the U.S. Census Bureau.

In April 2011, the ACS conducted an Internet test intended to measure the impact on self-response rates of offering an Internet mode. Paradata were collected as a by-product of this test, which allowed us to assess the quality of the instrument design and identify potential modifications to make the instrument easier to use and more attractive to potential respondents. Many researchers have used paradata from Web surveys in this manner (Heerwegh 2003, Couper *et al.*, 2001, Redline *et al.*, 2009, Dillman and Bowker, 2000). Specifically, Heerwegh (2003) analyzed response times to help determine whether participants knew the correct answer to a question. Additionally, Heerwegh (2002) found that changed answers could indicate usability problems with a question and Jeavons (1999) found that the number of changed answers, backtracks, and confusion due to response option format led to higher rates of abandoned interviews. Identifying potentially problematic screens can not only lead to higher quality data for those specific items, but better overall response across the survey.

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 were a product of extensive research, laboratory testing, and two field tests, but has not yet been used in ACS production. This paper analyzes the paradata from the first of the two field tests. These data allowed us to fully evaluate how well the instrument worked and whether there are modifications that need to be made due to problematic questions or features. In this paper, we discuss general findings from our paradata analysis, many of which support findings from prior research<sup>2</sup>. For more general information on the April 2011 ACS Internet Test, please see Tancreto *et al.* (2012).

## 2. Analysis

### 2.1 Breakoff Rates

A break-off is defined as anyone who begins a survey but does not finish it. For the purposes of this paradata analysis, breakoffs are defined as any respondent that accessed the survey (had a successful login), but did not reach the presummary screen, which appeared after the respondent had seen all applicable screens for all people in the household. A total of 19,406 households accessed the instrument and 16,253 completed the survey on the Internet, resulting in an overall breakoff rate of 15.8 percent.

This break-off rate is on the low end of many Internet studies. Specifically, two meta-analyses have been conducted on breakoff rates in Web surveys. One reported average breakoff rates of 16 percent and the other 34 percent (Musch and Reips, 2000; Lozar-Manfreda and Vehovar, 2002). It is encouraging that the breakoff rate for the ACS Internet Test is not higher than the average. The breakoff rate may have been reduced as a result of the ACS being a mandatory government survey. It should also be noted that 10.9 percent of the respondents categorized as Internet breakoffs ended up completing the survey using a mail form.

We also calculated a breakoff rate by household size (Table 1). The breakoff rate is calculated as the total number of breakoffs for each size household as compared to the total number of respondents in each size of household.

---

<sup>2</sup> The rates provided in this paper are unweighted estimates. Weights were not used due to the experimental treatments; some of the Internet respondents received a paper form and an invitation to participate via the Internet at the same time, whereas others only had the option to respond by Internet (although they later would receive a mail form). Therefore, the Internet respondents for this test are not generalizable to what we expect in production.

Table 1. Breakoff Rate by Household Size

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

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

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

### 2.1.1 Breakoff Rate by Screen

The breakoff rate itself does not provide any information on which questions were potentially difficult for respondents. Therefore, in addition to the overall rate, an instrument breakoff rate and screen breakoff rate were calculated as well. The screen breakoff rate compares the number of breakoffs on each screen to the total number of visits to that screen, whereas the percent of breakoffs compares the number of breakoffs on each screen to the total number of breakoffs in the instrument.

Table 2 provides eight screens with some of the highest breakoff rates, which account for a third of all breakoffs.

Table 2. Cumulative Breakoff Rates by Screen

<b>Breakoff Page</b>	<b>Breakoff Frequency</b>	<b>Number of Screen Visits</b>	<b>Screen Breakoff Rate (%)</b>	<b>Percent of Breakoffs</b>
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 <sup>3</sup>

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

These screens fall into two major groups: transitional pages and sensitive questions. The Pick Next Person, Finished Person, and PIN screens are all transitional, whereas the Respondent Name, Wages Amount, Contact, and Contact Later are all potentially sensitive. Transitional screens provide the respondent with information, but do not ask a survey question and there is no survey task for the respondent to complete. On the Pick Next Person screen, respondents can select the next household member for which they

<sup>3</sup> Less than 0.05 percent of breakoffs occurred on this page.

answer questions. While this was designed to increase flexibility if the respondent did not know the answers for one or more people in the household, the language on the page explicitly tells them to log out if they cannot answer for someone in the household. Similarly, the Finished Person screen informs the respondents that their answers have been saved and they will be able to edit their responses at the end of the survey. For all three 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. Additionally, the PIN screen states that the survey takes an average of 38 minutes to complete, which may have deterred some respondents. One way of reducing breakoffs on these transitional pages is to also include a survey question so there is not an obvious break.

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 being the result of an additional participation request. Respondents decided to participate in the survey when they first began, but since 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 (2009) suggests this is especially true of pages where there is not a question displayed, as on the Finished Person screen.

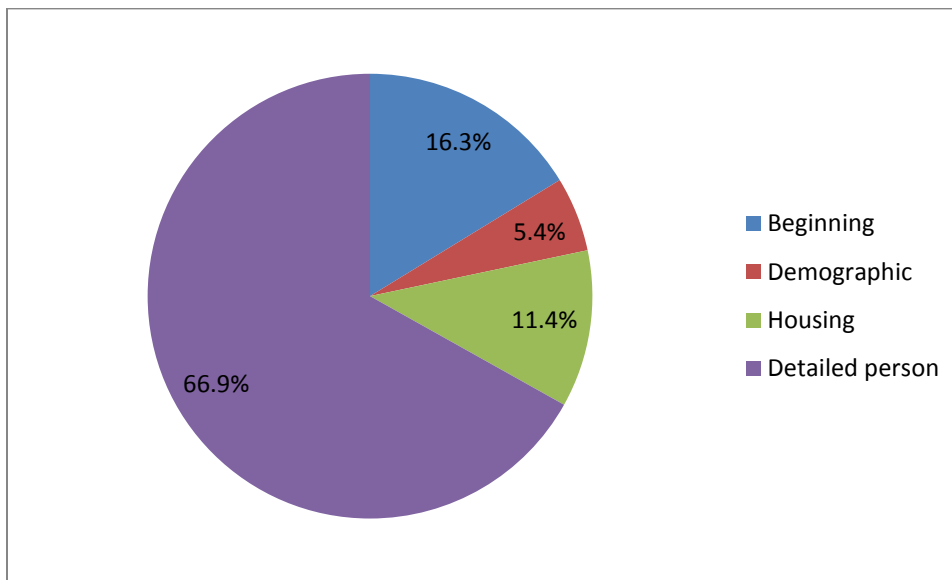
Additionally, it is very possible that respondents get to the transitional pages and simply do not know the responses for other household members and therefore leave the instrument before starting the next set of questions. This is likely more common for unrelated households than for related households. The data support this hypothesis for Internet-reporting households, as they are more likely to not 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 cases, whereas relatives were not reported in approximately 11 percent of cases and self reports were blank in just over 2 percent of cases. This discrepancy is not seen in mail, however. It may be that mail reporters are more likely to give the questionnaire to other household members because it is easier than providing other household members with the URL and login information.

The other questions in Table 2 ask sensitive information, such as name, addresses, phone numbers, and the amount received in wages. One major problem with breakoffs on these screens is that all of them, other than Wages Amount, appear at the very beginning of the survey, so we are not able to collect any data on these households. It is likely respondents simply do not want to provide this type of identifying and personal information. Additional text on why name and phone number are needed, along with security or privacy information, may help respondents be more comfortable sharing both these pieces of information. Further, some respondents may not know they can skip a question and still continue with the survey. While we do not want to give respondents permission to not answer questions, it is also important to keep them in the instrument.

### **2.1.2 Breakoff Rates by Section**

While the breakoffs by screen identify which pages may be problematic, we also looked at where in the instrument breakoffs occurred. Figure 1 provides a distribution of where in the instrument respondents broke off (beginning, demographic questions, housing questions, and detailed person questions) shown in sequential order.

Figure 1. 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. People who break off during the beginning roster section might just be seeing what the survey is about. Once they see the types of questions being asked, they break off. We expect this is due to needing to provide names or realizing the survey will be time consuming, but more research is needed to know exactly why people leave in the roster section. The demographic section is very short and the questions are fairly basic, so it follows there would be fewer breakoffs. Similarly, other than a series of questions on the amount paid in utilities and rent or mortgage, the housing section is not invasive. However, respondents could grow frustrated if they do not know the answers to these questions. Finally, the detailed person section has the most questions and may be difficult for one person to answer for everyone in the household, especially when the respondent is not related to the other household members and simply does not know the answers to the questions. 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 breakoffs occur in this section.

### 2.1.3 Summary of breakoffs

As with any survey, breakoffs occurred throughout the instrument for a variety of reasons. However, we were able to identify a handful of screens on which breakoffs occurred most frequently. These screens were either transitional in nature or asked for sensitive information. Overall, the breakoff rate was not alarmingly high and the majority of the breakoffs occurred where we expected them to. However, these breakoffs are still problematic for item nonresponse and data quality, so adjustments need to be made to keep respondents in the instrument.

## 2.2 Error Messages

The April 2011 Internet Test instrument had an error message associated with all of the screens containing questions deemed critical for the ACS, screens that included a skip

pattern, and all questions asking for a dollar amount<sup>4</sup> (56 screens). These error messages were soft edits in that the respondent did not need to make any changes in order to proceed. In total, there were 24,001<sup>5</sup> error messages rendered throughout the instrument. About 45 percent of respondents received at least one error message and those that received at least one, received an average of three. Additionally, about 2.4 percent of all the screens with a possible error message that were visited resulted in the display of an error message. For this analysis, however, we identified six screens on which the highest percent of error messages were rendered. Table 3 shows where the highest proportion of errors are occurring while Table 4 focuses on how many errors respondents are receiving on these pages. The “Percent of Errors” column in Table 3 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 3. Screens on which Respondents Frequently Received an Error Message

<b>Page Error Rendered</b>	<b>Number of Errors Rendered</b>	<b>Number of Screen Visits</b>	<b>Percent of Errors (%)</b>	<b>Percent of Screen Visits (%)</b>
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 <sup>6</sup>	8.7

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

Changing focus to how many error messages respondents received on the pages discussed in Table 3, we see that respondents who received an error message received about 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.

<sup>4</sup> For dollar amounts, a message appeared if a respondent entered a non-numeric value in the write-in field. More critical items had additional error messages.

<sup>5</sup> 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.

<sup>6</sup> Less than 0.05 percent of errors were on the Contact (if business) screen.

Table 4. Number of Errors Received by Respondents

<b>Page Error Rendered</b>	<b>Number of Errors Rendered</b>	<b>Number of Respondents</b>	<b>Avg Number of Errors per Respondent (for respondents that received an error message)</b>	<b>Avg Number of Errors per Respondent that Accessed the Survey</b>
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, American Community Survey Internet Test, April to May 2011

The majority of the errors across all of these screens are due to leaving the question blank (52.6 - 92.7 percent), as opposed to providing an invalid entry. This means that if the error messages are working as intended, there will be a reduction in item nonresponse. Table 5 shows the percent of entries that respondents corrected after receiving an error message. Corrected does not necessarily mean that the data are fully complete or accurate, but rather a change was made that corrected the issue that triggered the error.

Table 5. Percent of Corrected Entries

<b>Screen</b>	<b>% Corrected</b>
Place of birth	91.8
Year built	91.5
Respondent name	79.7
Date of birth	74.4
Wages amount	67.3

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

The high percent of errors and corrections for place of birth and year built are likely because the question format differs from what respondents are used to seeing. For place of birth (Figure 2), rather than just needing to click a response option, there is a second task to complete; respondents must select a state from the drop down menu or write in a country. Most of the errors on these screens are due to respondents not completing the second part of the task, likely because they did not realize there was an additional step. Therefore, when they realized there was more to do, they had no problem completing the entire task. Although respondents are correcting their answers, we may want to change the format of these questions to avoid additional error messages.

Figure 2. Place of Birth Screenshot

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

Correction rates drop for the remaining questions. Both respondent name and the amount received in wages can be sensitive questions for respondents and they simply may not want to provide this information. However, the correction rate is still relatively high considering the sensitivity of the items. It is possible that the accuracy of these corrections is low, but this hypothesis cannot be easily verified.

Finally, respondents to the ACS need to report for all members of their household. Many households contain unrelated people and answering questions about wages and date of birth may be difficult for many people in this type of living situation. This provides another reason, in addition to question sensitivity, as to why the correction rate is lower for these two questions. It could be that after receiving the error, some respondents try to guess, possibly because they do not know they can continue through the survey without providing a response.

While it is promising that the messages are prompting respondents to correct their errors, it is important to make sure the error messages do not have unintended negative consequences. Specifically, Mooney, Rogers, and Trunzo (2003) found that error prompts, while decreasing item nonresponse, may be increasing breakoff rates. To assess the impact of error messages on breakoffs, we measured the percentage of breakoffs that occurred immediately after receiving an error message as well as the average number of error messages rendered for respondents who broke off compared to respondents who completed the survey.

Of all the breakoffs, 3.6 percent occurred immediately after receiving an error message<sup>7</sup>. The page with one of the highest breakoffs immediately following an error message was respondent name (28.8 percent of error messages resulted in a breakoff). This is especially problematic since this screen appears so early in the instrument so essentially no data is captured for these households. The other pages with some of the highest frequencies of breakoffs following an error message included wages amount (28.0 percent) and date of birth (24.0 percent). These breakoffs could be a result of not wanting

<sup>7</sup> This percentage is likely an undercount due to respondents selecting other links on the page before leaving the survey, which is difficult to capture.



to provide sensitive information, especially if respondents are not aware they can continue without providing an answer.

In addition to immediate breakoffs, there is also concern that there is a cumulative effect of error messages. That is, the more errors a respondent receives as they move through the instrument, the more likely they are to break off. Respondents who completed the survey received 1.2 error messages on average, as compared to 1.5 errors for those that broke off, which is significantly different at the 0.1 level with a p-value of less than 0.0001. While there is a tradeoff between reduced item nonresponse and breakoffs, the error messages are resulting in much more data for only a slight increase in breakoffs, suggesting they are doing what we intended them to do.

### 2.3 Use of Help

Past research has shown respondents do not frequently use available help features in Internet surveys (Conrad *et al.*, 2006, Lind *et al.*, 2001). The research has provided many variations on the location and format of help as well as variations on how to activate the help. For the April 2011 Internet Test, we placed the help link immediately following the question stem, where we felt it would be easy to find (see Figure 2 for the location and format of the 'Help' link).

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

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

Contrary to much of the research, the Case Help Rate was relatively high at 40.0 percent, meaning that 40 percent of respondents selected Help at least once. This 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 instrument, but rather related to the questions themselves.

To determine how useful the help was to respondents, we examined the percentage of cases in which a respondent selected an answer, then 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 an answer change (i.e., if respondents clicked Help and then another link, such as Instructions, before changing their answer, they are not included in this analysis). In 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 either respondents were using the information to verify the selection they already made or the Help 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 hypotheses so additional cognitive testing is needed to uncover respondents' intentions.

## **2.4 General Observations**

In addition to identifying potentially problematic screens, we were also interested in how respondents interacted with the instrument. Specifically, we looked at the devices used to access the survey, the use of the Review & Edit feature, and completion behavior including the average number of sessions needed to complete the survey and pages respondents frequently left the survey by logging out or timing out.

### **2.4.1 Devices**

With smartphones and tablets becoming more popular, it is interesting to see what devices respondents used to access the ACS Internet instrument. While there was no mobile version of the survey, tracking these data can help inform whether it is important to create mobile-friendly versions of the instrument for future use. For this survey, mobile device users had to work with an instrument designed to fit on a standard monitor. We found that 97.3 percent of respondents did use a personal computer to complete the survey. Of the 2.7 percent of respondents that used a mobile device, 34.6 percent used an iPad (0.8 percent of respondents). With more tablets coming onto the market, this number should be increasing over time. While preliminary testing by Census Bureau 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. Despite not having a mobile version of the survey, the breakoff rate for mobile device users was consistent with the personal computer breakoff rate (about 15 percent). Past research has shown an average breakoff rate of 24 percent for mobile devices, so these findings are encouraging, especially since there was no mobile version available (Callegaro, 2010).

### **2.4.2 Review & Edit**

Once respondents completed the ACS questions, they were given the option to submit the survey or to review their responses ('Presummary' screen). When they chose to review, they were first presented with a 'Summary' screen, which let them select exactly what data they would like to review (the housing questions or the person information for each person in the household individually). They could either select the housing data or they could click on a household member's name and be taken to that person's data. On the detailed review screens, respondents could click on an answer they wanted to change and be taken to that question. After making a correction, they were redirected back to the detailed review screen.

In total, 21.5 percent of the respondents who completed the survey chose to review their answers, which brought them to the Summary screen (Figure 3). Of the people who reached the Summary screen, the vast majority opted to review the detailed responses for either the housing questions, person information, or both. Specifically, Table 6 shows what data respondents chose to review and whether they made changes. The "Percent of reviews" column represents the proportion of people who reviewed the housing and person sections of all the respondents who selected 'Review'.

Figure 3. Summary Screenshot

USCENSUSBUREAU  
Helping You Make Informed Decisions

**AMERICAN COMMUNITY SURVEY**

Instructions      FAQs      Save & Logout

## Summary

➡ Please click [Review Answers](#) below to view a list of your answers. Then click **SUBMIT** to complete the American Community Survey.

Housing Questions	<a href="#">Review Answers</a>
<b>Person Info</b>	
• John T Doe	<a href="#">Review Answers</a>
• Jane N Doe	<a href="#">Review Answers</a>
• Tommy C Doe	<a href="#">Review Answers</a>

**SUBMIT**

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

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

Table 6. Percent of Cases 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, American Community Survey Internet Test, April to May 2011

Table 6 shows that many more respondents reviewed the person data as compared to the housing data. This was expected as the person sections contain much more information than the housing section, so there is more room for error or needing to follow up with other household members. These data are very encouraging in that respondents that visit the Review & Edit section are using it to not only review, but also change their data. Unfortunately, we cannot say if respondents are changing answers or answering questions they had previously skipped. However, this is an area of future research to help understand how people are going through the instrument and using the Review & Edit feature.

It should also be noted that just because respondents are updating their responses, this is not always a good thing. A handful of cases were identified where respondents used the Review & Edit to delete prior responses, such as the amount received in wages and date of birth.

Another concern was whether to provide a 'Submit' button on the detailed review pages or force respondents to return to the 'Summary' screen to submit the survey. The 'Submit' button lets us know the respondent is completely finished with the survey. Once respondents select this link, they cannot return to the survey. If respondents were forced to return to the 'Summary' screen they might just close out of the survey either because they could not find the 'Submit' button or because it was not easily available to them. However, 97.7 percent of the people who reviewed their answers submitted the survey.

Therefore, making people return to the ‘Summary’ screen does not appear to be an issue.

### 2.4.3 Completion Behavior

Of 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 leaving 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 7 shows the four pages where the greatest proportion of logouts 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. Table 7 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 7. Pages on which Respondents Logged Out

<b>Logout Page</b>	<b>Total Number of Logouts</b>	<b>Percent of All Logouts (%)</b>	<b>Percent of Respondents (%)</b>
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, American Community Survey 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. 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 break off out of frustration, but rather intended to return and forgot. For example, as discussed in Section 2.1, 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. 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.

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 needed 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 choose to continue with the survey. Of the 86.3 percent that logged out of the survey, 82.2 percent eventually returned. While the majority of respondents chose to log out, the 13.7 percent continue rate 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 have<sup>8</sup>.

### **3. Discussion and Conclusions**

Paradata collected from the April 2011 ACS Internet Test provide a unique and exciting new method of evaluating the survey instrument in the field. While the instrument went through a series of lab tests, the number of participants and their experiences were limited, so only some problems were uncovered. The paradata, on the other hand, allow us to understand how the actual survey population interacts with the instrument. By analyzing the paradata, we were able to uncover certain problems with the instrument and also better understand how people navigate through the survey.

We discovered that respondents broke off more frequently on transitional screens and questions that asked sensitive information. Tancreto and her colleagues (2012) found that breakoffs increase item nonresponse and therefore hurt data quality because we are unable to collect most of the survey information for these people. Therefore, knowing on which screens respondents leave provides the opportunity to alter the transitional screens to help keep respondents from leaving. Additionally, by looking at error messages, we found that the format of the Place of Birth and Year Built questions was problematic because respondents did not notice they had two tasks to complete on each screen.

In addition to helping identify problematic questions, error messages helped reduce item nonresponse. Overall, the error messages led to more complete data than would have been obtained without the messages, although we do not know the quality of the data provided after receiving an error message.

While paradata can assist in identifying problematic questions, it can also inform researchers as to how respondents move through the instrument and interact with it. We learned that respondents did not have difficulty locating the 'Help' link and generally used the help to formulate a response, rather than to change a response. Additionally, most respondents completed the survey in one session. For respondents who completed the survey in multiple sessions, we can begin to understand their thought process by comparing pages on which respondents timed out or logged off to the pages on which

---

<sup>8</sup> The logout and leave rates are calculated by counting the number of people who saw the welcome back screen. However, due to complexities with the data, some respondents did log out and return, but did not receive the welcome back screen. Therefore, the rates are slightly underestimated as they do not count these respondents.

they broke off. The similarities suggest that in some cases respondents did not break off because they were frustrated, but rather took a break and forgot to come back. Finally, we saw a small, but likely growing proportion of respondents used a mobile device to complete the survey, which suggests it may be a good idea to develop a mobile-friendly version of the survey.

Together, the information provided from the paradata offers a complete picture of how the ACS Internet Test instrument performed. Some of the results will inform future changes to improve the instrument, while other results indicate certain aspects are working well. By continuing to monitor paradata in future Internet tests and in ACS production, we can continue to develop the instrument to make it as easy as possible for respondents to use.

### 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 Ruiter, Andrew Roberts, Brian Wilson, Tony Tersine, John Studds, Chris Butler, Joe Mistichelli, Brian Ridgeway, Michael Coan, Fred Meier, and Samantha Barron.

### References

- Callegaro, M. (2010). "Do you Know Which Device your Respondent has Used to take your Online Survey?" *Survey Practice*, December.
- 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. (2000). "Usability evaluation of computer-assisted survey instruments." *Social Science Computer Review* 18(4): 384-96.
- Couper, M., Traugott, M., & Lamias, M. (2001). "Web Survey Design and Administration." *The Public Opinion Quarterly* 65:230-53.
- Dillman, D. A., & Bowker, D. K. (2000). "The Web questionnaire challenge to survey methodologists" [Online]. Available: <http://survey.sesrc.wsu.edu/dillman/papers.htm>
- Galesic, M., & 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.
- Groves, R.M., & Kahn, R.L. (1979). *Surveys by Telephone*. New York: Academic Press.
- Heerwegh, D. (2003). "Explaining Response Latencies and Changing Answers Using Client-Side Paradata From a Web Survey." *Social Science Computer Review*. 21: 360-373.
- Heerwegh, D. (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.

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. Mannheim, Germany, October 2002.

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, C., Tourangeau, R., Couper, M., Conrad, F., & Ye, C. (2009). "The Effects of Grouping Response Options in Factual Questions with Many Options." Federal Committee on Statistical Methodology (FCSM).

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](http://www.census.gov/acs/www/Downloads/library/2012/2012_Tancreto_01.pdf)