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

April 23, 2015

2015 AMERICAN COMMUNITY SURVEY RESEARCH AND EVALUATION REPORT MEMORANDUM SERIES
#ACS15-RER-04

DSSD 2015 AMERICAN COMMUNITY SURVEY MEMORANDUM SERIES
#ACS15-MP-02

MEMORANDUM FOR ACS Research and Evaluation Steering Committee

From: Patrick J. Cantwell ***/Signed/***
Chief, Decennial Statistical Studies Division

Prepared by: Rachel Horwitz
Census Experiments Branch
Decennial Statistical Studies Division

Subject: Usability of the American Community Survey Internet Instrument on Mobile Devices

Attached is the final American Community Survey Research and Evaluation report "Usability of the American Community Survey Internet Instrument on Mobile Devices." Since the Internet instrument was first tested in 2011, an increasing percentage of respondents are using mobile devices to respond to the survey. The object of this research was to determine whether using a mobile device to complete a survey designed to be completed on a desktop or laptop computer had a negative impact on data quality and increased respondent burden. It also examined the demographic characteristics of mobile respondents. The content of this report was presented at the 2014 Methodology Symposium in Ottawa, Canada.

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

Attachment

cc:
ACS Research and Evaluation Work Group

Todd Hughes	(ACSO)
Stephanie Baumgardner	
Agnes Kee	
Dave Raglin	
Dameka Reese	
Andrew Roberts	
Jennifer Tancreto	
Brian Wilson	
Mary Davis	(CSM)
Erica Olmsted Hawala	
Temika Holland	
Beth Nichols	
Victor Quach	
Nancy Bates	(ADRM)
Tony Tersine	(DSSD)
Michael Bentley	
Steven Hefter	
Eli Poehler	
Mary Frances Zelenak	

American Community Survey Research and Evaluation Program
April 16, 2015

Usability of the American Community Survey Internet Instrument on Mobile Devices

FINAL REPORT

Rachel Horwitz
Decennial Statistical Studies Division

Intentionally Blank

Abstract
The American Community Survey (ACS) added an Internet data collection mode as part of a sequential mode design in 2013. The ACS currently uses a single web application for all Internet respondents, regardless of whether they respond on a personal computer or on a mobile device. However, as market penetration of mobile devices increases, more survey respondents are using tablets and smartphones to take surveys that are designed for personal computers. Using mobile devices to complete these surveys may be more difficult for respondents due to longer load times, small font sizes, using a finger to select the proper response option, and increased scrolling. These difficulties may translate to reduced data quality if respondents become frustrated or cannot navigate around the issues.

The ACS provides a unique opportunity to measure the impact of answering survey questions on a mobile device across a national probability sample. Specifically, this study uses breakoffs, completion time, how often respondents switch to a different device, average number of changed answers, and average number of error messages rendered to compare data quality indicators across computers, tablets, and smartphones. Using a large, national sample also allows us to explore which demographic groups use mobile devices to answer the survey. Some of the traditionally hard-to-interview groups have higher mobile device penetration. If a survey focuses on these populations, it may be even more important to ensure the survey has high usability on all devices.

I. Introduction

Only a few years ago, optimizing websites and surveys for mobile instruments was not at the forefront of survey designers' minds. In May of 2011, only 35 percent of U.S. residents owned a smartphone and eight percent owned a tablet (Pew Research Internet Project, 2014). However, two years later, 56 percent of residents owned a smartphone and 34 percent owned a tablet. The majority of these owners is younger and has a higher income than those who do not own such devices (Pew Research Internet Project, 2014). Originally, there were only a handful of options for mobile devices; the vast majority being Apple or Android. However, as consumer interest has increased, more devices are coming onto the market at more affordable rates, resulting in increased market penetration.

With more users browsing the Internet on mobile devices, Web designers have developed versions of their webpages that are optimized for viewing on mobile devices. This change is necessary because loading standard websites designed for personal computers (PCs) on a mobile device can be frustrating to users. Additionally, there are other potential issues with viewing standard websites on mobile devices if there is not enough space on the screen to display all of the information legibly. This can lead to increased scrolling and zooming to be able to see the information clearly.

Based on usability issues with viewing standard websites on mobile devices, both web designers and survey administrators have created optimized sites that limit features to make them easier to view and quicker to load on mobile devices (Johansson, 2013). To measure the impact of optimizing a survey to be taken on a mobile device, Baker-Prewitt (2013) randomly assigned respondents to complete a survey on a computer, tablet, smartphone, or optimized smartphone. She found less straightlining (selecting the same response option for all items in a table) and fewer breakoffs on the smartphone using an optimized site than the non-optimized. Similarly, Wells and his colleagues (2013) found fewer breakoffs and shorter response times when respondents used a mobile-optimized instrument. However, in both of these studies, the optimized instrument did not perform as well as a standard computer on the measures of interest.

While these studies provide some insight into the effects of not optimizing survey instruments for mobile devices and the limitations of doing so, the current literature focuses on small, relatively short web surveys that use specific respondent populations, such as online panels or college students. No research, to date, has used a random national sample from a multi-mode household survey to assess the response and usability differences between devices. This limits our knowledge of which demographic groups use which devices to respond to surveys and whether those differences could be contributing to usability issues and overall data quality. Additionally, prior research has either focused on iPad tablets or not specified what types of devices they are including in their analyses. Therefore, this paper aims to compare data quality and respondent burden associated with using iPads, other tablets, and smartphones to respond to the American Community Survey. The specific research questions this study answers are:

1) Do mobile device respondents take longer to complete the ACS than computer respondents?
2) Do mobile device respondents breakoff at a higher rate than computer respondents?
3) Do respondents that start the survey on a mobile device switch to a computer?
4) Do mobile device respondents change their answers more frequently than computer respondents?
5) Do mobile device respondents render more error messages than computer respondents?
6) Do the demographic characteristics of mobile device respondents differ from those of computer respondents?

These research questions will either help us evaluate respondent burden, data quality, or both. Specifically, increased completion times has been related to both respondent burden and lower data quality (Crawford *et al.*, 2001). An increased number breakoffs can lead to higher rates of missing data, which leads to lower overall data quality (Tancreto *et al.*, 2012b). Respondents that switch from a mobile device to a computer may do so because they are experiencing higher burden, but logging out of the survey increases the chances the respondent will forget to return, resulting in breakoffs and item nonresponse. Finally, increased answer changes and error messages rendered can lead to frustration with the instrument. Additionally, if respondents are not able to select the correct response option or do not realize they miss-selected, data quality could suffer. If we do find significant differences in data quality and burden for mobile respondents, we will show a need for a mobile-optimized ACS instrument, especially if mobile respondents differ demographically from computer respondents.

II. Literature Review

Although "smartphones" have been available since the late 1990s, the first modern, touchscreen smartphone was the iPhone, which was released in 2007, while the first touchscreen tablet, the iPad, was released in 2010. The success of Apple's mobile devices led other manufacturers to develop their own touchscreen models and the use of these products has continued to increase as more devices come onto the market and prices become more affordable.

Prior to the introduction of these devices, web designers created websites to be viewed on a standard computer or laptop that directly connects to an Internet source (ranging from dial-up to a cable modem or fiber optic connection). However, with increased mobile device penetration, users can now access the Internet almost anywhere. Many of these places do not have fast connection speeds such as WiFi or 4G, which can make loading websites, especially those with graphics, very slow and frustrating. Additionally, Johansson (2013) found potential usability issues that come with smaller screens. Specifically, users often need to scroll both horizontally and vertically more than they typically would. He noted that due to how easy it is to scroll on mobile devices, users can easily lose their place and scroll past what they are looking for. In cases where the links are difficult to select or navigate, zooming can make this easier, but the user then loses the context surrounding their focus. This also adds burden due to the additional steps needed to view the information.

Many of the same usability issues with viewing webpages on mobile devices also exist in taking surveys on mobile devices. Although surveys typically do not include many graphics, load times can still be quite slow compared to computers. Additionally, users may need to scroll and zoom more than they would on a computer, which could increase response time, especially on smartphones. Specifically, McClain and her colleagues (2012) found that college student respondents using a mobile device took four minutes longer to complete a 28-minute survey than did students responding on a computer while Mavletova (2013) found that mobile respondents took three times as long to answer than computer users did. In a Dutch panel survey, Bruijne and Wijnant (2013) found that both the perceived time spent on a survey and the actual time spent were longer for mobile devices. This increase in response time may lead respondents to break off, especially for longer surveys. However, this could also be an issue for shorter surveys if perceived response time is high.

While increased response time may frustrate respondents and lead to breakoffs, other usability issues may directly affect data quality. To complete surveys on mobile devices, respondents need to read small font, scroll, zoom, and select small radio and navigation buttons with their fingers. These actions can decrease data quality so much that Callegaro (2013) has suggested researchers at a minimum flag mobile device cases, but also consider eliminating them from the analysis or directing the respondents to another device. Either due to long response times, difficulty using a mobile device to complete a survey, or a combination of factors, many studies have found higher breakoffs on mobile devices than computers (Baker-Prewitt 2013; Callegaro 2013; Mavletova 2013; Wells et al., 2013; Guidry, 2012). In all of these studies, breakoff rates for smartphones were higher than those for tablets and computers while some researchers did not find significant differences in breakoff rates between tablet users and computer users (Wells et al., 2013; Guidry 2012; Baker-Prewitt 2013). Although it is promising that there may not be a difference in breakoff rates between tablet and computer users, it is not clear which tablets respondents are using in these studies. Wells et al. (2013) and Guidry (2012) both specify their tablet data only come from iPads. However, other researchers do not specify the different types of devices used. This may be important because iPad tablets have a different operating system and software than non-Apple products, which could result in a different user experience.

For respondents who do complete the survey, researchers have found more straightlining (McClain et al., 2012; Baker-Prewitt 2013), fewer write-ins (Mavletova, 2013; Maxl, 2013), and less text input in write-ins (Peytchev and Hill 2010) for mobile device respondents (especially smartphones) as compared to computer respondents. Additionally, Peytchev and Hill (2010) found that some mobile device respondents did not scroll horizontally, so they did not see response options or question text that was to the right of what was visible on the screen, and that the perceived difficulty of responding on a mobile device was greater than responding on a computer.

The findings from these studies suggest that responding to surveys on a smartphone could lead to a reduction in data quality, while there is a minimal difference in data quality between tablet and computer respondents. However, many of these studies (Wells et al., 2013; Peytchev and Hill 2010; Baker-Prewitt 2013) were designed with web respondents in mind: they are short in overall length, use short questions with short response lists, do not include grid items, and require minimal scrolling across all devices. Additionally, the studies typically either use nonprobability panel samples or sample a specific population (Baker-Prewitt 2013; Mavletova 2013; Peytchev and Hill 2010; Wells et al., 2013; Guidry, 2012; Maxl, 2013). Therefore, this study aims to use a national, probability sample to compare data quality indicators and respondent burden on four different devices (smartphone, iPad, other tablet, and computer). The results of this analysis will help researchers determine whether they need to take action to either improve data quality or handle mobile device respondents and data differently.

III. American Community Survey Background

Prior to the January 2013 panel, the ACS collected data using three response modes: paper questionnaire, computer-assisted telephone interview (CATI), and computer-assisted personal interview (CAPI). In the 2013 January panel, ACS production added a new Internet data collection mode. Under the current design, respondents first receive an invitation in the mail to complete the survey over the Internet. Those that do not respond to the initial request then receive a paper questionnaire. Households that do not respond on the Internet or by mail, and for which we are able to obtain a telephone number, are then contacted for CATI and a subset of CATI nonrespondents, unmailable addresses, and nonrespondents to the self-response modes receive a personal visit.

In 2008, the ACS began preliminary planning for an Internet reporting option. The motivation for this addition came from several government mandates (the 1998 Government Paperwork Reduction Act of 1998 and the E-GOV Act of 2002), increases in efficiency, and reduced costs (reduced printing, postage, and data capture). Before the Internet option was included in production, the ACS conducted two tests in 2011 to determine the best way to notify sampled households of the Internet option, measure usability of the instrument, and measure the data quality of the Internet responses compared to the mail responses, which were found to be equivalent (Horwitz et al., 2013a). The Internet instrument's design is consistent with the paper form, but uses some features similar to the CATI instrument to make the survey more user-friendly (Tancreto et al., 2012a). Many of the questions on the ACS are long, complex, or have long lists of response options, all of which can be difficult to view and understand on smaller screens. These aspects of the questions cannot be modified to accommodate a smaller screen because monthly data are aggregated to create annual estimates and changes could lead to a break in series. Additionally, such changes could alter the stimuli compared to other modes. There is no secondary instrument that is optimized for mobile devices because when the instrument was originally designed, there was very low mobile device usage. However, in the April 2011 Internet test, approximately 2.2 percent of respondents used a mobile device while only seven months later in the November 2011 test approximately 4.5 percent of respondents used a mobile device (Horwitz et al., 2013b). Of these mobile device users, most used a tablet to complete the survey (80.0 percent).

IV. Methods

ACS Data

The ACS collects data in 12 panels throughout the year. Each panel is comprised of data collected throughout a three-month period where data in the first month are collected through self-response modes (Internet and paper questionnaires), data in the second month are collected through the self-response modes and CATI, and data in the third month are collected through the self-response modes and CAPI. The data used in this analysis come from Internet responses throughout the complete November 2013, December 2013, and January 2014 ACS production panels, which consists of data collected between October 2013 and March of 2014. In total, there were 227,151 Internet respondents in these three panels. All estimates in this report use base weights that reflect each household's probability of selection into the sample.

Devices Used

Along with response data, we also used paradata to determine which devices respondents used to access and complete the survey. Specifically, we used the user-agent string corresponding to each individual login (Appendix A provides example user-agent strings for each device). Although these strings do not explicitly identify the type of device used, we were able to pull out key words, such as "mobile," "iPad," and "galaxy" to begin to subset mobile devices from computers. Using these key words, we then used a user-agent string translator, found online[1], to identify which combinations of key words corresponded to various phones and tablets. When the translator did not clarify the type of device, we looked up the specific device to determine into which category it fell.

Once we had identified all of the devices used to complete the survey, we grouped them into four categories: computer, tablet, iPad, and phone. We separated tablets and iPads because both studies that specified which devices they used only included iPads (Wells et al., 2013; Guidry, 2012). The iPad's operating system, software, and screen size may influence respondent behavior and data quality compared to other tablets[2]. We counted hybrid laptops/tablets as computers because their external keyboards and trackpads make the responding experience more similar to using a computer.

Analysis

We first calculated the average size of households that used the different devices to complete the survey. We include all households for which we know the household size in this analysis. If households that use a computer to respond to the survey are typically smaller than households that used a mobile device, then it follows that it would take them less time to complete the survey and they would answer fewer questions. Therefore, in order to compare results across devices, we needed to ensure the households are similar. If household size does vary across devices, we will then compare the total number of questions answered, on average, on each device.

We next compared the breakoff rate for the different devices. The numerator, breakoffs, includes any respondent that did not click the submit link or reach the last applicable question in

[1] For this analysis, we used the following translator: http://user-agent-string.info/.

[2] Although iPhones also use a different operating system and software than other smartphones, we were unable to similarly separate them because of the overall low incidence rate of smartphones.

the survey, while the denominator includes anyone that accessed the Internet instrument and saw at least the first question. Some respondents started the survey on a mobile device and then switched to a computer at some point during the survey. The structure of our data does not allow us to determine which questions were answered on which device, so any respondent that switched from a mobile device to a computer was removed from the breakoff analysis.

To determine which respondents switched devices, we compared the user-agent strings at each login point. This analysis served two purposes: the first, to identify the correct population for analyzing breakoffs and answer changes, and second, to determine whether respondents could not or did not want to complete the survey on a mobile device and needed to switch to a computer, likely because it was easier. The first switching measure we calculated was the percent of respondents that switched from each mobile device to a computer. Here we compared the number of people that first logged in using each mobile device and subsequently logged in using a computer to the total number of respondents that logged in using each mobile device. This measure indicates whether respondents found any particular device more difficult to use to respond than the others.

Another measure we used is the number of multiple logins for each device, which can be attributed to switching. Specifically, for each device, the numerator is the number of switches from a mobile device to a computer, while the denominator is the total number of multiple logins. We want as many respondents as possible to respond in one session because this reduces the number of people that leave and do not return. Therefore, if many of the multiple logins are due to switching devices, we can assume that even more respondents abandoned the mobile device and did not switch to a computer but rather broke off.

To this point, our measures have focused on people who either abandoned the survey entirely or abandoned their mobile device to complete the survey on a computer. However, completed cases can also tell us about respondent burden and data quality. Therefore, we compared the average completion time across the four devices. We focus only on respondents who submitted the survey and answered all of the questions in one session. Inexplicable outliers (respondents that took longer than 70 hours to complete the survey) were removed from the analysis. These outliers likely arise because the paradata occasionally miss an event, such as logging out or logging back in. Due to the long response times, we assumed these respondents did not actually complete the survey in only one session.

The final two measures of burden and data quality focus more on individual responses than the overall survey. First, we look at rate of changed answers, which was calculated by comparing the total number of changed answers on each device to the total number of respondents that completed the survey on each device. Respondents can change their answers for a variety of reasons, but in the case of mobile devices, it is likely that more of these changes are a result of difficulty touching the smaller radio buttons or check boxes. This can lead to decreased data quality if mobile respondents do not realize they selected the wrong response option and it increases burden because they need to answer the question multiple times. To determine the rate of changed answers, we only include immediate changes[3]. We calculate the changed answer rate twice, once including only completed cases (submitted or answered all

[3] If a respondent selects a response option and then their next action is to select a different response option, we count it as a change. However, if a respondent selected a response option, then selected the Help link, and then changed their answer, it was not counted. We also eliminated respondents who have more than 40 consecutive changes because we found irregular patterns for these respondents that did not reflect true answer changes.

applicable questions) to control the total number of questions respondents of each device saw[4] and once including only breakoffs. This can help inform whether frustration due to changing answers may have led to the breakoffs.

Next, we look at the error rate for each device, calculated by comparing the total number of errors rendered on each device to the total number of respondents that completed the survey on each device. Again, we calculate the rate once including only completed cases and once using breakoff cases.

Finally, we look at the demographic characteristics of mobile respondents compared to computer respondents. Given Callegaro's (2013) recommendation that mobile cases be flagged or not included in analyses, we need to know whether mobile respondents are similar to other respondents. We compare respondent age, education, race, Hispanic origin, household income and whether the home is rented across the four devices. To compare age and household income, we use the median as the basis of comparison, using a Wilcoxon score to test the difference between the income distributions of respondents using each device. For education, we look at the percent of respondents that have less than a high school education. Finally, to measure race, we look at the proportion of respondents who are Black; and for Hispanic origin, we look at the percent of respondents who are of Hispanic descent across all of the devices.

All comparisons in this report use Proc GLM in SAS to account for multiple comparisons and the inclusion of the ACS base weights. The GLM procedure provides an F-statistic that measures whether there is any difference in the variable of interest across the different devices. Additionally, it provides t-tests that can be used to compare differences between two devices once it has been established that there is a difference between the devices overall.

V. Results

In the November 2013, December 2013, and January 2014 ACS data collection panels, 85.9 (0.08)[5] percent of Internet respondents used a computer, 7.6 (0.06) percent used an iPad, 3.9 (0.05) used another type of tablet, and 2.6 (0.04) percent used a mobile phone to access the survey[6]. This compares to a total of 2.2 percent of respondents using any mobile device to access the survey in the April 2011 ACS Internet Test (Horwitz et al., 2013a) and 4.5 percent in the November 2011 ACS Internet Test[7] (Horwitz, et al., 2013b), and 11.3 (0.04) percent in the January 2013 ACS data collection panel (Horwitz, 2014). Although we cannot statistically compare the mobile usage rates across this complete time period, there appears to be an upward trend in the percent of respondents using mobile devices to respond to the survey. However, there was a significant increase in the percent of mobile respondents between January 2013 and November 2013 through January 2014 ($t=39.3$, $p=<0.0000$).

Table 1 provides the average household sizes of respondents that used each device to access the survey.

---

[4] The total number of questions each respondent saw varies depending on the age and situation of the household members. However, if mobile respondents break off more frequently than computer respondents do, they will necessarily answer fewer questions, which can affect the results of this analysis.

[5] The parentheses contain the standard error of the estimate.

[6] The device data do not account for respondents using multiple devices to complete the survey. For these estimates, we count the smallest device used by each respondent (phone, tablet, iPad, computer).

[7] The percent of responding using mobile devices for both the April and November tests are unweighted estimates due to a problem merging datasets.

Table 1. Average Household Size by Device

Device	Average Household Size	Std Err
Phone	2.79	0.01
Tablet	2.65	0.02
iPad	2.67	0.02
Computer	2.49	0.00[8]

Source: American Community Survey Data November 2013-March 2014.

A multiple comparison of average household size shows that the household size of respondents using the different devices varies significantly (F=177.19, p<0.0001), however a t-test comparing the household size of iPad and tablet respondents was not significantly different (t=5.9, p=0.3113). The average household size for computer respondents was significantly smaller than the household size for all three mobile device respondents. This is likely because mobile device users tend to be younger (Rainie, 2012; Pew Research Internet Project, 2013) and younger individuals typically live in larger households (Nichols et al., forthcoming). Although these differences are significant, they are very small, especially given the large sample size.

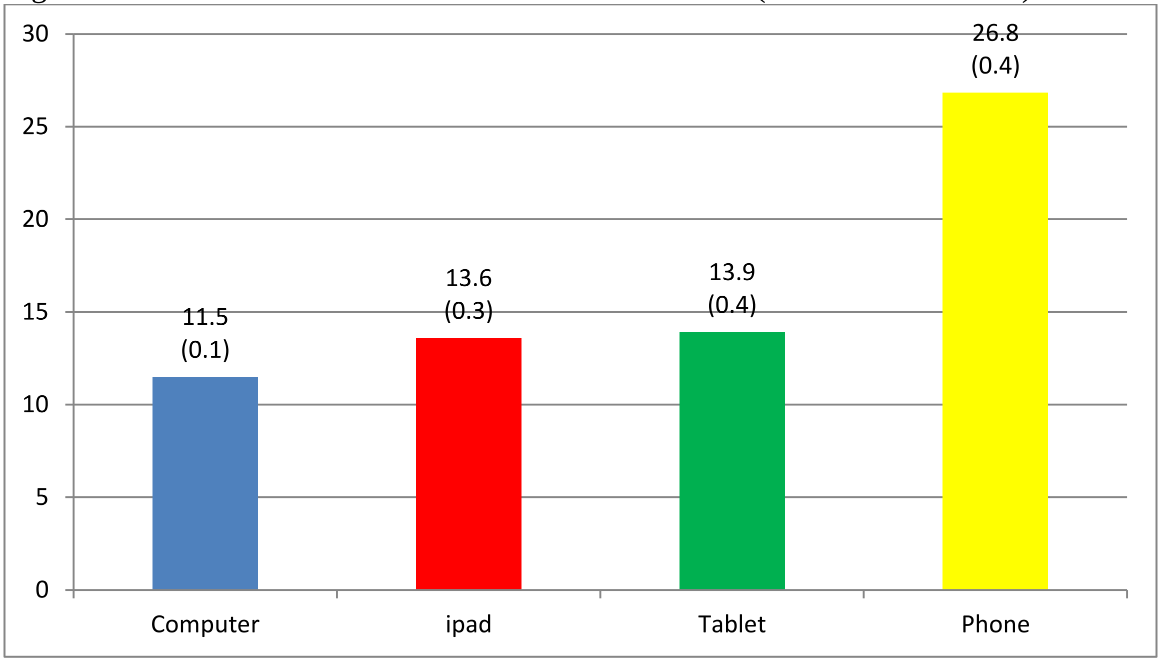
Because household size varies across devices, we look at the number of questions answered. A multiple comparison of means shows that there are differences in the number of questions answered across devices (F=18.92, p<0.001), with phone respondents answering the fewest questions (roughly 8.5 questions fewer than tablet and iPad respondents and 5.5 fewer than computer respondents), on average, even though they have among the largest household sizes. This could be due to the composition of the household (more young children). Although household size does vary across device, we do not believe there will be an issue in comparing the different measures across all households and devices because the difference is small and the number of questions answered does not appear to be associated with household size.

Do mobile device respondents breakoff at a higher rate than computer respondents?

Traditionally, the average breakoff rate for the ACS is around 12 percent (Clark, forthcoming). If burden and data quality are consistent across different devices, we expect to see a similar breakoff rate for each device. Figure 1 provides the difference in the breakoff rate between computer respondents and the other devices.

[8] Standard error rounds to 0.0.

Figure 1. Percent of Cases that Broke off for each Device (with standard errors)



Source: American Community Survey Data November 2013-March 2014.

The figure shows that the breakoff rate for phone respondents was over 15 percentage points higher than that for computers (t=34.13, p<0.0001). Additionally, phone respondents break off approximately 13 percentage points more than both tablet and iPad respondents (t=22.81, p<0.0001 and t=25.95, p<0.0001, respectively). The breakoff rate for tablet respondents is not significantly different from the rate for iPad respondents (t=0.72, p=0.4649), but there is a small difference between iPad and tablet respondents and computer respondents. Although it makes sense that phone respondents would break off at a higher rate than other devices, due to screen size and connection speed, there may be a confounding factor that is making the difference appear larger than it actually is. A higher percentage of phone respondents live in unrelated households (17.7 percent) compared to computer respondents (10.1 percent) and research has shown that unrelated households break off more frequently than related households (Horwitz et al., 2013a). This may account for some of the difference in response rates, but the percent of iPad respondents living in unrelated households is significantly less than computer respondents (t=2.97, p<0.0029), while iPad respondents breakoff more often. Therefore, while the demographic differences between phone and computer respondents may contribute to the large difference in breakoff rates, it is unlikely that they account for all of the difference. Increased breakoff rates are concerning for data quality because questions later in the survey that the respondent did not reach will have higher missing data rates.

Do respondents that start the survey on a mobile device switch to a computer?

Some respondents that have trouble completing the survey on a mobile device may not break off without returning, but rather switch to a computer. Table 2 shows the percent of respondents that switched from each device to a computer.

Table 2. Percent of Respondents Switching from a Mobile Device to a Computer

Device	Mean	Std Err
Phone	8.54	0.41
Tablet	3.91	0.24
iPad	3.52	0.16

Source: American Community Survey Data November 2013-March 2014.

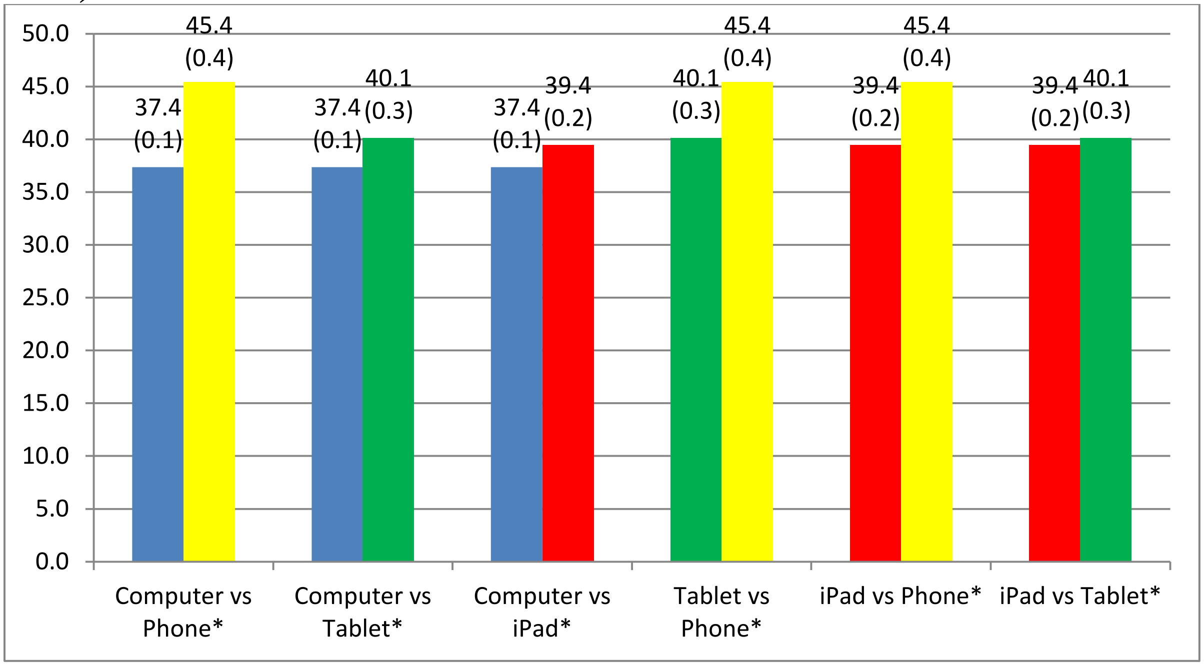
From the table, we see phone respondents switch to a computer more frequently than the other devices. Based on a logistic regression model, we estimate that phone respondents are 2.3 times more likely to abandon their phone than tablet respondents are and 2.6 times more likely than iPad respondents are (p<0.0001). Additionally, iPad respondents are less likely to switch to a computer than tablet respondents as well (p< 0.0001). This may be because respondents find the iPad interface more user-friendly, the browser is better optimized for non-mobile websites, or something related to the characteristics of iPad respondents. It is concerning that other mobile device users are switching to computers because not everyone has multiple devices on which they can answer questions. Therefore, some of these respondents are likely breaking off if they cannot switch.

Another way to look at the same switching phenomenon is by looking at the percent of multiple logins. Of the multiple logins for the different devices, 28.9 percent for phones were to switch to a computer, followed by 16.9 percent for tablets, and 15.9 percent for iPads. Again, while these respondents did return to the survey, many in similar situations likely did not and we want to make it as easy for respondents to complete the survey in one session as we can.

Do mobile device respondents take longer to complete the ACS than computer respondents?

For the respondents that did complete the entire survey in one session, we can measure their burden by the time it took to complete the survey. The ACS advertises that respondents can expect the survey to take approximately 40 minutes to complete. The average completion time for tablets, iPads, and computers was close to the expected 40 minutes, ranging from 37 minutes for computers to just over 40 minutes for tablets. While a multiple comparison of response times shows that the time taken to complete the survey significantly differed by device, (F=194.9, p<0.0001), the completion time for phone respondents is notably longer than the other devices. Figure 2 provides a comparison of the time to complete the survey across the four devices.

Figure 2. Comparison of Average Completion Time (in minutes) across Devices (with standard errors)



*Significant at the α = 0.10 level. Takes multiple comparisons into account.
Source: American Community Survey Data November 2013-March 2014.

The average completion time for all the mobile devices is significantly greater than the time for computers. However, the difference that really stands out is for phone respondents. Phone respondents took more than eight minutes longer to complete the survey than computer respondents ($t=20.5$, $p<0.0001$) did and almost six minutes longer than tablet respondents did ($t=11.1$, $p<0.0001$). We expected to see this type of result because smaller phone screens require more scrolling in order to see the entire question and response options, and more zooming in order to touch the correct radio button or check box. Additionally, page load times could be slower if respondents are on a mobile network instead of Wi-Fi.

Do mobile device respondents change their answers more frequently than computer respondents?

If respondents are taking longer to respond because they need to zoom in to touch the radio buttons, they might accidentally select the wrong option, which would require them to change their answer to the correct option. The majority of respondents, across all devices, changed at least one answer during the survey. However, they do not necessarily all change answers as frequently. Table 3 provides the average number of changed answers for all respondents that completed the survey. Both a multiple comparison test and t-test comparisons between all of the devices show the average number of answer changes between devices are significantly different, with p-values less than 0.0001.

Table 3. Average Number of Changed Answers of Completed Interviews by Device

Device	Avg Number of Changed Answers	se
Phone	8.56	0.06
Tablet	4.62	0.04
iPad	5.11	0.03
Computer	3.78	0.01

Source: American Community Survey Data November 2013-March 2014.

As expected given the size of the response options on smaller devices, phone respondents changed their answers significantly more than other respondents did. It should be noted, that phone respondents changed their answers more often than other respondents did, yet answered fewer questions, suggesting this is even more of an issue on such a small device. However, this problem is not isolated to phone respondents, both tablet and iPad respondents changed their answers significantly more than computer respondents did. We suspect the increase in changed answers for these respondents is because the radio buttons are small compared to an adult's finger size, making it easy to accidentally select the response option above or below the one they intended. It is surprising that iPad respondents changed their answers more frequently than other tablet respondents ($t=8.95$, $p<0.0001$), given that they have performed equivalently or better than other tablets in the indicators discussed so far. While these changes increase burden, it is even more of an issue if mobile respondents do not realize they selected the wrong answer or decide not to correct their error, which can lead to lower data quality.

We also looked at the number of changed answers for breakoffs, which followed the same pattern. This suggests that the breakoffs likely were not the result of frustration selecting an answer category.

Do mobile device respondents render more error messages than computer respondents?

Another data quality measure we examined was how often respondents rendered errors. Table 4 provides the average number of error messages rendered for respondents who completed the entire survey by device.

Table 4. Average Number Errors Rendered throughout the Survey by Device

Device	Avg Number of Errors Rendered	se
Phone	1.09	0.03
iPad	1.21	0.02
Computer	1.21	0.01
Tablet	1.33	0.02

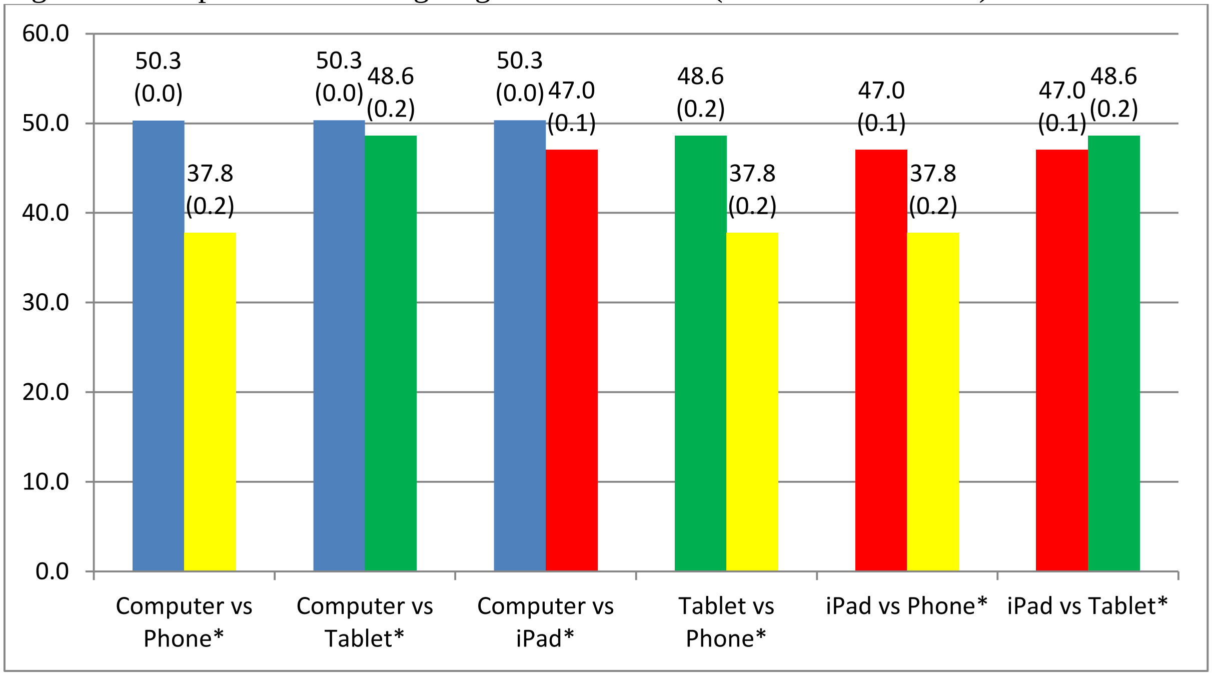
Source: American Community Survey Data November 2013-March 2014.

All of the differences in the number of errors rendered across device are relatively small. Individual t-tests comparing the average number of errors rendered show there was no difference in the number of messages rendered by iPad and computer respondents ($t=0.01$, $p=0.9886$), while tablet respondents rendered significantly more errors than iPad respondents ($t=4.09$, $p<0.0001$) and computer respondents ($t=4.93$, $p<0.0001$), and phone respondents rendered fewer errors than iPad respondents ($t=3.16$, $p=0.0016$), tablet respondents ($t=4.93$, $p<0.0001$), and computer respondents ($t=3.57$, $p=0.0004$). These findings are surprising as we expect computer respondents to have the fewest error messages because it is easier for them to read the questions and select answers. However, the number of questions answered may be contributing to this finding. Tablet respondents answered significantly more questions (148.2) than phone respondents (139.0) and computer respondents (144.6), while phone respondents answered significantly fewer questions than respondents of all the different devices. We also compared the number of errors rendered across breakoff cases and completed cases separately and found the same pattern of results.

Do the demographic characteristics of mobile device respondents differ from those of computer respondents?

In accordance with the research that mobile penetration is higher among younger people (Pew Research Internet Project, 2014; Zickuhr and Rainie, 2014), Figure 3 shows that computer respondents are significantly older than mobile device respondents.

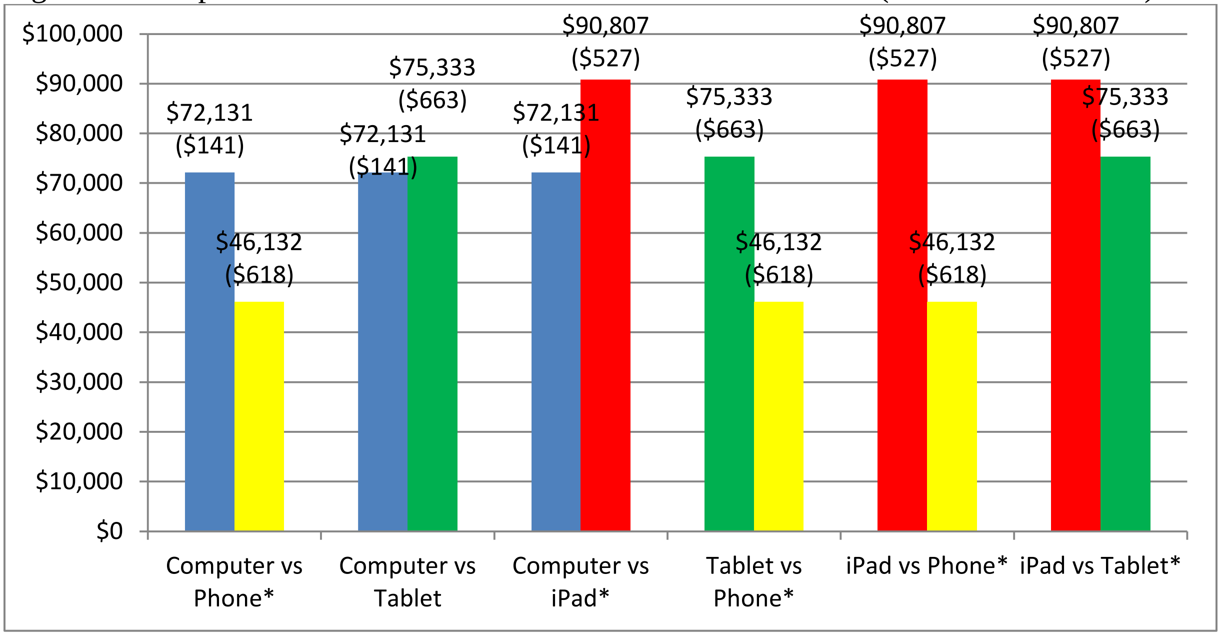
Figure 3. Comparison of Average Age across Devices (with standard errors)



*Significantly different at the α = 0.10 level. Takes multiple comparisons into account.
Source: American Community Survey Data November 2013-March 2014.

However, computer respondents are only a year or two older, on average, than tablet and iPad respondents, while phone respondents are around 10 years younger. This may be related to income, in which we see phone respondents make significantly less money than respondents of the other devices (Figure 4). Specifically, the median household income for phone respondents is $46,132, while the median income for all of the other devices is greater than $72,000. It is possible this cohort cannot afford tablets or computers like the other respondents can.

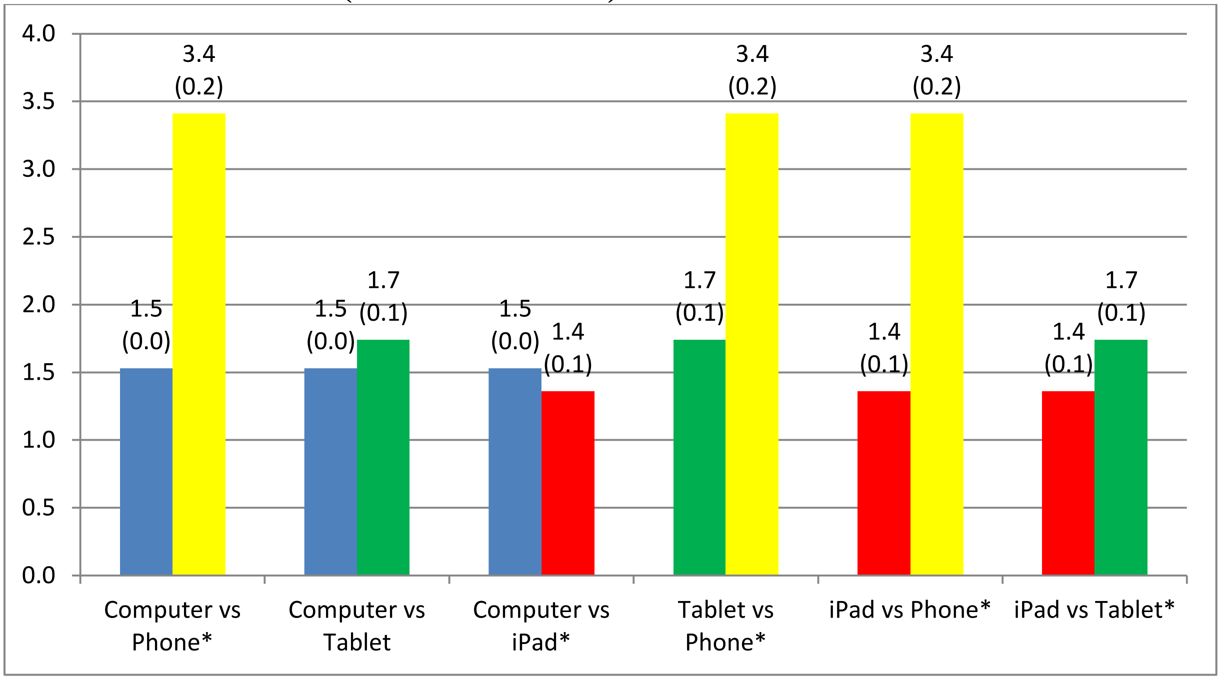
Figure 4. Comparison of Median Household Income across Device (with standard errors)



*Significantly different at the α = 0.10 level. Takes multiple comparisons into account.
Source: American Community Survey Data November 2013-March 2014.

We also find that a higher percentage of iPad respondents have completed high school than other respondents (Figure 5). Specifically, t-test comparisons of the percent of tablet respondents who did not complete high school is significantly higher than the percent of iPad respondents (t=2.26, p=0.0239) and phone respondents have more than twice as many respondents who did not complete high school as compared to iPad respondents (t=10.43, p<0.0001). Similarly, combining two reports from Pew Research Center, both using data from January 2014, 29 percent of tablet respondents had a high school education or less (Zickuhr and Rainie, 2014) , while 44 percent of smartphone users had a high school education or less (Pew Research Center, 2014). However, these results are not surprising given phone respondents' lower income and age compared to the other respondents.

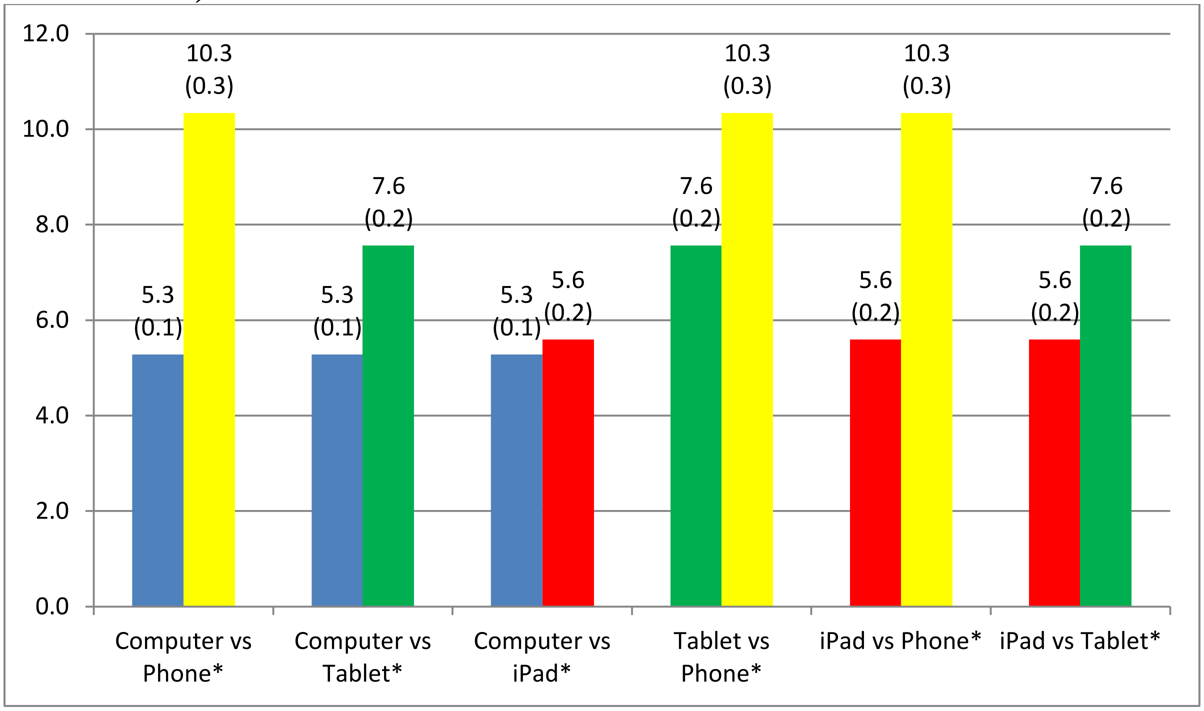
Figure 5. Comparison of the Percent of Respondents that have Less than a High School Education across Devices (with standard errors)



*Significantly different at the α = 0.10 level. Takes multiple comparisons into account.
Source: American Community Survey Data November 2013-March 2014.

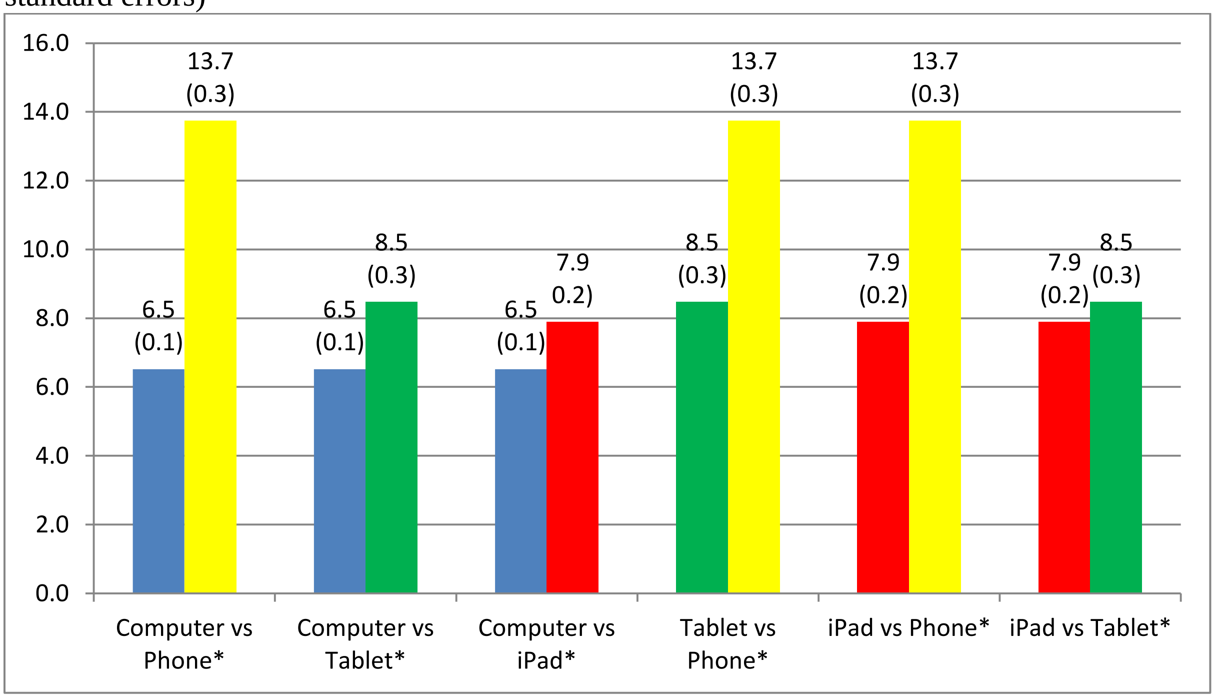
Again looking at the Pew Research Center's studies, they found that a higher proportion of smartphone users are African American or Hispanic (Pew Internet Research Project, 2014; Zickuhr and Rainie, 2014). It follows that if a higher proportion of these groups own smartphones, a higher proportion would use them to complete the survey, which is what we found (Figures 6 and 7).

Figure 6. Comparison of the Percent of Respondents that are Black across Devices (with standard errors)



*Significantly different at the α = 0.10 level. Takes multiple comparisons into account.
Source: American Community Survey Data November 2013-March 2014.

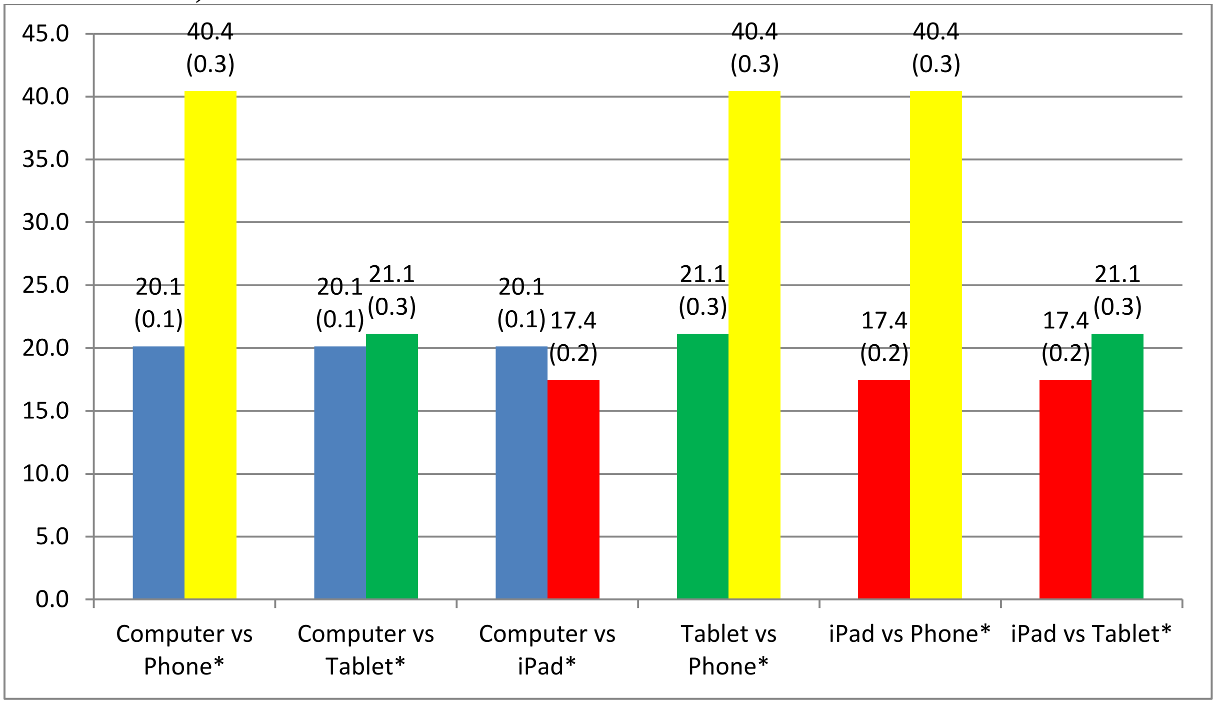
Figure 7. Comparison of the Percent of Respondents that are Hispanic across Devices (with standard errors)



*Significantly different at the α = 0.10 level. Takes multiple comparisons into account.
Source: American Community Survey Data November 2013-March 2014.

Finally, Figure 8 shows that tablet and phone respondents are more frequently renters as compared to iPad and computer respondents, with the proportion of phone respondents that are renters approximately double the other groups. Given the income and age distribution, we would not expect there to be more tablet respondents that are renters.

Figure 8. Comparison of the Percent of Respondents that are Renters across Devices (with standard errors)



*Significantly different at the α = 0.10 level. Takes multiple comparisons into account.
Source: American Community Survey Data November 2013-March 2014.

Although there are significant differences in the demographic characteristics of respondents across devices, the phone respondents stand out in particular. Many of the differences between tablet, iPad, and computer respondents, while significant, are relatively small. However, phone respondents are much more likely to be younger, less educated, a minority, a renter, and have lower income compared to respondents of the other devices.

VI. Discussion

Today's technological environment is constantly changing and it is important for surveys to keep up with the changes in order to meet respondents' expectations and maintain data quality. Research shows that market penetration of mobile devices, such as smartphones and tablets, is quickly on the rise. We have already seen increases in the use of these devices to complete surveys, so as penetration increases, it is likely that this trend will continue.

The results from the analyses conducted for this paper suggest that responding to the ACS on a mobile device likely results in higher burden and lower data quality as compared to answering on a desktop or laptop computer. Mobile users typically broke off at a higher rate, took longer to complete the survey, and changed their answers more often than computer users. However, many of the differences between computer users and iPad or tablet users, while significant, were minimal, especially considering the large sample size used in this analysis.

While the differences between iPads and other tablets and computers were not alarming in terms of burden and data quality, the results from respondents that used a phone to complete the survey were. Phone respondents broke off almost 14 percent more often than computer respondents, they took more than 8 minutes longer to complete the survey (even though computer respondents have smaller households), and had more than twice as many changed answers. They did render fewer error messages than computer respondents, but the difference was very small and they answered fewer questions.

There are several limitations to this analysis that may impact our findings. First, there was no experimental design assigning respondents to devices. Therefore, it is possible that phone respondents did not breakoff more frequently or take longer to complete the survey because they were on a phone, but rather the types of people who used a phone may exhibit this type of behavior on any device. Additionally, there may be possible confounding variables, such as household size, questions answered, and some of the demographic characteristics. For example, it is possible that phone respondents rendered fewer error messages because they have answered fewer questions, on average, than other respondents did. Similarly, it is possible that less educated respondents have more difficulty completing the survey because of their education, not because they used a phone. These confounding variables could be teased out in the future by controlling for them in analyses or by using an experimental design.

Although phone respondents appear to experience significantly more burden completing the survey and likely provide lower quality data as a result, they only make up 2.6 percent of all respondents. However, given their demographic characteristics, these respondents tend to be some of the hardest to interview, especially in the self-response modes. Joshipura (2008) identified several demographic groups that are more likely to respond in the interviewer-administered modes of data collection, resulting in higher costs to reach them. These groups include younger households, Blacks, and households with lower income and education. As we saw, a higher proportion of these same groups responded by phone than by the other devices. Additionally, given the characteristics of this group, it is possible that they do not have an alternative device to use if they are struggling with the phone. This can result not only in breakoffs from the survey, but in the worst case, they may be frustrated enough that they refuse to complete the survey in other modes as well. While the percent of respondents that use a phone to respond to the survey is small, it is important to obtain responses from these people. Not only are they demographically different from other respondents, but they will cost considerably more in interviewer-administered phases of data collection.

Given the findings discussed in this paper, we believe an optimized mobile version of the ACS would provide higher quality data at less burden to the respondents. To determine how to develop this instrument, we propose usability testing to identify exactly what issues mobile respondents, specifically phone respondents, are having with the instrument. Using this information, an optimized instrument could be developed and tested again in a lab setting to see if the problems have been reduced or eliminated after optimization. Although an optimized instrument would not solve all the problems associated with answering the ACS on a phone, especially because of the question length, it may increase the data quality of hard to interview groups responding online.

References

Baker-Prewitt, J. (2013). Mobile Research Risk: What Happens to Data Quality When Respondents Use a Mobile Device for a Survey Designed for a PC. Presented at 2013 CASRO Online Research Conference.

Bruijne, M. and Wijnant, A. (2013). Comparing Survey Results Obtained via Mobile Devices and Computers: An Experiment With a Mobile Web Survey on a Heterogeneous Group of Mobile Devices Versus a Computer-Assisted Web Survey. *Social Science Computer Review*, 31(4).

Callegaro, M. (2013). Do You Know Which Device Your Respondent Has Used to Take Your Online Survey? *Survey Practice* 3(6).

Clark, S. (Forthcoming). Internet Breakoffs. ACS Research and Evaluation.

Crawford, S.D., Couper, M.P., and Lamias, M.J. (2001). Web Surveys: Perceptions of Burden. *Social Science Computer Review* 19(2).

Guidry, K.R. (2012). Response quality and demographic characteristics of respondents using a mobile device on a web-based survey. Paper presented at the Annual Meeting of the American Association for Public Opinion Research, Orlando, FL.

Horwitz, R. (2014). Devices Used for Internet Collection. Email correspondence, May 22, 2014.

Horwitz, R., Tancreto, J.G., Zelenak, M.F., Davis, M. (2013a). Use of Paradata to Assess the Quality and Functionality of the American Community Survey Internet Instrument. Available at: http://www.census.gov/acs/www/Downloads/library/2013/2013_Horwitz_01.pdf

Horwitz, R., Tancreto, J.G., Zelenak, M.F., Davis, M. (2013b). Using Paradata to Identify Potential Issues and Trends in the American Community Survey Internet Instrument. Available at: http://www.census.gov/acs/www/Downloads/library/2013/2013_Horwitz_02.pdf

Johansson, J. (2013). How to make your Websites Faster on Mobile Devices. *Smashing Magazine*, April 3 issue.

Joshipura, M. (2008). "2005 ACS Respondent Characteristics Evaluation," U.S. Census Bureau: Decennial Statistical Studies Division. #ACS-RE-2. September 15, 2008.

Mavletova, A. (2013). Data Quality in PC and Mobile Web Surveys. *Social Science Computer Review*, 31(6).

Maxl, E., Baumgartner, B. (2013). Influence of Mobile Devices on Online Surveys. Presented at the 2013 General Online Research Conference.

McClain, C., Crawford, S.D., Dugan, J.P. (2012). Use of Mobile Devices to Access Computer-optimized Web Instruments: Implications for Respondent Behavior and Data Quality. Paper presented at the Annual Meeting of the American Association for Public Opinion Research, Orlando, FL.

Nichols, E., Horwitz, R., Tancreto, J.G. (Forthcoming). Do hard-to-interview groups self-respond more when offered an Internet reporting option?

Pew Research Internet Project (2014). Mobile Technology Fact Sheet. Released January 27, 2014. Available at: http://www.pewinternet.org/fact-sheets/mobile-technology-fact-sheet/

Peytchev, A. and Hill, C.A. (2010). Experiments in mobile Web survey design similarities to other modes and unique considerations. *Social Science Computer Review* 28(3).

Rainie, L. (2012). Tablet and E-book reader Ownership Nearly Doubled over the Holiday Gift-Giving Period. Pew Internet & American Life Project. Available at: http://libraries.pewinternet.org/2012/01/23/tablet-and-e-book-reader-ownership-nearly-double-over-the-holiday-gift-giving-period/

Tancreto, J.G., Davis, M., Zelenak, M.F. (2012a). Design of the American Community Survey Internet Instrument. Available at: http://www.census.gov/acs/www/Downloads/library/2012/2012_Tancreto_02.pdf

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

Wells, T., Bailey, J.T., Link, M.W. (2013). Filling the Void: Gaining a Better Understanding of Tablet-based Surveys. *Survey Practice*, 6(1).

Zickuhr, K. and Rainie, L. (2014). E-Reading Rises as Device Ownership Jumps. Available at: http://www.pewinternet.org/2014/01/16/tablet-and-e-reader-ownership/

Appendix A. Examples of User-agent Strings for Each Device

Computer:
Windows Personal Computer - Mozilla/5.0 (compatible; MSIE 9.0; Windows NT 6.0; Trident/5.0; SLCC1; .NET CLR 2.0.50727; .NET CLR 3.5.30729; .NET4.0C; .NET CLR 3.0.30729)

Apple Personal Computer - Mozilla/5.0 (Macintosh; Intel Mac OS X 10_7_5) AppleWebKit/536.26.17 (KHTML, like Gecko) Version/6.0.2 Safari/536.26.17

iPad:
Mozilla/5.0 (iPad; CPU OS 6_0_1 like Mac OS X) AppleWebKit/536.26 (KHTML, like Gecko) Version/6.0 Mobile/10A523 Safari/8536.25

Mozilla/5.0 (iPad; CPU OS 5_1_1 like Mac OS X) AppleWebKit/534.46 (KHTML, like Gecko) Version/5.1 Mobile/9B206 Safari/7534.48.3

Tablet:
Windows Tablet - Mozilla/4.0 (compatible; MSIE 8.0; Windows NT 5.1; Trident/4.0; Tablet PC 1.7; .NET CLR 1.0.3705; .NET CLR 1.1.4322; .NET CLR 2.0.50727; .NET CLR 3.0.4506.2152; .NET CLR 3.5.30729; BRI/2)

Android Nexus Tablet - Mozilla/5.0 (Linux; Android 4.2.1; Nexus 7 Build/JOP40D) AppleWebKit/535.19 (KHTML, like Gecko) Chrome/18.0.1025.166 Safari/535.19

Phone:
iPhone - Mozilla/5.0 (iPhone; CPU iPhone OS 6_0_1 like Mac OS X) AppleWebKit/536.26 (KHTML, like Gecko) Version/6.0 Mobile/10A523 Safari/8536.25

Huawei Android Smartphone - Mozilla/5.0 (Linux; U; Android 2.3.6; en-us; M865 Build/HuaweiM865) AppleWebKit/533.1 (KHTML, like Gecko) Version/4.0 Mobile Safari/533.1