04/10/2014

2014 AMERICAN COMMUNITY SURVEY RESEARCH AND EVALUATION REPORT
MEMORANDUM SERIES #ACS14-RER-13

MEMORANDUM FOR   ACS Research and Evaluation Advisory Group

From: James Treat **(signed 4/11/2014)**
                      Chief, American Community Survey Office

Prepared by: Stephanie Baumgardner
                      Special Studies Staff
                      American Community Survey Office

Subject: Evaluation of the use of Optical Character Recognition to capture
                      American Community Survey numeric write-ins in the 2013
                      Questionnaire Design

Attached is the final American Community Survey Research and Evaluation report for the
Evaluation of the use of Optical Character Recognition to capture American Community Survey
numeric write-ins in the 2013 Questionnaire Design Test. This evaluation examines the use of
optical character recognition (OCR) software to capture ACS numeric write-ins and reports on
the accuracy of the OCR software as well as the effect of using OCR on the quality of ACS data,
including the distribution of numeric values.

If you have any questions about this report, please contact David Raglin ((301) 763-4226) or
Stephanie Baumgardner ((301) 763-5893).

Attachment

cc:
ACS Research and Evaluation Work Group
Debbie Griffin (ACSO)   Colleen Hughes (POP)
Todd Hughes   Louisa Miller
Agnes Kee   Tori Velkoff
David Raglin   Rose Kreider (SEHSD)
Andrew Roberts   Tavia Simmons
Jennifer Tancreto (DSSD)   Patty Goerman (CSM)
Tony Tersine   Yuling Pan
Beth Tyszka

Evaluation of the use of Optical Character Recognition to capture American Community Survey numeric write-ins in the 2013 Questionnaire Design Test

FINAL REPORT

Stephanie Baumgardner
American Community Survey Office

United States Census Bureau

Table of Contents

I.   Introduction

Since 2007, the American Community Survey (ACS) has employed the Integrated Computer Assisted Data Entry (iCADE) system to capture data on paper questionnaires. This system makes use of Optical Mark Recognition (OMR) technology to detect a presence (mark) in checkboxes. Clerks at the National Processing Center (NPC) currently key all write-in responses from scanned images of the paper questionnaires. The developers of the iCADE system have now introduced Optical Character Recognition (OCR) technology to interpret numeric fields. Other Census Bureau surveys using the iCADE system have incorporated OCR in the data capture of their questionnaires with enhanced efficiencies and reduced costs. It is desirable for the ACS to consider OCR technology to reduce survey costs. Prior to any shift in methods, it is critical to demonstrate that the use of OCR to capture ACS numeric write-ins would not negatively affect the quality of ACS data, including the distribution of numeric values. This evaluation will also inform 2020 decision makers on the current usability of OCR technology in the iCADE system.

II.   Background

The 2013 Questionnaire Design Test (QDT) included a treatment to test OCR technology in the data capture of numeric write-ins on ACS paper questionnaires. The control treatment in the QDT employs the same questionnaire currently used in production (the 10.25 inch by 10.5 inch 28-page questionnaire). The OCR treatment paper questionnaire closely resembles the control treatment questionnaire with enhancements to enable the use of OCR (described below). Other treatments in the QDT tested different sizes of questionnaires (number of pages/size of paper) and new/revised content. Please see Davis and Wakim (2014) for a full description of the QDT methodology and more detail on the other treatments.

In current production, clerks in NPC key all numeric and character write-in fields from scanned images. For the OCR treatment in this test, the software electronically extracted the scanned numeric write-in fields and applied OCR technology to interpret the entries. As part of this test, clerks also independently keyed the numeric fields. This evaluation compares the values of the numeric fields resulting from keying and OCR. Numeric write-ins are associated with the items in Table 1 below. Attachment A shows where these numeric write-ins occur on the form.

Table 1. ACS Numeric Write-In Fields

Cover Page Items	Housing Items	Detailed Person Items
Month	Year built write-in	Citizenship write-in
Day	Move in month	Year of entry
Year	Move in year	Grade attend write-in
Area code	Rooms	Grade comp write-in
Phone prefix	Bedrooms	Migration zip
Phone number	Electricity amount	Year last married
Number of people	Gas amount	Place of work zip
	Water amount	Total riders
Basic Person Items	Heat fuel amount	Leave home hour
Age	Condo amount	Leave home minute
Date of birth month	Rent amount	Travel time to work
Date of birth day	Value amount	Hours worked
Date of birth year	Tax amount	Wages amount
	Insurance amount	Self-employment amount
	Mortgage amount	Interest amount
	Second mortgage amount	SS amount
	Mobile home amount	SSI amount
		Public assist amount
		Retirement amount
		Other income amount
		Total income amount

In order to test OCR technology, ACS staff changed the format of numeric write-ins on the ACS form. To aid in the OCR interpretation of numbers, separate boxes (called dentils) for each individual number replace the single segmented box in current production. See Figures 1 and 2 below for an example of the change[1]. Cognitive testing of the new format resulted in no negative comments (Terry, 2013).

---

[1] In addition to the introduction of dentils in the test, the test format also changed the color of the border of checkboxes and entry boxes from black to green. The color of the page border and vertical lines separating the three columns on the page also changed from black to green in the test format. This change is an enhancement of the form to aid in data capture by the iCADE system.

Figure 1. Current Production Numeric Field   **Figure 2. OCR Test Numeric Field**

In production, the iCADE data capture system would use OCR technology to first attempt to read all numeric entries. OCR may not read some numeric entries if the respondent's answer is undeterminable. All values read by OCR receive a confidence level of accuracy. Field values not read by OCR or below minimum acceptable confidence and any fields that fail survey-defined edits[2] would not receive an answer during OCR and would go directly to a keyer. Field values that are above the minimum acceptable confidence but below the high confidence threshold are subject to a process named OCR Review. During OCR Review, a clerk reviews individual digits captured by OCR and makes a determination on the accuracy of that digit. The clerk can either confirm that OCR captured the digit correctly or send incorrect values to a keyer.

In normal production, clerks will key only the fields not read by OCR, the OCR fields that were below minimum acceptable confidence, fields that failed survey-defined edits, and fields captured incorrectly as determined by OCR Review. We followed this basic procedure in the test environment, but to determine 'truth', clerks also keyed all numeric entries. Then, the Quality Assurance (QA) staff in NPC adjudicated differences between the OCR and keyed values. Current production procedures captured all other data on the forms designed for the OCR treatment.

III. Methodology

This report focuses on two of the QDT sample treatments, the control treatment and the OCR treatment. We mailed each treatment to about 10,000 addresses. For both the control and OCR treatments, about 24 percent of addresses responded by mail[3]. The research questions below refer only to the mail responses received in each of the treatments.

The data used to answer research questions A, B, C, and D was contained on an OCR analysis file that had the following values for each numeric field:
- the pre-OCR Review OCR value (if sent to OCR Review and found to be incorrect),
- the post-OCR Review OCR value,

---

[2] Survey-defined edits used as part of the OCR module determine if the OCR-read answer makes contextual sense for the survey. For this test, these edits are answer length and legal value edits.
[3] About 28 percent of the control treatment and 29 percent of the OCR treatment responded by Internet. The total self-response rate for the control and OCR treatment is about 52 percent and about 53 percent respectively.

- the independently keyed value, and, if necessary,
- the adjudicated value(s) as determined by the quality assurance staff in NPC.

As these are operational data, they are not weighted and we performed no statistical testing.

To answer research question E, we used unedited response data received from NPC for the control and OCR treatments (mail forms only). These data are weighted.

This report answers the following research questions:

A. For a given item with a numeric write-in, how often does the OCR value match exactly to the keyed value? How often does OCR fail to read a numeric value?

We calculate the following metrics for each item and overall. All rates compare the OCR value (post-OCR Review value if sent to OCR Review) to the keyed value. Keyed values refer to both values keyed when OCR does not read a value and values keyed in the 100 percent verification of OCR values. The number of non-missing keyed values in the denominator represents the universe of all numeric values requiring data capture.

$$\text{OCR Match Rate} = \frac{\text{Number of identical OCR and keyed values}}{\text{Number of non-missing keyed values}} * 100$$

$$\text{OCR Value Not Read Rate} = \frac{\text{Number of missing OCR values}}{\text{Number of non-missing keyed values}} * 100$$

$$\text{OCR Non-Match Rate} = \frac{\text{Number of differing OCR and keyed values}}{\text{Number of non-missing keyed values}} * 100$$

Some OCR values have leading zeros due to respondents filling in each dentil provided for a given item. Identical OCR and keyed values include values that match when these leading zeros are dropped from the OCR value.

B. Where values differ, how often is the OCR value correct?

To answer this question, we show, of the non-matching OCR and keyed values, which value is 'truth' (as determined by the QA staff in NPC) as a percentage of all non-missing keyed values. We calculated the following metrics for each item and overall.

$$\text{OCR Incorrect Rate} = \frac{\begin{array}{c}\text{Number of differing OCR and keyed values}\\ \text{where OCR value is in error}\end{array}}{\text{Number of non-missing keyed values}} * 100$$

$$\text{OCR Correct Rate} = \frac{\text{Number of differing OCR and keyed values where OCR value is correct}}{\text{Number of non-missing keyed values}} * 100$$

C. Of all values read by OCR, how often is the OCR value incorrect?

To answer this question, we consider only those values read by OCR and determine the percentage of OCR values that were read in error. We calculated the following metric for each item and overall.

$$\text{OCR Read in Error Rate} = \frac{\text{Number of OCR values read in error}}{\text{Number of non-missing OCR values}} * 100$$

D. Where values differ, what is the magnitude of that difference?

When the OCR and keyed values differ, we calculate the following metrics:

- Where the OCR value is correct, distributions of differences (minimum, maximum, and quartiles) in the OCR and keyed values by item

- Where the keyed value is correct, distributions of differences (minimum, maximum, and quartiles) in the OCR and keyed values by item

E. Does the distribution of data for numeric items differ due to the format change needed to accommodate OCR technology?

Using unedited response data received from NPC for the control and OCR treatments (data on paper forms only), we compare data distributions and item nonresponse rates for select numeric items. These unedited data represent data provided by the respondent not yet subjected to logic to account for nonresponse and data inconsistencies. For the item nonresponse rates, we calculate standard errors using replicate base weights[4] and perform statistical testing to determine if the rates are statistically significant at the 0.1 significance level. For the data distributions (not including item nonresponse), we use the chi-square test (significance level of 0.1) to determine if the weighted distributions are different due to the format change.

The numeric items selected for this comparison of distributions are:

- Age
- DOB month

---

[4] See U.S Census Bureau (2009), chapter 12, for more information on variance estimation.

- DOB year
- Year moved in
- Rooms
- Electricity cost
- Rent
- Value
- Mortgage
- Educational attainment write-in (grades 1-11)
- Year last married
- Commute time
- Hours worked
- Wages
- Self-employment income
- Interest income
- Social Security income
- Supplemental Security income (SSI)
- Public assistance income
- Retirement income
- Other income
- Total income

IV.   Results

We captured numeric write-in fields on paper forms in the QDT OCR treatment using OCR technology. Clerks also keyed all numeric entries.   The term 'keyed value' in this section refers to both values keyed when OCR does not read a value and values keyed in the 100 percent verification of OCR values. The total number of more than 71,000 keyed values represents the universe of all numeric values requiring data capture. Table 2 shows these keyed values by item. It is important to understand that a small subset of respondents answer some of these items which can make the subsequent results more variable for those items as compared to other items that have a larger universe of responses.

Table 2. Number of Keyed Values For Each ACS Numeric Write-In Fields (Paper Forms in the OCR Treatment)

Cover Page Items		Housing Items		Detailed Person Items	
Month	1,867	Year built write-in	610	Citizenship write-in	386
Day	1,838	Move in month	1,798	Year of entry	799
Year	1,862	Move in year	1,874	Grade attend write-in	717
Area code	1,886	Rooms	1,900	Grade comp write-in	900
Phone prefix	1,881	Bedrooms	1,903	Migration zip	506
Phone number	1,878	Electricity amount	1,773	Year last married	2,516
Number of people	1,931	Gas amount	1,210	Place of work zip	1,444
		Water amount	1,296	Total riders	1,399
Basic Person Items		Heat fuel amount	511	Leave home hour	1,434
Age	4,332	Condo amount	145	Leave home minute	1,415
Date of birth month	4,112	Rent amount	647	Travel time to work	1,524
Date of birth day	4,127	Value amount	1,326	Hours worked	1,930
Date of birth year	4,141	Tax amount	1,233	Wages amount	1,855
		Insurance amount	1,036	Self-employ amount	360
		Mortgage amount	683	Interest amount	609
		Second mortgage amount	174	SS amount	1,059
		Mobile home amount	203	SSI amount	380
				Public assist amount	222
				Retirement amount	691
				Other income amount	352
				Total income amount	2,641

Source: U.S. Census Bureau, 2013 American Community Survey Questionnaire Design Test, July to August 2013

For a given item with a numeric write-in, how often does the OCR value match exactly to the keyed value? How often does OCR fail to read a numeric value?

Figure 3 shows how often the OCR value matched the keyed value. The figure groups the items by where they occur on the form (cover, basic person, housing, detailed person). Item names ending in 'amount' indicate monetary values. Across more than 71,000 values, the OCR value matched the keyed value about 73 percent of the time[5]. OCR did not read almost 27 percent of all values most often because it could not determine one or more of the numbers in the field with an acceptable level of confidence or because the value read failed an edit. When OCR cannot determine the value for a field, the field goes to a keyer. The OCR value and

---

[5] Fields requiring the keyer to enter a value with a specific length are often keyed with a leading zero. Likewise, OCR will read leading zeros provided by the respondent that keyers may not key (for fields not requiring a specific length). For comparison purposes, leading zeros are dropped in the determination of matching OCR and keyed values.

keyed value did not match for 0.3 percent of all keyed values.

Across items, the degree to which OCR was able to read and match the keyed value varied greatly. For the cover and basic person items, the OCR value matched the keyed value about 81 percent of the time and OCR did not read about 19 percent of values. Among these items, the phone number components on the cover had the highest rates of the OCR value not matching the keyed value (about 0.8 percent). This is due to keying rules that direct a keyer to blank all three phone number component fields if any of the fields are blank or incomplete. This occurred for about half of the OCR and keyed values that do not match so a better estimate of the rate of non-matching values for these fields would be about 0.4 percent. For the other items on the cover and in the basic person section, the OCR and keyed values do not match for about 0.2 percent of values.

In the housing section of the form, a few items stand out from the rest. OCR did not read about 70 percent of the year built write-in fields. Upon further investigation, about half of these received an OCR error code indicating that the value determined by OCR failed an edit. This write-in instructs the respondent to provide a year only if the housing unit was built in 2000 or later. While we instruct keyers to key what they see, any OCR values that were before 2000 caused the edit to fail, which sent these fields to a keyer without an OCR answer. If not for this edit for this item, the OCR value would have matched about 67 percent (up from about 30 percent) of the keyed values and the percentage of the values where the OCR and keyed value did not match would have increased from 0.2 percent to 0.8 percent. The percent of year built write-in fields not read by OCR would have decreased by 37 percent or about 225 fields. The legal value edits affected other items in the housing section similarly. If these edits were not in place, the larger percentages of values not read by OCR for heating fuel amount and insurance amount would have seen reductions of about 40 percent and 55 percent respectively.

Condo amount and mobile home amount in the housing section also have large percentages of values not read by OCR due mostly to OCR read failure (one or more characters could not be determined) or OCR confidence failure (one or more characters have a low confidence value). Because few respondents answer these questions, the fields have small universes and therefore these measures may be more variable. In the housing section, the item with the highest percentage of non-matching OCR and keyed values is mobile home amount (1.5 percent). However, this percentage accounts for only three cases where the values did not match.

The detailed person section also shows varying levels of matching OCR and keyed values. For the items where OCR did not read a value for a large percentage of keyed values, it generally had more to do with OCR read failure (one or more characters could not be determined) or OCR confidence failure (one or more characters have a low confidence value) rather than an edit failure. The citizenship write-in and year of entry fields expect a year entry like the year field on the cover and the date of birth year field in the basic person section. However, these fields in the detailed person section have a much higher percentage of values not read by OCR due to OCR read or confidence failure. Upon reviewing images for some of these fields, it was not immediately apparent why OCR had difficulty reading the value. The leave home hour field is a field requiring two digits (keyers will key a leading zero if the respondent writes 7 instead of 07). When OCR reads just the one digit provided by the respondent, it fails a field length

edit and consequently does not provide a value. This occurs for about three-quarters of all values not read by OCR for this field. Other fields in this section with a large percentage of values not read by OCR are some of the income fields, specifically, self-employment, SSI, public assistance, retirement, and other income. For all of these income fields, over half of the values not read by OCR were due to OCR read or confidence failure. Legal value edits affected the retirement and SSI fields the most, accounting for 30 percent and 16 percent of the non-reads respectively. Other OCR non-reads occur when OCR detected a presence in a field but determined it to be a blank value. These detected blanks (all confirmed blanks by a keyer) account for 23 percent of self-employment values not read, 13 percent of SSI values not read, 23 percent of public assistance values not read, and 12 percent of other income values not read. Among all detailed person items, social security amount had the largest percentage of non-matching values (10 values did not match).

Figure 3. Matching OCR and Keyed Values by Item and Overall

Source: U.S. Census Bureau, 2013 American Community Survey Questionnaire Design Test, July to August 2013

As seen from the discussion above, sponsor-defined value and answer length edits implemented by OCR had a great effect on some items. Originating from the iCADE Matrix, most of the legal value and answer length edits are intended as soft edits for keyers. This means that after

checking the value for error, the keyer can input a value that falls outside of the legal value range or expected answer length. These soft edits simply alert a keyer to a possible miskeyed value. A few of the edits require the keyer to enter a value with a specific length. These edits mainly affect date, time, and zipcode fields. In the test, OCR's implementation of these edits blocked the output of values that failed a legal value or answer length edit. The values that would have been captured by OCR had the edit not been in place were included on the analysis file. Therefore, we can simulate the elimination of these edits on the OCR data and observe the effects on the rate of matching OCR and keyed values and the rate of values not read by OCR.

Figure 4 shows the simulation of the elimination of the legal value and answer length edits. Compared to Figure 3, we see that the percentage of values not read by OCR decreases dramatically for some items (year built write-in, heat fuel amount, insurance amount, leave home hour, and retirement amount) and about four percent overall. Across all items, the percent of keyed values that do not match the OCR value increases to about 0.6 percent (compared to 0.3 percent in Figure 3). Items that see some of the biggest increases in the rate of non-matching keyed and OCR values include some of the year items (cover page, date of birth, year built, and move-in year), condo amount, place of work zip, SSI amount, and public assistance amount. We see the largest increase in the rate of non-matching keyed and OCR values for the public assistance amount field when the edits are eliminated.   All of these non-matches are cases where the OCR value is '1' when the keyed value indicates that the field is blank. A review of the images showed that OCR read a line that the respondent marked through the field to denote that the field was not applicable as a '1'.

Figure 4. Simulated Matching OCR and Keyed Values by Item and Overall - No Legal Value and Answer Length Edits

Source: U.S. Census Bureau, 2013 American Community Survey Questionnaire Design Test, July to August 2013

Where values differ, how often is the OCR value correct?

Figure 5 shows, of all keyed values, the percent where the OCR and keyed value did not match and how often the OCR value was correct. These data are based on the original test data, which includes sponsor-defined edits. The heat fuel, second mortgage, citizenship write-in, public assistance, and other income fields did not have any non-matching OCR and keyed values. Across all fields, there are varying levels of OCR accuracy. For the condo amount, year of entry, migration zip, leave home hour, interest amount, and retirement amount fields, the OCR value was always incorrect. The OCR value was always correct for the year (cover page), bedrooms, and grade completion write-in fields. The mobile home amount field had the highest rate of inconsistent OCR and keyed values (1.5 percent of all keyed values) and for about a third of them, the OCR value was correct. Overall, the OCR value was correct for about 34 percent of the non-matching OCR and keyed values.

Figure 5. Non-Matching OCR and Keyed Values and OCR Accuracy Result by Item and Overall

Source: U.S. Census Bureau, 2013 American Community Survey Questionnaire Design Test, July to August 2013

Of all values read by OCR, how often is the OCR value incorrect?

Figure 6, based on the original test data including sponsor-defined edits, considers only the values that OCR read and shows how often those values were in error. This is analogous to the keying error rate that is required to be less than one percent for all items keyed in a batch, which typically consists of 50 forms. Across all items, the OCR error rate is 0.3 percent. Most items have an OCR error rate less than one percent. Only four items have an individual OCR error rate greater than one percent – condo amount, mobile home amount, self-employ amount, and SS amount. However, small portions of the population generally answer these four items, which together account for just 2.4 percent of all numeric items keyed for this treatment in the test.
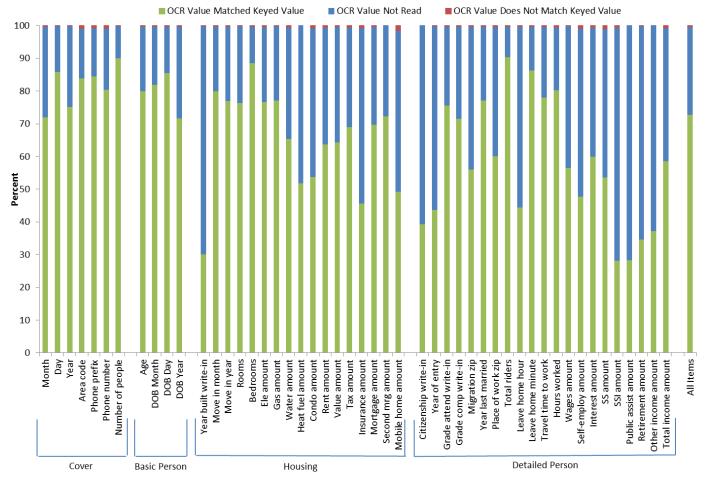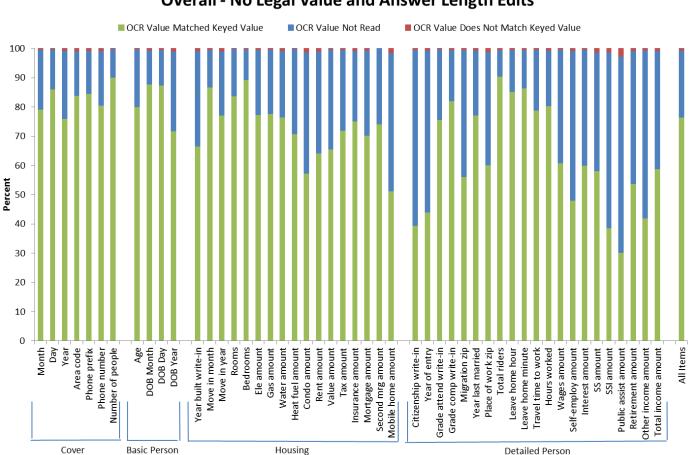
Figure 6. OCR Values Read in Error by Item and Overall

Source: U.S. Census Bureau, 2013 American Community Survey Questionnaire Design Test, July to August 2013

Figure 7 shows what the worst-case scenario OCR error rate would be if the legal value and answer length edits were removed. Because there is not an adjudicated value for values that were not read by OCR due to the edits, we cannot know if the simulated OCR value or the keyed value is correct when the values do not agree. These error rates are worst-case scenario error rates because they assume that the OCR value is always in error when the simulated OCR value and keyed value does not match. The items with the highest error rates include some of the year fields, place of work zip, condo amount, mobile home amount, and some of the income fields. The public assistance amount field has the highest error rate but this is somewhat misleading as the error rate accounts for just six values where the OCR and keyed values did not match. A similar situation also exists for the SSI amount field. Based on this data, we may want to consider keeping the edits for the year fields (with a change to the year built write-in edit to accept values prior to 2000) and the two zip code fields. In addition, an edit for amount fields that sends values of '1' to a keyer would avoid the occurrences of OCR reading a line that marked through the field (denoting that the field was not applicable) as a '1', thus reducing the OCR error rates for those fields. Even without these proposed edits, the OCR error rate without any edits across all items is 0.6 percent.

13

Figure 7. Simulated OCR Values Read in Error by Item and Overall - No Legal Value and Answer Length Edits

Where values differ, what is the magnitude of that difference?

Table 3 shows the magnitude of the differences (OCR value minus keyed value) when the OCR and keyed values differ and the OCR value is correct. In some cases, the keyer incorrectly blanked a field or simply failed to key a field while OCR captured the correct value. We show these cases separately in the table. The biggest difference occurs for the housing value amount field. For this particular value, OCR read 950,000 while the keyer incorrectly keyed 9,500,000. Many of the other differences are the result of the keyer keying an extra digit, omitting a digit, or simply miskeying one digit of the number. In these cases, OCR improves the quality of data on paper forms.

Table 3. Magnitude of Differences when ORC and Keyed Values Differ and OCR Value is Correct

Item	Total	Keyer Incorrectly Blanked Field; OCR Captured Correct Value						OCR Captured Correct Value				
		Number	Differences					Differences				
			Min	Q1	Median	Q3	Max	Min	Q1	Median	Q3	Max
Cover Items												
Month	1							2	2	2	2	2
Day	2							-5	-5	-2	1	1
Year	2							-2	-2	-0.5	1	1
Area code	3	1	77	77	77	77	77	-500	-500	-385	-270	-270
Phone prefix	4	1	611	611	611	611	611	-100	-100	27	30	30
Phone number	5	1	3,639	3,639	3,639	3,639	3,639	-710	-255	1,100	2,649	3,298
Number of people	1							-10	-10	-10	-10	-10
Basic Person Items												
Age	2							1	1	18.5	36	36
DOB Month	2	1	12	12	12	12	12	1	1	1	1	1
DOB Day	6							-10	-5	-2.5	2	3
DOB Year	3							-4	-4	-1	20	20
Housing Items												
Move in month	1							1	1	1	1	1
Move in year	3							-10	-10	3	10	10
Rooms	1							-1	-1	-1	-1	-1
Bedrooms	4	1	3	3	3	3	3	-30	-30	-3	-1	-1
Ele amount	2	1	105	105	105	105	105	-50	-50	-50	-50	-50
Gas amount	1	1	0	0	0	0	0					
Water amount	2	1	0	0	0	0	0	100	100	100	100	100
Rent amount	1							50	50	50	50	50
Value amount	2							-8,550,000	-8,550,000	-4,274,997	6	6
Tax amount	1							-9	-9	-9	-9	-9
Insurance amount	3	1	592	592	592	592	592	-50	-50	5	60	60
Mortgage amount	2							2	2	31	60	60
Mobile home amount	1	1	5,337	5,337	5,337	5,337	5,337					
Detailed Person Items												
Grade attend write-in	1							-2	-2	-2	-2	-2
Grade comp write-in	2	2	5	5	7.5	10	10					
Year last married	3							-1	-1	6	20	20
Place of work zip	2							-60	-60	-35	-10	-10
Total riders	1							-10	-10	-10	-10	-10
Leave home minute	2	1	0	0	0	0	0	30	30	30	30	30
Travel time to work	2							-3	-3	32.5	68	68
Hours worked	1							10	10	10	10	10

Wages amount	1							-1,000	-1,000	-1,000	-1,000	-1,000

Table 3. Continued…

		Keyer Incorrectly Blanked Field; OCR Captured Correct Value						OCR Captured Correct Value				
					Differences					Differences		
Item	Total	Number	Min	Q1	Median	Q3	Max	Min	Q1	Median	Q3	Max
Self-employ amount	1							-9,000	-9,000	-9,000	-9,000	-9,000
SS amount	4							-15,200	-7,625	-20	5,005	10,000
SSI amount	1							1	1	1	1	1
Total income amount	6							-19,000	-10,000	-997	20	14,517

Source: U.S. Census Bureau, 2013 American Community Survey Questionnaire Design Test, July to August 2013

Table 4 shows the magnitude of the differences (OCR value minus keyed value) when the OCR and keyed values differ and the keyed value is correct. In some cases, OCR captured a value that the keyer correctly blanked. We show these cases separately in the table. The phone number component fields on the cover page account for nearly half of these keyer-blanked fields. The keying procedure instructs keyers to blank all three of these fields if any one of the three fields are incomplete. More than a quarter of the keyer-blanked fields have a difference of -1 which indicates that OCR read the value as '1'. Reviewing the images for some of these fields show that OCR read as a value of '1' a line that the respondent marked through the field (denoting that the field was not applicable). For those values where the keyer entered the correct value, the difference for the SSI amount field and the largest difference for the SS amount field is due to the respondent entering a monthly value (indicated as monthly) instead of the requested yearly amount. The keyer in each of these cases added the letter 'm' to the keyed value, which ultimately converts the monthly amount to a yearly amount. The difference shown in Table 4 is the difference between the yearly value and the OCR read monthly value. The majority of the other differences between the OCR value and keyed value are due to the misread of a single digit of the value. Overall, the occurrence of discrepancies between the OCR and keyed values where the OCR value is in error is very low and the magnitude of those discrepancies appears to be nominal in most cases.

Table 4. Magnitude of Differences when ORC and Keyed Values Differ and Keyed Value is Correct

		Keyer Correctly Blanked Field						Keyer Entered Correct Value				
					Differences					Differences		
Item	Total	Number	Min	Q1	Median	Q3	Max	Min	Q1	Median	Q3	Max
Cover Items												
Month	5							-2	-2	-2	5	5
Day	1							3	3	3	3	3
Area code	14	10	-915	-803	-530	-202	0	-50	-30	-8	-1	5
Phone prefix	10	8	-788	-503	-159	0	0	-60	-60	-29	2	2
Phone number	11	4	-428	-214	0	0	0	-6	2	60	500	2,000
Number of people	1							-9	-9	-9	-9	-9
Basic Person Items												
Age	6	6	-83	-7	-4	-1	-1					
DOB Month	6	1	-4	-4	-4	-4	-4	1	5	6	10	10

Item	Total											
DOB Day	3							-10	-10	-4	5	5

Table 4. Continued…

		Keyer Correctly Blanked Field						Keyer Entered Correct Value				
		Number	Differences						Differences			
Item	Total	Number	Min	Q1	Median	Q3	Max	Min	Q1	Median	Q3	Max
DOB Year	9	1	-1,940	-1,940	-1,940	-1,940	-1,940	-10	0	11	50	50
Housing Items												
Year built write-in	1							1	1	1	1	1
Move in month	3							-1	-1	-1	3	3
Move in year	3							1	1	2	70	70
Rooms	1							5	5	5	5	5
Ele amount	4							-600	-396	-106	-15	-10
Gas amount	1							-10	-10	-10	-10	-10
Water amount	6							-123	-40	0	267	385
Condo amount	1							-10	-10	-10	-10	-10
Rent amount	3	1	-1,950	-1,950	-1,950	-1,950	-1,950	-50	-50	-35	-20	-20
Value amount	1							-900	-900	-900	-900	-900
Tax amount	5	2	-7	-7	-4	-1	-1	-90	-90	-50	10	10
Insurance amount	4	3	-11	-11	-7	-1	-1	100	100	100	100	100
Mortgage amount	1							600	600	600	600	600
Mobile home amount	2	1	-1	-1	-1	-1	-1	2,200	2,200	2,200	2,200	2,200
Detailed Person Items												
Year of entry	1							50	50	50	50	50
Grade attend write-in	1	1	-9	-9	-9	-9	-9					
Migration zip	1							-50	-50	-50	-50	-50
Year last married	3	1	-1,960	-1,960	-1,960	-1,960	-1,960	-6	-6	27	60	60
Place of work zip	3							-5,000	-5,000	-700	-2	-2
Total riders	3	1	-1	-1	-1	-1	-1	-6	-6	-5	-3	-3
Leave home hour	3							2	2	4	9	9
Leave home minute	2							-8	-8	-8	-7	-7
Travel time to work	4	1	-1	-1	-1	-1	-1	1	1	2	4	4
Hours worked	11	3	-1	-1	-1	-1	-1	-338	-263	-75	3	20
Wages amount	2							10,000	10,000	55,000	100,000	100,000
Self-employ amount	2	2	-1	-1	-1	-1	-1					
Interest amount	3	1	-1	-1	-1	-1	-1	-4	-4	248	500	500
SS amount	6							-80	4	3,662	12,309	24,200
SSI amount	1							9,427	9,427	9,427	9,427	9,427
Retirement amount	1							2	2	2	2	2
Total income amount	12	1	-1	-1	-1	-1	-1	-5,000	-100	-60	-2	6,000

Source: U.S. Census Bureau, 2013 American Community Survey Questionnaire Design Test, July to August 2013

Does the distribution of data for numeric items differ due to the format change needed to accommodate OCR technology?

To accommodate OCR technology for numeric values, the numeric write-in boxes on the paper forms for this treatment changed from a single box meant to contain the entire numeric value to multiple separate boxes, called dentils, meant to contain single digits of the numeric value. See Figure 1 and Figure 2 in the Background section for an example of this change. While this change in format is not radical, we wanted to ensure that it did not cause respondents to answer the numeric questions differently.

We compared the item nonresponse rates from the control and OCR treatments for these selected numeric items:

- age,
- date of birth month,
- date of birth year,
- year moved in,
- rooms,
- electricity cost,
- rent, value,
- mortgage,
- educational attainment write-in,
- year last married,
- commute time,
- hours worked,
- and all income fields (wages, self-employment income, interest income, Social Security income, Supplemental Security income, public assistance income, retirement income, other income, and total income).

T-tests performed at a significance level of 0.1 indicate that the item nonresponse rates do not differ between the control and OCR treatments for any of the items listed. See Attachment B for a table containing the nonresponse rates and associated test statistics for each item.

We also compared the response distributions from the control and OCR treatments for the selected numeric items noted above. The chi-square test, performed at a significance level of 0.1, rejected the null hypothesis of independence for four items: age, rooms, hours worked, and self-employment income. This suggests that respondents may have answered these items differently due to the format of the numeric write-in boxes. However, these items are not concentrated in any one section of the form that would indicate a specific issue in using dentils for a series of questions (i.e. income questions). For most of the items analyzed, the format change did not cause respondents to answer the numeric questions differently. Attachment B contains a table showing the distributions for each of the items evaluated.

V.   Conclusions

The use of OCR technology for numeric items in the OCR treatment of the QDT displays great promise. Overall, the OCR value matched the keyed value about 73 percent of the time. The OCR value and keyed value did not match for just 0.3 percent of all keyed values and for about 34 percent of these, the OCR value was correct. When the OCR value was not correct, the magnitude of the differences between the OCR and keyed value appears to be nominal in most cases. Additionally, of all values read by OCR, just 0.3 percent were in error, which is less than the one percent error rate required for keyed batches.

By simulating the elimination of legal value and answer length edits that are used by OCR, we see that the percentage of OCR and keyed values that matched increases by about four percent, the percentage of keyed values not read by OCR decreases by about four percent while the OCR error rate increases by about 0.3 percent. The elimination of these edits adversely affects some items more than others. Retaining just an edit for the year fields and an edit restricting OCR values of '1' in the amount fields would mitigate many of those adverse effects.

Another area of possible improvement to the OCR methodology is not sending values that OCR reads as a blank to a keyer. In the test, keyers keyed all values that OCR detected as a blank. For the self-employment amount, SSI amount, public assistance amount, and other income amount fields, keyers confirmed all of the OCR detected blanks. Further research into all OCR detected blanks shows that keyers confirmed all but four of the 601 detected blanks (0.7 detected blank in error). By accepting the values that OCR detects as blanks, we can further decrease the keying workload with little change to the OCR error rate.

References

Davis, Mary C. and Wakim, Anne (2014). 2014 American Community Survey Research and Evaluation Report Memorandum Series #ACS14-RER-03, DSSD 2014 American Community Survey Memorandum Series #ACS14-MP-02, "2013 American Community Survey Questionnaire Design Final Report", U.S. Census Bureau.
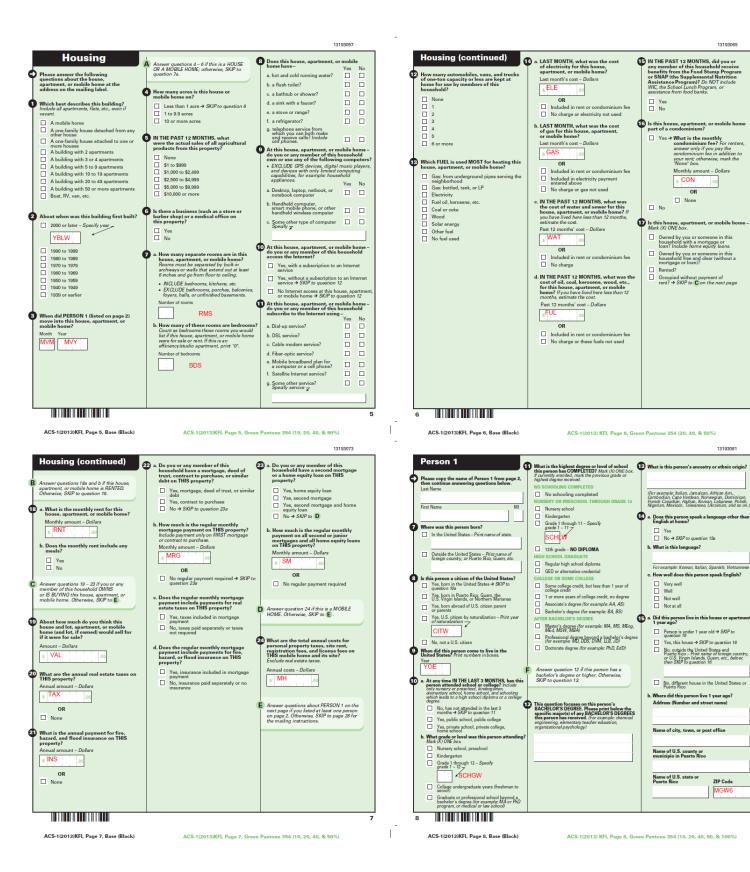
Terry, Rodney L. (2013), "Cognitive Pretesting for Navigation of 2013 ACS Questionnaire Design Test Questionnaires," U.S. Census Bureau, http://www.census.gov/srd/papers/pdf/ssm2013-17.pdf
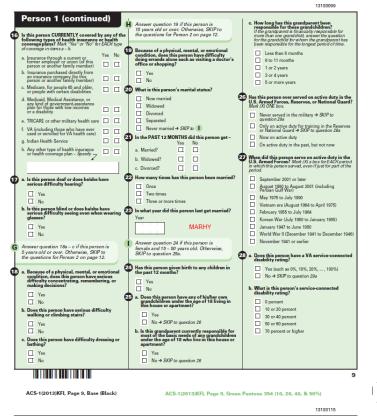
U.S. Census Bureau (2009), "(ACS) Design and Methodology," available at: http://www.census.gov/acs/www/methodology/methodology_main/ last accessed in April 2014.

Attachment A – Numeric Items on the Control Form and Associated Field Names (variable name in red)

Page 1

13193008

U.S. DEPARTMENT OF COMMERCE
Economics and Statistics Administration
U.S. CENSUS BUREAU

THE American Community Survey

Start Here

Respond online today at:
https://respond.census.gov/acs
OR
Complete this form and mail it back as soon as possible.

This form asks for information about the people who are living or staying at the address on the mailing label and about the house, apartment, or mobile home located at the address on the mailing label.

If you need help or have questions about completing this form, please call 1-800-354-7271. The telephone call is free.

Telephone Device for the Deaf (TDD): Call 1-800-582-8330. The telephone call is free.

¿NECESITA AYUDA? Si usted habla español y necesita ayuda para completar su cuestionario, llame sin cargo alguno al 1-877-833-5625. Usted también puede completar su entrevista por teléfono con un entrevistador que habla español. O puede responder por Internet en: https://respond.census.gov/acs

For more information about the American Community Survey, visit our web site at: http://www.census.gov/acs/www/

➡ **Please print today's date.**
Month Day Year
IMN IDY IYR

➡ **Please print the name and telephone number of the person who is filling out this form.** We may contact you if there is a question.
Last Name
First Name MI
Area Code + Number
RTEL

➡ **How many people are living or staying at this address?**
• INCLUDE everyone who is living or staying here for more than 2 months.
• INCLUDE yourself if you are living here for more than 2 months.
• INCLUDE anyone else staying here who does not have another place to stay, even if they are here for 2 months or less.
• DO NOT INCLUDE anyone who is living somewhere else for more than 2 months, such as a college student living away or someone in the Armed Forces on deployment.
Number of people
RPER

➡ Fill out pages 2, 3, and 4 for everyone, including yourself, who is living or staying at this address for more than 2 months. Then complete the rest of the form.

FORM ACS-1(2013)KFI (06-15-2012)
OMB No. 0607-0810

ACS-1(2013)KFI, Page 1, Base (Black)   ACS-1(2013)KFI, Page 1, Green Pantone 354 (20 & 40%)

Page 2

13193024

Person 1

(Person 1 is the person living or staying here in whose name this house or apartment is owned, being bought, or rented. If there is no such person, start with the name of any adult living or staying here.)

1 **What is Person 1's name?**
Last Name (Please print)   First Name   MI

2 **How is this person related to Person 1?**
☒ Person 1

3 **What is Person 1's sex?** Mark (X) ONE box.
☐ Male ☐ Female

4 **What is Person 1's age and what is Person 1's date of birth?**
Please report babies as age 0 when the child is less than 1 year old.
Print numbers in boxes.
Age (in years)   Month Day Year of birth
AGE   DBM DBD DBY

➡ NOTE: Please answer BOTH Question 5 about Hispanic origin and Question 6 about race. For this survey, Hispanic origins are not races.

5 **Is Person 1 of Hispanic, Latino, or Spanish origin?**
☐ No, not of Hispanic, Latino, or Spanish origin
☐ Yes, Mexican, Mexican Am., Chicano
☐ Yes, Puerto Rican
☐ Yes, Cuban
☐ Yes, another Hispanic, Latino, or Spanish origin – Print origin, for example, Argentinean, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on.

6 **What is Person 1's race?** Mark (X) one or more boxes.
☐ White
☐ Black, African Am., or Negro
☐ American Indian or Alaska Native – Print name of enrolled or principal tribe.

☐ Asian Indian ☐ Japanese ☐ Native Hawaiian
☐ Chinese ☐ Korean ☐ Guamanian or Chamorro
☐ Filipino ☐ Vietnamese ☐ Samoan
☐ Other Asian – Print race, for example, Hmong, Laotian, Thai, Pakistani, Cambodian, and so on. ☐ Other Pacific Islander – Print race, for example, Fijian, Tongan, and so on.

☐ Some other race – Print race.

Person 2

1 **What is Person 2's name?**
Last Name (Please print)   First Name   MI

2 **How is this person related to Person 1?** Mark (X) ONE box.
☐ Husband or wife ☐ Son-in-law or daughter-in-law
☐ Biological son or daughter ☐ Other relative
☐ Adopted son or daughter ☐ Roomer or boarder
☐ Stepson or stepdaughter ☐ Housemate or roommate
☐ Brother or sister ☐ Unmarried partner
☐ Father or mother ☐ Foster child
☐ Grandchild ☐ Other nonrelative
☐ Parent-in-law

3 **What is Person 2's sex?** Mark (X) ONE box.
☐ Male ☐ Female

4 **What is Person 2's age and what is Person 2's date of birth?**
Please report babies as age 0 when the child is less than 1 year old.
Print numbers in boxes.
Age (in years)   Month Day Year of birth
AGE   DBM DBD DBY

➡ NOTE: Please answer BOTH Question 5 about Hispanic origin and Question 6 about race. For this survey, Hispanic origins are not races.

5 **Is Person 2 of Hispanic, Latino, or Spanish origin?**
☐ No, not of Hispanic, Latino, or Spanish origin
☐ Yes, Mexican, Mexican Am., Chicano
☐ Yes, Puerto Rican
☐ Yes, Cuban
☐ Yes, another Hispanic, Latino, or Spanish origin – Print origin, for example, Argentinean, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on.

6 **What is Person 2's race?** Mark (X) one or more boxes.
☐ White
☐ Black, African Am., or Negro
☐ American Indian or Alaska Native – Print name of enrolled or principal tribe.

☐ Asian Indian ☐ Japanese ☐ Native Hawaiian
☐ Chinese ☐ Korean ☐ Guamanian or Chamorro
☐ Filipino ☐ Vietnamese ☐ Samoan
☐ Other Asian – Print race, for example, Hmong, Laotian, Thai, Pakistani, Cambodian, and so on. ☐ Other Pacific Islander – Print race, for example, Fijian, Tongan, and so on.

☐ Some other race – Print race.

2

ACS-1(2013)KFI, Page 2, Base (Black)   ACS-1(2013)KFI, Page 2, Green Pantone 354 (20 & 40%)

Page 3

13193032

Person 3

1 **What is Person 3's name?**
Last Name (Please print)   First Name   MI

2 **How is this person related to Person 1?** Mark (X) ONE box.
☐ Husband or wife ☐ Son-in-law or daughter-in-law
☐ Biological son or daughter ☐ Other relative
☐ Adopted son or daughter ☐ Roomer or boarder
☐ Stepson or stepdaughter ☐ Housemate or roommate
☐ Brother or sister ☐ Unmarried partner
☐ Father or mother ☐ Foster child
☐ Grandchild ☐ Other nonrelative
☐ Parent-in-law

3 **What is Person 3's sex?** Mark (X) ONE box.
☐ Male ☐ Female

4 **What is Person 3's age and what is Person 3's date of birth?**
Please report babies as age 0 when the child is less than 1 year old.
Print numbers in boxes.
Age (in years)   Month Day Year of birth
AGE   DBM DBD DBY

➡ NOTE: Please answer BOTH Question 5 about Hispanic origin and Question 6 about race. For this survey, Hispanic origins are not races.

5 **Is Person 3 of Hispanic, Latino, or Spanish origin?**
☐ No, not of Hispanic, Latino, or Spanish origin
☐ Yes, Mexican, Mexican Am., Chicano
☐ Yes, Puerto Rican
☐ Yes, Cuban
☐ Yes, another Hispanic, Latino, or Spanish origin – Print origin, for example, Argentinean, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on.

6 **What is Person 3's race?** Mark (X) one or more boxes.
☐ White
☐ Black, African Am., or Negro
☐ American Indian or Alaska Native – Print name of enrolled or principal tribe.

☐ Asian Indian ☐ Japanese ☐ Native Hawaiian
☐ Chinese ☐ Korean ☐ Guamanian or Chamorro
☐ Filipino ☐ Vietnamese ☐ Samoan
☐ Other Asian – Print race, for example, Hmong, Laotian, Thai, Pakistani, Cambodian, and so on. ☐ Other Pacific Islander – Print race, for example, Fijian, Tongan, and so on.

☐ Some other race – Print race.

Person 4

1 **What is Person 4's name?**
Last Name (Please print)   First Name   MI

2 **How is this person related to Person 1?** Mark (X) ONE box.
☐ Husband or wife ☐ Son-in-law or daughter-in-law
☐ Biological son or daughter ☐ Other relative
☐ Adopted son or daughter ☐ Roomer or boarder
☐ Stepson or stepdaughter ☐ Housemate or roommate
☐ Brother or sister ☐ Unmarried partner
☐ Father or mother ☐ Foster child
☐ Grandchild ☐ Other nonrelative
☐ Parent-in-law

3 **What is Person 4's sex?** Mark (X) ONE box.
☐ Male ☐ Female

4 **What is Person 4's age and what is Person 4's date of birth?**
Please report babies as age 0 when the child is less than 1 year old.
Print numbers in boxes.
Age (in years)   Month Day Year of birth
AGE   DBM DBD DBY

➡ NOTE: Please answer BOTH Question 5 about Hispanic origin and Question 6 about race. For this survey, Hispanic origins are not races.

5 **Is Person 4 of Hispanic, Latino, or Spanish origin?**
☐ No, not of Hispanic, Latino, or Spanish origin
☐ Yes, Mexican, Mexican Am., Chicano
☐ Yes, Puerto Rican
☐ Yes, Cuban
☐ Yes, another Hispanic, Latino, or Spanish origin – Print origin, for example, Argentinean, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on.

6 **What is Person 4's race?** Mark (X) one or more boxes.
☐ White
☐ Black, African Am., or Negro
☐ American Indian or Alaska Native – Print name of enrolled or principal tribe.

☐ Asian Indian ☐ Japanese ☐ Native Hawaiian
☐ Chinese ☐ Korean ☐ Guamanian or Chamorro
☐ Filipino ☐ Vietnamese ☐ Samoan
☐ Other Asian – Print race, for example, Hmong, Laotian, Thai, Pakistani, Cambodian, and so on. ☐ Other Pacific Islander – Print race, for example, Fijian, Tongan, and so on.

☐ Some other race – Print race.

3

ACS-1(2013)KFI, Page 3, Base (Black)   ACS-1(2013)KFI, Page 3, Green Pantone 354 (20 & 40%)

Page 4

13193040

Person 5

1 **What is Person 5's name?**
Last Name (Please print)   First Name   MI

2 **How is this person related to Person 1?** Mark (X) ONE box.
☐ Husband or wife ☐ Son-in-law or daughter-in-law
☐ Biological son or daughter ☐ Other relative
☐ Adopted son or daughter ☐ Roomer or boarder
☐ Stepson or stepdaughter ☐ Housemate or roommate
☐ Brother or sister ☐ Unmarried partner
☐ Father or mother ☐ Foster child
☐ Grandchild ☐ Other nonrelative
☐ Parent-in-law

3 **What is Person 5's sex?** Mark (X) ONE box.
☐ Male ☐ Female

4 **What is Person 5's age and what is Person 5's date of birth?**
Please report babies as age 0 when the child is less than 1 year old.
Print numbers in boxes.
Age (in years)   Month Day Year of birth
AGE   DBM DBD DBY

➡ NOTE: Please answer BOTH Question 5 about Hispanic origin and Question 6 about race. For this survey, Hispanic origins are not races.

5 **Is Person 5 of Hispanic, Latino, or Spanish origin?**
☐ No, not of Hispanic, Latino, or Spanish origin
☐ Yes, Mexican, Mexican Am., Chicano
☐ Yes, Puerto Rican
☐ Yes, Cuban
☐ Yes, another Hispanic, Latino, or Spanish origin – Print origin, for example, Argentinean, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on.

6 **What is Person 5's race?** Mark (X) one or more boxes.
☐ White
☐ Black, African Am., or Negro
☐ American Indian or Alaska Native – Print name of enrolled or principal tribe.

☐ Asian Indian ☐ Japanese ☐ Native Hawaiian
☐ Chinese ☐ Korean ☐ Guamanian or Chamorro
☐ Filipino ☐ Vietnamese ☐ Samoan
☐ Other Asian – Print race, for example, Hmong, Laotian, Thai, Pakistani, Cambodian, and so on. ☐ Other Pacific Islander – Print race, for example, Fijian, Tongan, and so on.

☐ Some other race – Print race.

➡ If there are more than five people living or staying here, print their names in the spaces for Person 6 through Person 12. We may call you for more information about them.

Person 6
Last Name (Please print)   First Name   MI
Sex ☐ Male ☐ Female   Age (in years) AGE

Person 7
Last Name (Please print)   First Name   MI
Sex ☐ Male ☐ Female   Age (in years) AGE

Person 8
Last Name (Please print)   First Name   MI
Sex ☐ Male ☐ Female   Age (in years) AGE

Person 9
Last Name (Please print)   First Name   MI
Sex ☐ Male ☐ Female   Age (in years) AGE

Person 10
Last Name (Please print)   First Name   MI
Sex ☐ Male ☐ Female   Age (in years) AGE

Person 11
Last Name (Please print)   First Name   MI
Sex ☐ Male ☐ Female   Age (in years) AGE

Person 12
Last Name (Please print)   First Name   MI
Sex ☐ Male ☐ Female   Age (in years) AGE

4

ACS-1(2013)KFI, Page 4, Base (Black)   ACS-1(2013)KFI, Page 4, Green Pantone 354 (20 & 40%)

Housing

➡ Please answer the following questions about the house, apartment, or mobile home at the address on the mailing label.

1 Which best describes this building? *Include all apartments, flats, etc., even if vacant.*
- [ ] A mobile home
- [ ] A one-family house detached from any other house
- [ ] A one-family house attached to one or more houses
- [ ] A building with 2 apartments
- [ ] A building with 3 or 4 apartments
- [ ] A building with 5 to 9 apartments
- [ ] A building with 10 to 19 apartments
- [ ] A building with 20 to 49 apartments
- [ ] A building with 50 or more apartments
- [ ] Boat, RV, van, etc.

2 About when was this building first built?
- [ ] 2000 or later – *Specify year* ⟶
  YBLW
- [ ] 1990 to 1999
- [ ] 1980 to 1989
- [ ] 1970 to 1979
- [ ] 1960 to 1969
- [ ] 1950 to 1959
- [ ] 1940 to 1949
- [ ] 1939 or earlier

3 When did PERSON 1 (listed on page 2) move into this house, apartment, or mobile home?

Month   Year
MVM   MVY

A Answer questions 4 – 6 if this is a HOUSE OR A MOBILE HOME; otherwise, SKIP to question 7a.

4 How many acres is this house or mobile home on?
- [ ] Less than 1 acre – *SKIP to question 6*
- [ ] 1 to 9.9 acres
- [ ] 10 or more acres

5 IN THE PAST 12 MONTHS, what were the actual sales of all agricultural products from this property?
- [ ] None
- [ ] $1 to $999
- [ ] $1,000 to $2,499
- [ ] $2,500 to $4,999
- [ ] $5,000 to $9,999
- [ ] $10,000 or more

6 Is there a business (such as a store or barber shop) or a medical office on this property?
- [ ] Yes
- [ ] No

7 a. How many separate rooms are in this house, apartment, or mobile home? *Rooms must be separated by built-in archways or walls that extend out at least 6 inches and go from floor to ceiling.*
- *INCLUDE bedrooms, kitchens, etc.*
- *EXCLUDE bathrooms, porches, balconies, foyers, halls, or unfinished basements.*

Number of rooms
[   ]   RMS

b. How many of these rooms are bedrooms? *Count as bedrooms those rooms you would list if this house, apartment, or mobile home were for sale or rent. If this is an efficiency/studio apartment, print "0".*

Number of bedrooms
[   ]   BDS

8 Does this house, apartment, or mobile home have –

	Yes	No
a. hot and cold running water?	[ ]	[ ]
b. a flush toilet?	[ ]	[ ]
c. a bathtub or shower?	[ ]	[ ]
d. a sink with a faucet?	[ ]	[ ]
e. a stove or range?	[ ]	[ ]
f. a refrigerator?	[ ]	[ ]
g. telephone service from which you can both make and receive calls? *Include cell phones.*	[ ]	[ ]

9 At this house, apartment, or mobile home – do you or any member of this household own or use any of the following computers?
- *EXCLUDE GPS devices, digital music players, and devices with only limited computing capabilities, for example: household appliances.*

	Yes	No
a. Desktop, laptop, netbook, or notebook computer	[ ]	[ ]
b. Handheld computer, smart mobile phone, or other handheld wireless computer	[ ]	[ ]
c. Some other type of computer *Specify* ⟶	[ ]	[ ]

[   ]

10 At this house, apartment, or mobile home – do you or any member of this household access the Internet?
- [ ] Yes, with a subscription to an Internet service
- [ ] Yes, without a subscription to an Internet service → *SKIP to question 12*
- [ ] No Internet access at this house, apartment, or mobile home → *SKIP to question 12*

11 At this house, apartment, or mobile home – do you or any member of this household subscribe to the Internet using –

	Yes	No
a. Dial-up service?	[ ]	[ ]
b. DSL service?	[ ]	[ ]
c. Cable modem service?	[ ]	[ ]
d. Fiber-optic service?	[ ]	[ ]
e. Mobile broadband plan for a computer or a cell phone?	[ ]	[ ]
f. Satellite Internet service?	[ ]	[ ]
g. Some other service? *Specify service* ⟶	[ ]	[ ]

[   ]

Housing (continued)

12 How many automobiles, vans, and trucks of one-ton capacity or less are kept at home for use by members of this household?
- [ ] None
- [ ] 1
- [ ] 2
- [ ] 3
- [ ] 4
- [ ] 5
- [ ] 6 or more

13 Which FUEL is used MOST for heating this house, apartment, or mobile home?
- [ ] Gas: from underground pipes serving the neighborhood
- [ ] Gas: bottled, tank, or LP
- [ ] Electricity
- [ ] Fuel oil, kerosene, etc.
- [ ] Coal or coke
- [ ] Wood
- [ ] Solar energy
- [ ] Other fuel
- [ ] No fuel used

14 a. LAST MONTH, what was the cost of electricity for this house, apartment, or mobile home?

Last month's cost – Dollars
$ ELE
  OR
- [ ] Included in rent or condominium fee
- [ ] No charge or electricity not used

b. LAST MONTH, what was the cost of gas for this house, apartment, or mobile home?

Last month's cost – Dollars
$ GAS
  OR
- [ ] Included in rent or condominium fee
- [ ] Included in electricity payment entered above
- [ ] No charge or gas not used

c. IN THE PAST 12 MONTHS, what was the cost of water and sewer for this house, apartment, or mobile home? *If you have lived here less than 12 months, estimate the cost.*

Past 12 months' cost – Dollars
$ WAT
  OR
- [ ] Included in rent or condominium fee
- [ ] No charge

d. IN THE PAST 12 MONTHS, what was the cost of oil, coal, kerosene, wood, etc., for this house, apartment, or mobile home? *If you have lived here less than 12 months, estimate the cost.*

Past 12 months' cost – Dollars
$ FUL
  OR
- [ ] Included in rent or condominium fee
- [ ] No charge or these fuels not used

15 IN THE PAST 12 MONTHS, did you or any member of this household receive benefits from the Food Stamp Program or SNAP (the Supplemental Nutrition Assistance Program)? *Do NOT include WIC, the School Lunch Program, or assistance from food banks.*
- [ ] Yes
- [ ] No

16 Is this house, apartment, or mobile home part of a condominium?
- [ ] Yes → **What is the monthly condominium fee?** *For renters, answer only if you pay the condominium fee in addition to your rent; otherwise, mark the "None" box.*

  Monthly amount – Dollars
  $ CON
  - [ ] None
- [ ] No

17 Is this house, apartment, or mobile home – *Mark (X) ONE box.*
- [ ] Owned by you or someone in this household with a mortgage or loan? *Include home equity loans.*
- [ ] Owned by you or someone in this household free and clear (without a mortgage or loan)?
- [ ] Rented?
- [ ] Occupied without payment of rent? → *SKIP to* **C** *on the next page*

Housing (continued)

B Answer questions 18a and b if this house, apartment, or mobile home is RENTED. Otherwise, SKIP to question 19.

18 a. What is the monthly rent for this house, apartment, or mobile home?

Monthly amount – Dollars
$ RNT .00

b. Does the monthly rent include any meals?
- [ ] Yes
- [ ] No

C Answer questions 19 – 23 if you or any member of this household OWNS or IS BUYING this house, apartment, or mobile home. Otherwise, SKIP to **E**.

19 About how much do you think this house and lot, apartment, or mobile home (and lot, if owned) would sell for if it were for sale?

Amount – Dollars
$ VAL .00

20 What are the annual real estate taxes on THIS property?

Annual amount – Dollars
$ TAX .00
  OR
- [ ] None

21 What is the annual payment for fire, hazard, and flood insurance on THIS property?

Annual amount – Dollars
$ INS .00
  OR
- [ ] None

22 a. Do you or any member of this household have a mortgage, deed of trust, contract to purchase, or similar debt on THIS property?
- [ ] Yes, mortgage, deed of trust, or similar debt
- [ ] Yes, contract to purchase
- [ ] No → *SKIP to question 23a*

b. How much is the regular monthly mortgage payment on THIS property? *Include payment only on FIRST mortgage or contract to purchase.*

Monthly amount – Dollars
$ MRG .00
  OR
- [ ] No regular payment required → *SKIP to question 23a*

c. Does the regular monthly mortgage payment include payments for real estate taxes on THIS property?
- [ ] Yes, taxes included in mortgage payment
- [ ] No, taxes paid separately or taxes not required

d. Does the regular monthly mortgage payment include payments for fire, hazard, or flood insurance on THIS property?
- [ ] Yes, insurance included in mortgage payment
- [ ] No, insurance paid separately or no insurance

23 a. Do you or any member of this household have a second mortgage or a home equity loan on THIS property?
- [ ] Yes, home equity loan
- [ ] Yes, second mortgage
- [ ] Yes, second mortgage and home equity loan
- [ ] No → *SKIP to* **D**

b. How much is the regular monthly payment on all second or junior mortgages and all home equity loans on THIS property?

Monthly amount – Dollars
$ SM .00
  OR
- [ ] No regular payment required

D Answer question 24 if this is a MOBILE HOME. Otherwise, SKIP to **E**.

24 What are the total annual costs for personal property taxes, site rent, registration fees, and license fees on THIS mobile home and its site? *Exclude real estate taxes.*

Annual costs – Dollars
$ MH .00

E Answer questions about PERSON 1 on the next page if you listed at least one person on page 2. Otherwise, SKIP to page 28 for the mailing instructions.

Person 1

➡ **Please copy the name of Person 1 from page 2, then continue answering questions below.**

Last Name
[   ]

First Name   MI
[   ] [ ]

7 Where was this person born?
- [ ] In the United States – *Print name of state.*
  [   ]
- [ ] Outside the United States – *Print name of foreign country, or Puerto Rico, Guam, etc.*
  [   ]

8 Is this person a citizen of the United States?
- [ ] Yes, born in the United States → *SKIP to question 10a*
- [ ] Yes, born in Puerto Rico, Guam, the U.S. Virgin Islands, or Northern Marianas
- [ ] Yes, born abroad of U.S. citizen parent or parents
- [ ] Yes, U.S. citizen by naturalization – *Print year of naturalization* ⟶
  CITW
- [ ] No, not a U.S. citizen

9 When did this person come to live in the United States? *Print numbers in boxes.*

Year
YOE

10 a. At any time IN THE LAST 3 MONTHS, has this person attended school or college? *Include only nursery or preschool, kindergarten, elementary school, home school, and schooling which leads to a high school diploma or a college degree.*
- [ ] No, has not attended in the last 3 months → *SKIP to question 11*
- [ ] Yes, public school, public college
- [ ] Yes, private school, private college, home school

b. What grade or level was this person attending? *Mark (X) ONE box.*
- [ ] Nursery school, preschool
- [ ] Kindergarten
- [ ] Grade 1 through 12 – *Specify grade 1 – 12* ⟶
  SCHGW
- [ ] College undergraduate years (freshman to senior)
- [ ] Graduate or professional school beyond a bachelor's degree (for example: MA or PhD program, or medical or law school)

11 What is the highest degree or level of school this person has COMPLETED? *Mark (X) ONE box. If currently enrolled, mark the previous grade or highest degree received.*

NO SCHOOLING COMPLETED
- [ ] No schooling completed

NURSERY OR PRESCHOOL THROUGH GRADE 12
- [ ] Nursery school
- [ ] Kindergarten
- [ ] Grade 1 through 11 – *Specify grade 1 – 11* ⟶
  SCHLW
- [ ] 12th grade – **NO DIPLOMA**

HIGH SCHOOL GRADUATE
- [ ] Regular high school diploma
- [ ] GED or alternative credential

COLLEGE OR SOME COLLEGE
- [ ] Some college credit, but less than 1 year of college credit
- [ ] 1 or more years of college credit, no degree
- [ ] Associate's degree (for example: AA, AS)
- [ ] Bachelor's degree (for example: BA, BS)

AFTER BACHELOR'S DEGREE
- [ ] Master's degree (for example: MA, MS, MEng, MEd, MSW, MBA)
- [ ] Professional degree beyond a bachelor's degree (for example: MD, DDS, DVM, LLB, JD)
- [ ] Doctorate degree (for example: PhD, EdD)

F Answer question 12 if this person has a bachelor's degree or higher. Otherwise, SKIP to question 13.

12 This question focuses on this person's BACHELOR'S DEGREE. Please print below the specific major(s) of any BACHELOR'S DEGREES this person has received. *(For example: chemical engineering, elementary teacher education, organizational psychology)*

[   ]

13 What is this person's ancestry or ethnic origin?

[   ]

(For example: Italian, Jamaican, African Am., Cambodian, Cape Verdean, Norwegian, Dominican, French Canadian, Haitian, Korean, Lebanese, Polish, Nigerian, Mexican, Taiwanese, Ukrainian, and so on.)

14 a. Does this person speak a language other than English at home?
- [ ] Yes
- [ ] No → *SKIP to question 15e*

b. What is this language?
[   ]

For example: Korean, Italian, Spanish, Vietnamese

c. How well does this person speak English?
- [ ] Very well
- [ ] Well
- [ ] Not well
- [ ] Not at all

15 a. Did this person live in this house or apartment 1 year ago?
- [ ] Person is under 1 year old → *SKIP to question 16*
- [ ] Yes, this house → *SKIP to question 16*
- [ ] No, outside the United States and Puerto Rico – *Print name of foreign country, or U.S. Virgin Islands, Guam, etc., below; then SKIP to question 16*
  [   ]
- [ ] No, different house in the United States or Puerto Rico

b. Where did this person live 1 year ago?

Address (Number and street name)
[   ]

Name of city, town, or post office
[   ]

Name of U.S. county or municipio in Puerto Rico
[   ]

Name of U.S. state or Puerto Rico   ZIP Code
[   ]   MGW6

Person 1 (continued)

16 Is this person CURRENTLY covered by any of the following types of health insurance or health coverage plans? *Mark "Yes" or "No" for EACH type of coverage in items a – h.*

Yes No

a. Insurance through a current or former employer or union (of this person or another family member) ☐ ☐

b. Insurance purchased directly from an insurance company (by this person or another family member) ☐ ☐

c. Medicare, for people 65 and older, or people with certain disabilities ☐ ☐

d. Medicaid, Medical Assistance, or any kind of government-assistance plan for those with low incomes or a disability ☐ ☐

e. TRICARE or other military health care ☐ ☐

f. VA (including those who have ever used or enrolled for VA health care) ☐ ☐

g. Indian Health Service ☐ ☐

h. Any other type of health insurance or health coverage plan – *Specify* ☐ ☐

17 a. Is this person deaf or does he/she have serious difficulty hearing?
☐ Yes
☐ No

b. Is this person blind or does he/she have serious difficulty seeing even when wearing glasses?
☐ Yes
☐ No

G *Answer question 18a – c if this person is 5 years old or over. Otherwise, SKIP to the questions for Person 2 on page 12.*

18 a. Because of a physical, mental, or emotional condition, does this person have serious difficulty concentrating, remembering, or making decisions?
☐ Yes
☐ No

b. Does this person have serious difficulty walking or climbing stairs?
☐ Yes
☐ No

c. Does this person have difficulty dressing or bathing?
☐ Yes
☐ No

H *Answer question 19 if this person is 15 years old or over. Otherwise, SKIP to the questions for Person 2 on page 12.*

19 Because of a physical, mental, or emotional condition, does this person have difficulty doing errands alone such as visiting a doctor's office or shopping?
☐ Yes
☐ No

20 What is this person's marital status?
☐ Now married
☐ Widowed
☐ Divorced
☐ Separated
☐ Never married → *SKIP to* **I**

21 In the PAST 12 MONTHS did this person get –
Yes No
a. Married? ☐ ☐
b. Widowed? ☐ ☐
c. Divorced? ☐ ☐

22 How many times has this person been married?
☐ Once
☐ Two times
☐ Three or more times

23 In what year did this person last get married?
Year
[____] MARHY

I *Answer question 24 if this person is female and 15 – 50 years old. Otherwise, SKIP to question 25a.*

24 Has this person given birth to any children in the past 12 months?
☐ Yes
☐ No

25 a. Does this person have any of his/her own grandchildren under the age of 18 living in this house or apartment?
☐ Yes
☐ No → *SKIP to question 26*

b. Is this grandparent currently responsible for most of the basic needs of any grandchildren under the age of 18 who live in this house or apartment?
☐ Yes
☐ No → *SKIP to question 26*

c. How long has this grandparent been responsible for these grandchildren? *If the grandparent is financially responsible for more than one grandchild, answer the question for the grandchild for whom the grandparent has been responsible for the longest period of time.*
☐ Less than 6 months
☐ 6 to 11 months
☐ 1 or 2 years
☐ 3 or 4 years
☐ 5 or more years

26 Has this person ever served on active duty in the U.S. Armed Forces, Reserves, or National Guard? *Mark X ONE box.*
☐ Never served in the military → *SKIP to question 29a*
☐ Only on active duty for training in the Reserves or National Guard → *SKIP to question 28a*
☐ Now on active duty
☐ On active duty in the past, but not now

27 When did this person serve on active duty in the U.S. Armed Forces? *Mark (X) a box for EACH period in which this person served, even if just for part of the period.*
☐ September 2001 or later
☐ August 1990 to August 2001 (including Persian Gulf War)
☐ May 1975 to July 1990
☐ Vietnam era (August 1964 to April 1975)
☐ February 1955 to July 1964
☐ Korean War (July 1950 to January 1955)
☐ January 1947 to June 1950
☐ World War II (December 1941 to December 1946)
☐ November 1941 or earlier

28 a. Does this person have a VA service-connected disability rating?
☐ Yes (such as 0%, 10%, 20%, ... , 100%)
☐ No → *SKIP to question 29a*

b. What is this person's service-connected disability rating?
☐ 0 percent
☐ 10 or 20 percent
☐ 30 or 40 percent
☐ 50 or 60 percent
☐ 70 percent or higher

Person 1 (continued)

29 a. LAST WEEK, did this person work for pay at a job (or business)?
☐ Yes → *SKIP to question 30*
☐ No – Did not work (or retired)

b. LAST WEEK, did this person do ANY work for pay, even for as little as one hour?
☐ Yes
☐ No → *SKIP to question 35a*

30 At what location did this person work LAST WEEK? *If this person worked at more than one location, print where he or she worked most last week.*

a. Address (Number and street name)
[_____]

If the exact address is not known, give a description of the location such as the building name or the nearest street or intersection.

b. Name of city, town, or post office
[_____]

c. Is the work location inside the limits of that city or town?
☐ Yes
☐ No, outside the city/town limits

d. Name of county
[_____]

e. Name of U.S. state or foreign country
[_____]

f. ZIP Code
[PWW6]

31 How did this person usually get to work LAST WEEK? *If this person usually used more than one method of transportation during the trip, mark (X) the box of the one used for most of the distance.*
☐ Car, truck, or van
☐ Bus or trolley bus
☐ Streetcar or trolley car
☐ Subway or elevated
☐ Railroad
☐ Ferryboat
☐ Taxicab
☐ Motorcycle
☐ Bicycle
☐ Walked
☐ Worked at home → *SKIP to question 39a*
☐ Other method

J *Answer question 32 if you marked "Car, truck, or van" in question 31. Otherwise, SKIP to question 33.*

32 How many people, including this person, usually rode to work in the car, truck, or van LAST WEEK?
Person(s)
[____] JWRI

33 What time did this person usually leave home to go to work LAST WEEK?
Hour : Minute ☐ a.m. ☐ p.m.
[JWLH] [JWLM]

34 How many minutes did it usually take this person to get from home to work LAST WEEK?
Minutes
[____] JWMN

K *Answer questions 35 – 38 if this person did NOT work last week. Otherwise, SKIP to question 39a.*

35 a. LAST WEEK, was this person on layoff from a job?
☐ Yes → *SKIP to question 35c*
☐ No

b. LAST WEEK, was this person TEMPORARILY absent from a job or business?
☐ Yes, on vacation, temporary illness, maternity leave, other family/personal reasons, bad weather, etc. → *SKIP to question 38*
☐ No → *SKIP to question 36*

c. Has this person been informed that he or she will be recalled to work within the next 6 months OR been given a date to return to work?
☐ Yes → *SKIP to question 37*
☐ No

36 During the LAST 4 WEEKS, has this person been ACTIVELY looking for work?
☐ Yes
☐ No → *SKIP to question 38*

37 LAST WEEK, could this person have started a job if offered one, or returned to work if recalled?
☐ Yes, could have gone to work
☐ No, because of own temporary illness
☐ No, because of all other reasons (in school, etc.)

38 When did this person last work, even for a few days?
☐ Within the past 12 months
☐ 1 to 5 years ago → *SKIP to* **L**
☐ Over 5 years ago or never worked → *SKIP to question 47*

39 a. During the PAST 12 MONTHS (52 weeks), did this person work 50 or more weeks? Count paid time off as work.
☐ Yes → *SKIP to question 40*
☐ No

b. How many weeks DID this person work, even for a few hours, including paid vacation, paid sick leave, and military service?
☐ 50 to 52 weeks
☐ 48 to 49 weeks
☐ 40 to 47 weeks
☐ 27 to 39 weeks
☐ 14 to 26 weeks
☐ 13 weeks or less

40 During the PAST 12 MONTHS, in the WEEKS WORKED, how many hours did this person usually work each WEEK?
Usual hours worked each WEEK
[____] WKH

Person 1 (continued)

L *Answer questions 41 – 46 if this person worked in the past 5 years. Otherwise, SKIP to question 47.*

41 – 46 CURRENT OR MOST RECENT JOB ACTIVITY. *Describe clearly this person's chief job activity or business last week. If this person had more than one job, describe the one at which this person worked the most hours. If this person had no job or business last week, give information for his/her last job or business.*

41 Was this person –
Mark (X) ONE box.
☐ an employee of a PRIVATE FOR-PROFIT company or business, or of an individual, for wages, salary, or commissions?
☐ an employee of a PRIVATE NOT-FOR-PROFIT, tax-exempt, or charitable organization?
☐ a local GOVERNMENT employee (city, county, etc.)?
☐ a state GOVERNMENT employee?
☐ a Federal GOVERNMENT employee?
☐ SELF-EMPLOYED in own NOT INCORPORATED business, professional practice, or farm?
☐ SELF-EMPLOYED in own INCORPORATED business, professional practice, or farm?
☐ working WITHOUT PAY in family business or farm?

42 For whom did this person work?
If now on active duty in the Armed Forces, mark (X) this box ☐ and print the branch of the Armed Forces.
Name of company, business, or other employer
[_____]

43 What kind of business or industry was this? *Describe the activity at the location where employed. (For example: hospital, newspaper publishing, mail order house, auto engine manufacturing, bank)*
[_____]

44 Is this mainly – *Mark (X) ONE box.*
☐ manufacturing?
☐ wholesale trade?
☐ retail trade?
☐ other (agriculture, construction, service, government, etc.)?

45 What kind of work was this person doing? *(For example: registered nurse, personnel manager, supervisor of order department, secretary, accountant)*
[_____]

46 What were this person's most important activities or duties? *(For example: patient care, directing hiring policies, supervising order clerks, typing and filing, reconciling financial records)*
[_____]

47 INCOME IN THE PAST 12 MONTHS

Mark (X) the "Yes" box for each type of income this person received, and give your best estimate of the TOTAL AMOUNT during the PAST 12 MONTHS. (NOTE: The "past 12 months" is the period from today's date one year ago up through today.)

Mark (X) the "No" box to show types of income NOT received.

If net income was a loss, mark the "Loss" box to the right of the dollar amount.

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

a. Wages, salary, commissions, bonuses, or tips from all jobs. *Report amount before deductions for taxes, bonds, dues, or other items.*
☐ Yes → $ [WAG] .00
☐ No TOTAL AMOUNT for past 12 months

b. Self-employment income from own nonfarm businesses or farm businesses, including proprietorships and partnerships. *Report NET income after business expenses.*
☐ Yes → $ [SEM] .00 ☐ Loss
☐ No TOTAL AMOUNT for past 12 months

c. Interest, dividends, net rental income, royalty income, or income from estates and trusts. *Report even small amounts credited to an account.*
☐ Yes → $ [INT] .00 ☐ Loss
☐ No TOTAL AMOUNT for past 12 months

d. Social Security or Railroad Retirement.
☐ Yes → $ [SS] .00
☐ No TOTAL AMOUNT for past 12 months

e. Supplemental Security Income (SSI).
☐ Yes → $ [SSI] .00
☐ No TOTAL AMOUNT for past 12 months

f. Any public assistance or welfare payments from the state or local welfare office.
☐ Yes → $ [PA] .00
☐ No TOTAL AMOUNT for past 12 months

g. Retirement, survivor, or disability pensions. *Do NOT include Social Security.*
☐ Yes → $ [RET] .00
☐ No TOTAL AMOUNT for past 12 months

h. Any other sources of income received regularly such as Veterans' (VA) payments, unemployment compensation, child support or alimony. *Do NOT include lump sum payments such as money from an inheritance or the sale of a home.*
☐ Yes → $ [OI] .00
☐ No TOTAL AMOUNT for past 12 months

48 What was this person's total income during the PAST 12 MONTHS? *Add entries in questions 47a to 47h; subtract any losses. If net income was a loss, enter the amount and mark (X) the "Loss" box next to the dollar amount.*
☐ None OR $ [TI] .00 ☐ Loss
TOTAL AMOUNT for past 12 months

→ Continue with the questions for Person 2 on the next page. If no one is listed as person 2 on page 2, SKIP to page 28 for mailing instructions.

Attachment B – Weighted Distributions of Specific Items – Control vs. OCR Treatments
Source: U.S. Census Bureau, 2013 American Community Survey Questionnaire Design Test, July to August 2013

Age	Control Estimate	SE	OCR Estimate	SE	Difference (OCR - Control) Estimate	SE	Significant?
Nonresponse	1.1	0.2	1.5	0.3	0.4	0.4	no
Response Distribution							
0-17	18.2	1.1	16.9	1.0	Chi-Square Test Result: Distribution is statistically different (χ^2=12.4, p-value=0.0146)		
18-24	7.9	0.7	5.6	0.6			
25-44	18.1	1.0	18.7	1.0			
45-64	30.5	1.0	29.9	1.1			
65 and over	25.2	1.2	29.0	1.2			

Date of birth month	Control Estimate	SE	OCR Estimate	SE	Difference (OCR - Control) Estimate	SE	Significant?
Nonresponse	4.4	0.5	5.5	0.6	1.1	0.8	no
Response Distribution							
1	8.7	0.7	9.2	0.8			
2	8.5	0.7	8.0	0.7			
3	8.2	0.7	7.8	0.7			
4	9.0	0.7	8.7	0.7			
5	7.4	0.6	8.2	0.7	Chi-Square Test Result: Distribution is not statistically different (χ^2=7.9, p-value=0.7209)		
6	6.4	0.7	8.3	0.7			
7	8.7	0.6	7.5	0.7			
8	8.5	0.7	8.4	0.6			
9	8.1	0.7	8.9	0.7			
10	8.8	0.7	8.1	0.7			
11	8.3	0.7	8.5	0.7			
12	9.4	0.8	8.5	0.8			

Date of birth year	Control		OCR		Difference (OCR - Control)		
	Estimate	SE	Estimate	SE	Estimate	SE	Significant?
Nonresponse	4.0	0.5	4.7	0.5	0.7	0.8	no
Response Distribution							
1890-1899	0.0	0.0	0.0	0.0			
1900-1909	0.0	0.0	0.0	0.0			
1910-1919	0.3	0.1	0.3	0.2			
1920-1929	4.4	0.5	4.0	0.6			
1930-1939	9.8	0.8	10.3	0.7			
1940-1949	13.8	0.9	16.5	1.1	Chi-Square Test Result: Distribution is not statistically different (χ^2=10.5, p-value=0.3989)		
1950-1959	17.2	1.0	16.9	0.9			
1960-1969	13.0	0.8	13.3	0.8			
1970-1979	9.1	0.7	9.6	0.7			
1980-1989	8.3	0.7	8.7	0.8			
1990-1999	12.0	0.8	10.2	0.7			
2000-2009	8.7	0.9	7.8	0.7			
2010 or later	3.3	0.5	2.4	0.4			

Year moved in	Control		OCR		Difference (OCR - Control)		
	Estimate	SE	Estimate	SE	Estimate	SE	Significant?
Nonresponse	4.0	0.7	4.9	0.7	0.9	1.0	no
Response distribution							
Before 1970	8.6	1.0	8.2	1.1			
Between 1970 to 1979	7.0	1.0	8.0	1.0	Chi-Square Test Result: Distribution is not statistically different (χ^2=4.3, p-value=0.5013)		
Between 1980 to 1989	9.9	1.2	8.2	1.0			
Between 1990 to 1999	20.1	1.2	18.8	1.6			
Between 2000 to 2009	32.9	1.7	31.6	1.4			
After 2009	21.5	1.8	25.2	1.5			

Rooms	Control		OCR		Difference (OCR - Control)		
	Estimate	SE	Estimate	SE	Estimate	SE	Significant?
Nonresponse	4.3	0.7	3.5	0.7	-0.7	0.9	no
Response distribution							
1 rooms	1.3	0.5	0.8	0.3			
2 rooms	3.0	0.6	3.9	0.7			
3 rooms	7.3	1.0	6.8	1.0			
4 rooms	13.5	1.1	13.4	1.2	Chi-Square Test Result: Distribution is statistically different (χ^2=13.9, p-value=0.0856)		
5 rooms	15.6	1.4	19.7	1.3			
6 rooms	18.8	1.5	18.3	1.3			
7 rooms	17.0	1.2	12.6	1.1			
8 rooms	9.7	1.0	12.1	1.2			
9 rooms or more	13.9	1.2	12.5	1.3			

Electricity cost	Control		OCR		Difference (OCR - Control)		
	Estimate	SE	Estimate	SE	Estimate	SE	Significant?
Nonresponse	6.9	1.0	5.2	0.8	-1.7	1.3	no
Response distribution							
$1 to $99	41.9	2.0	46.1	1.8	Chi-Square Test Result: Distribution is not statistically different (χ^2=6.9, p-value=0.2298)		
$100 to $199	38.7	2.0	36.8	1.7			
$200 to $299	12.1	1.3	10.3	1.1			
$300 to $399	3.8	0.7	4.1	0.8			
$400 to $999	2.7	0.6	1.4	0.4			
$1,000 or more	0.7	0.3	1.3	0.4			

Rent	Control		OCR		Difference (OCR - Control)		
	Estimate	SE	Estimate	SE	Estimate	SE	Significant?
Nonresponse	2.8	1.2	5.5	1.6	2.7	2.1	no
Response distribution							
No rent	0.0	0.0	0.0	0.0	Chi-Square Test Result: Distribution is not statistically different (χ^2=0.2, p-value=0.9750)		
$1 to $999	70.5	3.6	68.6	2.9			
$1,000 to $1,999	23.1	3.4	25.0	3.1			
$2,000 to $2,999	3.5	1.3	3.7	1.4			
$3,000 to $3,999	0.0	0.0	0.0	0.0			
$4,000 or more	2.9	1.2	2.7	1.2			

Value	Control		OCR		Difference (OCR - Control)		
	Estimate	**SE**	**Estimate**	**SE**	**Estimate**	**SE**	**Significant?**
Nonresponse	11.6	1.3	11.9	1.5	0.3	2.2	no
Response distribution							
Less than $10,000	3.5	0.8	3.5	0.9			
$10,000 to $14,999	0.6	0.4	2.1	0.6			
$15,000 to $19,999	1.9	0.6	0.6	0.3			
$20,000 to $24,999	1.0	0.4	0.6	0.4			
$25,000 to $29,999	0.6	0.4	1.5	0.5			
$30,000 to $34,999	0.2	0.2	1.2	0.5			
$35,000 to $39,999	0.2	0.2	1.2	0.5			
$40,000 to $49,999	2.1	0.6	2.3	0.6			
$50,000 to $59,999	2.7	0.7	1.9	0.6			
$60,000 to $69,999	3.5	0.9	3.7	0.8			
$70,000 to $79,999	2.7	0.8	4.2	1.0	Chi-Square Test Result: Distribution is not statistically different (χ^2=29.1, p-value=0.1773)		
$80,000 to $89,999	4.8	1.0	5.2	1.0			
$90,000 to $99,999	5.0	0.9	4.2	0.9			
$100,000 to $124,999	9.8	1.5	9.4	1.4			
$125,000 to $149,999	6.3	1.0	8.3	1.2			
$150,000 to $174,999	7.5	1.3	8.1	1.3			
$175,000 to $199,999	7.1	1.2	7.5	1.2			
$200,000 to $249,999	12.1	1.5	7.7	1.3			
$250,000 to $299,999	6.5	1.2	4.2	1.0			
$300,000 to $399,999	9.6	1.3	10.4	1.5			
$400,000 to $499,999	5.0	1.0	4.8	1.1			
$500,000 to $749,999	2.7	0.7	3.1	0.8			
$750,000 to $999,999	1.9	0.6	1.7	0.6			
$1,000,000 or more	2.5	0.6	2.7	0.7			

Mortgage	Control		OCR		Difference (OCR - Control)		
	Estimate	**SE**	**Estimate**	**SE**	**Estimate**	**SE**	**Significant?**
Nonresponse	5.2	1.1	6.3	1.1	1.1	1.4	no
Response distribution							
No mortgage	50.1	2.2	50.9	2.3	Chi-Square Test Result: Distribution is not statistically different (χ^2=6.1, p-value=0.1944)		
$1 to $999	22.1	1.9	26.2	2.1			
$1,000 to $1,999	21.7	1.7	17.5	1.6			
$2,000 to $2,999	3.9	0.8	4.2	0.8			
$3,000 or more	2.2	0.6	1.2	0.5			

Educational attainment write-in (grades 1-11)	Control		OCR		Difference (OCR - Control)		
	Estimate	SE	Estimate	SE	Estimate	SE	Significant?
Nonresponse	4.6	1.5	3.2	1.2	-1.5	1.7	no
Response Distribution							
1	6.5	1.8	3.3	1.3			
2	7.0	1.8	3.7	1.5			
3	5.4	2.0	6.0	1.6			
4	7.5	1.9	8.4	1.8			
5	8.6	2.1	10.2	2.1	Chi-Square Test Result: Distribution is not statistically different (χ^2=8.6, p-value=0.6562)		
6	7.5	2.1	4.7	1.3			
7	5.3	1.6	4.2	1.3			
8	9.1	1.9	12.1	2.1			
9	8.6	1.9	12.6	1.9			
10	17.7	2.6	16.7	2.7			
11	16.1	2.8	17.2	2.6			
12	0.5	0.5	0.9	0.6			

Year last married	Control		OCR		Difference (OCR - Control)		
	Estimate	SE	Estimate	SE	Estimate	SE	Significant?
Nonresponse	4.1	0.6	4.1	0.7	0.0	1.0	no
Response Distribution							
1949 or before	3.1	0.5	3.4	0.6			
1950 to 1959	8.4	1.0	8.9	0.9			
1960 to 1969	15.2	1.1	13.5	1.2	Chi-Square Test Result: Distribution is not statistically different (χ^2=5.6, p-value=0.4660)		
1970 to 1979	13.5	1.2	16.0	1.3			
1980 to 1989	17.6	1.1	18.5	1.1			
1990 to 1999	17.8	1.0	15.7	1.0			
2000 or later	24.4	1.3	23.9	1.5			

Commute time	Control		OCR		Difference (OCR - Control)		
	Estimate	SE	Estimate	SE	Estimate	SE	Significant?
Nonresponse	3.7	0.8	5.6	0.9	1.9	1.3	no
Response distribution							
1 to 15 minutes	42.8	1.9	40.8	2.2			
16 to 30 minutes	34.4	2.2	35.1	2.0	Chi-Square Test Result: Distribution is not statistically different (χ^2=3.7, p-value=0.5988)		
31 to 45 minutes	12.0	1.3	14.3	1.5			
46 to 60 minutes	5.6	1.1	5.7	0.9			
61 to 120 minutes	4.7	1.0	3.2	0.8			
121 and over minutes	0.5	0.3	0.9	0.4			

Hours worked	Control		OCR		Difference (OCR - Control)		
	Estimate	SE	Estimate	SE	Estimate	SE	Significant?
Nonresponse	4.4	0.7	5.7	0.8	1.3	1.0	no
Response distribution							
1-14	8.8	1.0	5.7	0.8	Chi-Square Test Result: Distribution is statistically different (χ^2=7.1, p-value=0.0284)		
15-34	21.0	1.5	20.0	1.3			
35 or more	70.2	1.7	74.3	1.5			

Wages	Control		OCR		Difference (OCR - Control)		
	Estimate	SE	Estimate	SE	Estimate	SE	Significant?
Nonresponse	12.7	1.1	13.2	1.1	0.4	1.4	no
Response distribution							
$1-$9,999	19.7	1.5	15.3	1.3	Chi-Square Test Result: Distribution is not statistically different (χ^2=9.1, p-value=0.1658)		
$10,000-$24,999	20.5	1.5	19.4	1.4			
$25,000-$49,999	22.8	1.5	21.9	1.5			
$50,000-$99,999	17.9	1.3	19.8	1.3			
$100,000-$249,999	4.1	0.9	5.4	0.8			
$250,000 or more	1.0	0.3	0.7	0.3			
No wages	14.0	1.2	17.5	1.3			

Total income	Control		OCR		Difference (OCR - Control)		
	Estimate	SE	Estimate	SE	Estimate	SE	Significant?
Nonresponse	23.0	1.0	22.8	1.0	-0.2	1.5	no
Response distribution							
$1-$9,999	31.4	1.2	33.0	1.5	Chi-Square Test Result: Distribution is not statistically different (χ^2=2.9, p-value=0.7154)		
$10,000-$24,999	25.2	1.4	22.4	1.3			
$25,000-$49,999	22.0	1.2	22.1	1.3			
$50,000-$99,999	16.2	1.0	16.9	1.1			
$100,000-$249,999	4.2	0.8	4.8	0.7			
$250,000 or more	1.0	0.3	0.8	0.3			

Self-Employment Income	Control		OCR		Difference (OCR - Control)		
	Estimate	SE	Estimate	SE	Estimate	SE	Significant?
Nonresponse	17.9	1.4	18.8	1.4	0.9	1.8	no
Response distribution							
$1-$9,999	1.5	0.5	4.2	0.8	Chi-Square Test Result: Distribution is statistically different (χ^2=11.1, p-value=0.0868)		
$10,000-$24,999	2.7	0.7	1.9	0.6			
$25,000-$49,999	1.5	0.4	0.9	0.4			
$50,000-$99,999	0.7	0.3	0.6	0.3			
$100,000-$249,999	0.3	0.2	0.6	0.3			
$250,000 or more	0.3	0.2	0.1	0.1			
No income	93.0	0.9	91.6	1.1			

Interest Income	Control		OCR		Difference (OCR - Control)		
	Estimate	SE	Estimate	SE	Estimate	SE	Significant?
Nonresponse	25.5	1.1	26.1	1.0	0.6	1.5	no
Response distribution							
$1-$9,999	13.7	1.1	15.1	1.1			
$10,000-$24,999	2.4	0.5	1.8	0.4			
$25,000-$49,999	1.0	0.3	0.5	0.2	Chi-Square Test Result: Distribution is not statistically different (χ^2=4.3, p-value=0.5090)		
$50,000-$99,999	0.7	0.2	0.6	0.2			
$100,000-$249,999	0.3	0.2	0.5	0.2			
$250,000 or more	0.0	0.0	0.0	0.0			
No income	82.0	1.2	81.6	1.1			

Social Security Income	Control		OCR		Difference (OCR - Control)		
	Estimate	SE	Estimate	SE	Estimate	SE	Significant?
Nonresponse	21.4	1.0	19.8	1.0	-1.5	1.5	no
Response distribution							
$1-$9,999	12.5	1.0	13.7	0.9			
$10,000-$24,999	16.2	1.1	18.4	1.1			
$25,000-$49,999	1.2	0.3	0.9	0.3	Chi-Square Test Result: Distribution is not statistically different (χ^2=3.7, p-value=0.4470)		
$50,000-$99,999	0.2	0.1	0.2	0.1			
$100,000-$249,999	0.0	0.0	0.0	0.0			
$250,000 or more	0.0	0.0	0.0	0.0			
No income	70.0	1.4	66.8	1.3			

SSI Income	Control		OCR		Difference (OCR - Control)		
	Estimate	SE	Estimate	SE	Estimate	SE	Significant?
Nonresponse	22.0	1.0	20.9	1.1	-1.1	1.4	no
Response distribution							
$1-$9,999	6.2	0.7	6.3	0.5			
$10,000-$49,999	1.4	0.3	1.1	0.2			
$50,000-$99,999	0.0	0.0	0.0	0.0	Chi-Square Test Result: Distribution is not statistically different (χ^2=0.3, p-value=0.8692)		
$100,000-$249,999	0.0	0.0	0.0	0.0			
$250,000 or more	0.0	0.0	0.0	0.0			
No income	92.4	0.8	92.6	1.2			

Public Assistance Income	Control		OCR		Difference (OCR - Control)		
	Estimate	**SE**	**Estimate**	**SE**	**Estimate**	**SE**	**Significant?**
Nonresponse	21.2	0.9	19.6	1.1	-1.6	1.4	no
Response distribution							
$1-$24,999	3.7	0.6	3.4	0.5			
$25,000-$49,999	0.0	0.0	0.0	0.0	Chi-Square Test Result: Distribution is not statistically different (χ^2=0.1, p-value=0.7009)		
$50,000-$99,999	0.0	0.0	0.0	0.0			
$100,000-$249,999	0.0	0.0	0.0	0.0			
$250,000 or more	0.0	0.0	0.0	0.0			
No income	96.3	0.6	96.6	1.5			

Retirement Income	Control		OCR		Difference (OCR - Control)		
	Estimate	**SE**	**Estimate**	**SE**	**Estimate**	**SE**	**Significant?**
Nonresponse	21.1	1.0	19.9	1.1	-1.2	1.5	no
Response distribution							
$1-$9,999	7.9	0.8	8.4	0.7			
$10,000-$24,999	4.6	0.6	5.6	0.7	Chi-Square Test Result: Distribution is not statistically different (χ^2=2.8, p-value=0.7276)		
$25,000-$49,999	2.7	0.6	3.3	0.5			
$50,000-$99,999	1.1	0.3	0.9	0.3			
$100,000 or more	0.1	0.1	0.2	0.1			
No income	83.6	1.1	81.7	1.0			

Other Income	Control		OCR		Difference (OCR - Control)		
	Estimate	**SE**	**Estimate**	**SE**	**Estimate**	**SE**	**Significant?**
Nonresponse	21.0	1.0	19.9	1.2	-1.1	1.5	no
Response distribution							
$1-$9,999	4.8	0.6	4.4	0.7			
$10,000-$24,999	1.1	0.3	1.3	0.3	Chi-Square Test Result: Distribution is not statistically different (χ^2=0.4, p-value=0.9304)		
$25,000-$249,999	0.5	0.2	0.5	0.2			
$250,000 or more	0.0	0.0	0.0	0.0			
No income	93.7	0.7	93.8	0.7			