October 12, 2017

2017 AMERICAN COMMUNITY SURVEY RESEARCH AND EVALUATION REPORT
MEMORANDUM SERIES # ACS17-RER-12

MEMORANDUM FOR   Victoria Velkoff
                     Chief, American Community Survey Office

From: David Waddington
                     Chief, Social, Economic, and Housing Statistics Division (SEHSD)

Prepared by: Anthony Martinez
                     Social, Economic, and Housing Statistics Division (SEHSD)

Subject: 2016 American Community Survey Content Test Evaluation
                     Report: Class of Worker

Attached is the final report for the 2016 American Community Survey (ACS) Content Test for
Class of Worker. This report describes the results of the test for a revised version of the Class of
Worker questions.

If you have any questions about this report, please contact Lynda Laughlin at 301-763-6314,
Anthony Martinez at 301-763-3595, or Ana Montalvo at 301-763-5977.

Attachment

cc:
Agnes Kee (ACSO)
Jennifer Ortman (ACSO)
David Raglin (ACSO)
Patrick Cantwell (DSSD)
Broderick Oliver (DSSD)
Elizabeth Poehler (DSSD)
Anthony Tersine (DSSD)
Jennifer Day (SEHSD)
Kelly Holder (SEHSD)
Lynda Laughlin (SEHSD)
Nicole Scanniello (SEHSD)

Intentionally Blank

October 12, 2017

2016 American Community Survey Content Test Evaluation Report: Class of Worker

FINAL REPORT

United States Census Bureau
Population Division

Anthony Martínez,
Ana J. Montalvo
Social, Economic, and Housing Statistics Division
Andrew W. Roberts, Population Division

Intentionally Blank

Table of Contents

List of Tables

iii

List of Figures

EXECUTIVE SUMMARY

From February to June of 2016, the U.S. Census Bureau conducted the 2016 American Community Survey (ACS) Content Test, a field test of new and revised content. The primary objective was to test whether changes to question wording, response categories, and definitions of underlying constructs improve the quality of data collected. Both new and revised versions of existing questions were tested to determine if they could provide data of sufficient quality compared to a control version as measured by a series of metrics including item missing data rates, response distributions, comparisons with benchmarks, and response error. The results of this test will be used to help determine the future ACS content and to assess the expected data quality of revised questions and new questions added to the ACS.

The 2016 Content Test consisted of a nationally representative sample of 70,000 residential addresses in the United States, independent of the production ACS sample. The sample universe did not include Group Quarters, nor did it include housing units in Alaska, Hawaii, or Puerto Rico. The test was a split-panel experiment with one-half of the addresses assigned to the control treatment and the other half assigned to the test treatment. As in production ACS, the data collection consisted of three main data collection operations: 1) a six-week mailout period, during which the majority of self-response via internet and mailback were received; 2) a one-month Computer-Assisted Telephone Interview (CATI) period for nonresponse follow-up; and 3) a one-month Computer-Assisted Personal Interview (CAPI) period for a sample of the remaining nonresponse. For housing units that completed the original Content Test interview, a Content Follow-Up telephone reinterview was conducted to measure response error.

Class of Worker

This report discusses the Class of Worker question and the associated write-in question on Employer Name. Class of Worker categorizes people according to the type of ownership of the employing organization, helps specify whether an employed person is salaried or self-employed, and helps specify if the person works in the private sector or in government (i.e., the public sector). The Class of Worker question has been asked in its current version since the 1970 Census. This question is being revised to clarify the intent of the question and the response categories, clarify the definition of Unpaid Family Workers, and improve the definition of Active Duty military. It also included a revised question layout.

In the internet, mail, and CAPI modes, the test version of the Class of Worker question grouped response categories under three general headings. In the control, they were listed in one group with no headings.[1] In addition, the test added Active Duty as one of the response categories in the Government section, while the control used a checkbox as part of the Employer Name question. The wording for the question and response categories were revised in both the Class of Worker question and the Employer Name question.

---

[1] For CAPI, test treatment changes include the card shown to respondents when reading the Class of Worker question.

Class of Worker is part of a set of six employment questions, including Industry and Occupation. To signal that all employment questions refer to the same job, the series was renumbered from separate questions (control version) to a single series (test version). Lastly, the text and heading for the series immediately preceding Class of Worker was simplified in the test version.

In the CATI mode, the Class of Worker test question had five broad categories: Private Company or Organization, Government, Active Duty U.S. Armed Forces, Self-Employed, and Worked Without Pay in a For-Profit Family Business. Interviewers asked follow-up questions depending on which category the respondent selected. For the CATI control treatment, all categories except Worked Without Pay had a follow-up question. The test treatment included a follow-up question for Worked Without Pay, which asked if the respondent worked in this family business for 15 hours or more per week. If the response to this question was "No," then the respondent was out of universe for the rest of the employment series.

Research Questions and Results

This research was guided by several questions concerning item missing data rates, coding rates, response reliability, benchmark comparisons, and differences in response distributions between the control and test treatments.

Item missing data rates:
- Overall, the test produced no statistically significant differences in item missing data rates between the control and test treatments. However, for the mail mode, where multiple checkboxes marked are treated as a nonresponse, the test treatment (14.0 percent) had a significantly higher item missing data rate than the control (10.4 percent).

- For the Class of Worker question, cases with multiple marks are only possible in the mail mode. The test treatment had a higher item missing data rate in the mail mode mainly due to more respondents incorrectly marking two response categories instead of one. For example, incorrectly marking both the Private For-Profit and Self-Employed response categories.[2] We believe that this problem can be fixed in the editing process and should therefore not impact the final data quality.

Response reliability:
- The test treatment improved response reliability for Unpaid Family Workers.

---

[2] For the Class of Worker question, cases with multiple marks were treated as missing, as were responses of "Don't Know" or "Refused." In the control treatment, if the Active Duty checkbox was selected in the Employer Name question without a response to the main Class of Worker question, this was treated as a response for this analysis, in order to make the control and test categories comparable.

Response distributions, consistency checks, and benchmark comparisons:

- The proportion of Unpaid Family Workers was not significantly different between treatments, although 63.6 percent reported working at least 15 hours a week (consistent with the definition) whereas in the control it was 39.9 percent, a statistically significant difference.

- Consistency checks for wages, self-employment income, and Class of Worker were not significantly different between the control and test treatment.

- Lastly, because the ACS estimates for Unpaid Family Workers have been higher than Current Population Survey Annual Social and Economic Supplement (CPS ASEC) estimates, we examined whether the test treatment estimate was more consistent with these estimates. Overall, the estimates of Unpaid Family Workers for the test and control treatments fell outside of the CPS ASEC's confidence interval for these estimates (at the 90 percent confidence level).

Conclusion

We revised the instructions and wording of the questions to clarify the intent of the series of questions concerning employment characteristics, which includes the Class of Worker question. During cognitive testing, respondents preferred this format and it is similar in format to other federal surveys, such as the National Survey of College Graduates (NCSG).

The most notable improvement was the increase in the rate of Unpaid Family Workers working at least 15 hours or more in the test treatment. The test treatment's clarification of the definition of this group of workers likely led to better respondent understanding. In addition, adding Active Duty as a Class of Worker category and removing it from the Employer Name question appears to reduce respondent confusion as indicated by the lower rate of response error for Military industries in the *2016 American Community Survey Content Test Evaluation Report: Industry and Occupation.*

Any significant differences between the test and control estimates for Class of Worker were small in magnitude, suggesting future continuity in data reliability and consistency, and no break in series should the test version be implemented. Our expectation is that implementing the change in Class of Worker question format and the proposed modifications for the Industry and Occupation questions will improve the overall quality of the data collected by this series of questions.[3] The recommendation of the Industry and Occupation Statistics Branch is to move forward with the implementation of the test version of the Class of Worker question.

---

[3] Results from the test of the revised version of the Industry and Occupation questions are published in the following report, *2016 American Community Survey Content Test Evaluation Report: Industry and Occupation.*

Intentionally Blank

1 BACKGROUND

From February to June of 2016, the Census Bureau conducted the 2016 American Community Survey (ACS) Content Test, a field test of new and revised content. The primary objective was to test whether changes to question wording, response categories, and definitions of underlying constructs improve the quality of data collected. Both revised versions of existing questions and new questions were tested to determine if they could provide data of sufficient quality compared to a control version as measured by a series of metrics including item missing data rates, response distributions, comparisons with benchmarks, and response error. The results of this test will be used to help determine the future ACS content and to assess the expected data quality of revised questions and new questions added to the ACS.

The 2016 ACS Content Test included the following topics:
- Relationship
- Race and Hispanic Origin
- Telephone Service
- Computer and Internet Use
- Health Insurance Coverage
- Health Insurance Premium and Subsidy (new questions)
- Journey to Work: Commute Mode
- Journey to Work: Time of Departure for Work
- Number of Weeks Worked
- Class of Worker
- Industry and Occupation
- Retirement, Survivor, and Disability Income

This report discusses the Class of Worker topic.

1.1 Justification for Inclusion of Class of Worker in the Content Test

Data collected by the Class of Worker question make it possible to categorize people according to the type of ownership of the employing organization. The Bureau of Labor Statistics and the Bureau of Economic Analysis use these data to understand labor force trends. A question on the class of work has been asked on the census since 1910; the term Class of Worker first appeared in 1940. While the current question wording has been asked since 1970, the format and placement within the employment section changed. There is some concern that the estimates derived from ACS data do not agree, even approximately, with similar statistics from other sources, such as the CPS (U.S. Bureau of Labor Statistics, 2015), the Annual Survey of Public Employment and Payroll (ASPEP) (U.S. Census Bureau, 2015d), and the Survey of Business Owners (SBO) (U.S. Census Bureau, 2015c).

The revision was designed to improve question clarity by: (a) clarifying the definition of Unpaid Family Workers; (b) explicitly defining a category for those in Active Duty military and uniformed services status; (c) rewording the question and categories; and (d) reformatting the layout.

(a) *Unpaid Family Workers.* The Bureau of Labor Statistics defines Unpaid Family Workers as people "who work without pay for 15 or more hours per week on a farm or in a business operated by a member of the household to whom they are related by birth or marriage."[4] The current ACS question does not make clear the person must work 15 hours or more per week. Until 2008, the Labor Force question clearly defined the hours worked requirement for the Unpaid Family Worker category (U.S. Census Bureau, 2015b). When the reference to the hours worked requirement for Unpaid Family Workers was removed from the Labor Force question in 2008, this description was not added to the category in the Class of Worker question (Holder & Raglin 2007). Thus, it implicitly removed the 15 hours or more per week work requirement for Unpaid Family Workers from the questionnaire. Perhaps because of this, the ACS estimates for this group are higher than estimates from the CPS.[5] Clarifying the definition should bring the ACS more in line with CPS estimates.

(b) *Active Duty.* Since many people do not think of Active Duty military status as Federal Government work, respondents are unclear about which category to check in the Class of Worker question. This was apparent when the Class of Worker checkbox was compared to the write-in responses for the Industry and Occupation questions for internet and mail modes (self-response modes). There is additional confusion in the mail mode, which has an Active Duty status checkbox in the question on the Employer Name, which follows the Class of Worker question. Respondents, whether Active Duty or not, are confused by this checkbox, thinking it refers to employment status or veteran status (U.S. Census Bureau, 2014b; U.S. Census Bureau, 2015a). Active Duty status was integrated as a response category to the Class of Worker question on the 1996-1998 test versions of the ACS, but was removed to match Class of Worker categories on the Census 2000 long form in order to evaluate any differences between the ACS data and the information collected on the Census 2000 long form (U.S. Census Bureau, 2015b). For the 2016 ACS Content Test, we tested adding Active Duty back as a Class of Worker category and removing it from Employer Name question to reduce respondent confusion for both questions.

---

[4] The definition of Unpaid Family Workers is located in the Bureau of Labor Statistic's website - https://www.bls.gov/bls/glossary.htm#U.

[5] To establish this comparison, Unpaid Family Worker data for ACS 1-year estimates were obtained from Table B24080, Sex by Class of Worker for the Civilian Employed Population 16 years and over, for years 2011 through 2015, available at: https://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml. Data for CPS were extracted from DataFerrett using variable PEIO1COW (Main job) Class of Worker for years 2011 through 2015, available at: https://dataferrett.census.gov/.

(c) *Question Wording*. The question wording is confusing. In the CPS prior to 1994 and in Census 2000, the Class of Worker question was placed at the end of the employment characteristics series (Polivka & Rothgeb, 1993; U.S. Census Bureau, 2015b). In this placement, instructions and questions on the Industry and Occupation of the same job were asked first, providing context for the Class of Worker question. Asking the question "Was this person [*each Class of Worker category*]?" may not have been as confusing with context provided from the preceding employment questions. The 1996 National Content Test did not find any conclusive effects of different question order (Kirk, 1996).

In the CPS prior to 1994, Class of Worker was the last question, and interviewers often recorded answers to it, without asking the respondent (Polivka & Rothgeb, 1993). To address this issue, Class of Worker was placed first in the question series without any wording changes. The 1994 CPS version served as a model for the ACS. Because of the change in question placement, the new version on the CPS had statistically significantly higher estimates of the proportion of Self-Employed and significantly lower estimates of the proportion of wage and salary workers, and Unpaid Family Workers (Polivka & Miller, 1998).[6] Changing the question order without a corresponding change in question wording resulted in a question that was confusing to respondents (Raglin, 2014; U.S. Census Bureau, 2014b; U.S. Census Bureau, 2015a). We expect that rewording the question, and adding context and examples will improve data quality (Stapleton & Steiger, 2015).

(d) *Question Layout.* The response categories are difficult to read in the mail and internet versions of the questionnaire, as well as on the flashcard shown to respondents in the CAPI interview. Reformatting the categories with headings, grouping similar statuses, and providing more concise wording is expected to make the question easier to read and clarify its intent (Stapleton & Steiger, 2015).

Additional Considerations: Reference Period and Multiple Jobs. People currently unemployed, but who worked in the past five years (e.g., those who retired recently), find the reference period and job categories confusing (Raglin, 2014; U.S. Census Bureau, 2014b; U.S. Census Bureau, 2015a). During cognitive testing, several respondents reported the "main" job from which the individual retired or was laid-off, not the job held most recently (Stapleton & Steiger, 2015). These issues were considered in the design and cognitive testing of the final test question wording and format (Stapleton & Steiger, 2015).

---

[6] Starting in January 1994, the CPS underwent a major redesign both in the wording of the questionnaire and the methodology used to collect the data. The 60,000-household CPS sample was switched to the revised questionnaire and computerized collection procedure. Polivka and Miller (1998) compared data between the new and old CPS questionnaires.

1.2   Question Development

Initial versions of the new and revised questions were proposed by federal agencies participating in the U.S. Office of Management and Budget (OMB) Interagency Committee for the ACS. The initial proposals contained a justification for each change and described previous testing of the question wording, the expected impact of revisions to the time series and the single-year as well as five-year estimates, and the estimated net impact on respondent burden for the proposed revision.[7] For proposed new questions, the justification also described the need for the new data, whether federal law or regulation required the data for small areas or small population groups, if other data sources were currently available to provide the information (and why any alternate sources were insufficient), how policy needs or emerging data needs would be addressed through the new question, an explanation of why the data were needed with the geographic precision and frequency provided by the ACS, and whether other testing or production surveys had evaluated the use of the proposed questions.

The Census Bureau and the OMB, as well as the Interagency Council on Statistical Policy Subcommittee reviewed these proposals for the ACS. The OMB determined which proposals moved forward into cognitive testing. After OMB approval of the proposals, topical subcommittees were formed from the OMB Interagency Committee for the ACS, which included all interested federal agencies that use the data from the impacted questions. These subcommittees further refined the specific proposed wording that was cognitively tested.

The Census Bureau contracted with Westat to conduct three rounds of cognitive testing. The results of the first two rounds of cognitive testing informed decisions on specific revisions to the proposed content for the stateside Content Test (Stapleton and Steiger, 2015). In the first round, 208 cognitive interviews were conducted in English and Spanish and in two modes (self-administered on paper and interviewer-administered on paper). In the second round of testing, 120 cognitive interviews were conducted for one version of each of the tested questions, in English and Spanish, using the same modes as in the first round.

A third round of cognitive testing involved only the Puerto Rico Community Survey (PRCS) and the GQ versions of the questionnaire (Steiger, Anderson, Folz, Leonard, & Stapleton, 2015). Cognitive interviews in Puerto Rico were conducted in Spanish; GQ cognitive interviews were conducted in English. The third round of cognitive testing was carried out to assess the revised versions of the questions in Spanish and identify any issues with questionnaire wording unique to Puerto Rico and GQ populations.[8] The proposed changes identified through cognitive testing for each question topic were reviewed by the Census Bureau, the corresponding topical subcommittee, and the Interagency Council on Statistical Policy Subcommittee for the ACS. The OMB then provided final overall approval of the proposed wording for field testing.[9]

---

[7] The ACS produces both single and five-year estimates annually. Single-year estimates are produced for geographies with populations of 65,000 or more and five-year estimates are produced for all areas down to the block-group level, with no population restriction.

[8] Note that the field testing of the content was not conducted in Puerto Rico or in GQs. See the Methodology section for more information.

[9] A cohabitation question and domestic partnership question were included in cognitive testing but ultimately we decided not to move forward with field testing these questions.

For Class of Worker, two versions of the question were tested on paper in the first round of cognitive testing. Version 1 used three subheadings to describe the types of employers ("Private Sector," "Government," "Self-Employed or Other") and response options were more concisely worded. Version 2 had no subheadings, used more words to describe each category, and provided a category for "did not work for pay in the past 5 years." Additionally, the paper version tested two different introductions to the Class of Worker, Industry, and Occupation item series. Version 1 had a very specific instruction to "describe clearly this person's main job activity or business last week." Version 2 had a more general instruction and introduced the series of questions as being about "the type of business this person worked for and the type of work this person did." Moreover, two Computer-Assisted Interviewing (CAI) modes of the question were tested. The main difference between the two versions was whether the Active Duty category mentioned "Commissioned Corps Service" and whether there was a follow-up item asking for branch of the military.

The recommendation going into the second round was to use the introduction, instructions, and question stem from Version 2 and the response categories from Version 1 (including the Version 1 formatting for the mail mode). Most respondents preferred the version with the subheadings and more concise wording. It was also recommended that the emphasis in the introduction be on the job held "last week." In the CAI modes, we recommended asking respondents about military branch separate from the name of their employer or business, which is consistent with the approach used in Version 1 of the CAI modes in the first round of testing.

Suggested probes for cognitive testing included:
1. Were the introduction and instructions clear to the respondent, especially the time referenced in the question?
2. Did it take respondents longer to read and process this question because of the additional headers (i.e., "PRIVATE SECTOR employee:")?
3. Did respondents understand they only select one box and not one box from each category?
4. Did bolding the Class of Worker categories play the most important role in helping respondents make a selection?
5. What were respondents considering as Active Duty military status?
6. Did respondents understand the difference between for-profit and non-profit?
7. Was the distinction between Local and State Government Worker clear?
8. What were respondents counting as Unpaid Family Work in this category?

The subcommittee recommendation was to retain the wording from the second round of cognitive testing for the Content Test treatment (see Section 1.3 for the question wording). The revised classification of the Class of Worker item tested well, overall, in both modes and both languages. Most respondents appeared to interpret the questions as intended. There was no underlying pattern among those who did not answer accurately (Steiger, Anderson, Folz, Leonard, & Stapleton, 2015).

1.3   Question Content

The Class of Worker question was tested to clarify the intent of the question and the response categories while also improving the question layout. The control treatment of the Class of Worker question in the Content Test used the same format and procedures as all modes of data collection in the ACS production panel for March 2016. The test treatment incorporated the changes described here:

Changes to the Paper Questionnaire:

The Class of Worker response categories were grouped under three general headings rather than listed in one group with no headings (See Figures 1 and 2). The three headings were "PRIVATE SECTOR EMPLOYEE," "GOVERNMENT EMPLOYEE," and "SELF-EMPLOYED OR OTHER." Additionally, Active Duty was added as one of the response categories in the Government section, and the Active Duty checkbox was dropped from the Employer Name question. Question and response category wording was revised for clarity in both the Class of Worker question and the Employer Name question. Also, to signal that all six employment characteristics questions refer to the same job (including Industry and Occupation), the series was renumbered from separate questions (42 through 47) to a single series with subquestions (41a through f). Finally, the instructional text and heading for the series immediately preceding Class of Worker was simplified.

Figure 1: Mail Control Version of the Class of Worker Question

Figure 2: Mail Test Version of the Class of Worker Question

M *Answer questions 42 – 47 if this person worked in the past 5 years. Otherwise, SKIP to question 48.*

42 – 47 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.*

42 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?

43 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

_____

M *Answer questions 41a – f if this person worked in the past 5 years. Otherwise, SKIP to question 42.*

41 DESCRIPTION OF EMPLOYMENT

The next series of questions is about the type of employment this person had last week.

If this person had more than one job, describe the one at which the most hours were worked. If this person did not work last week, describe the most recent employment in the past five years.

a. Which one of the following best describes this person's employment last week or the most recent employment in the past 5 years?
Mark (X) ONE box.

PRIVATE SECTOR EMPLOYEE

☐ **For-profit** company or organization

☐ **Non-profit** organization (including tax-exempt and charitable organizations)

GOVERNMENT EMPLOYEE

☐ **Local government** (for example: city or county school district)

☐ **State government** (including state colleges/universities)

☐ **Active duty** U.S. Armed Forces or Commissioned Corps

☐ **Federal government** civilian employee

SELF-EMPLOYED OR OTHER

☐ **Owner of non-incorporated** business, professional practice, or farm

☐ **Owner of incorporated** business, professional practice, or farm

☐ Worked **without pay** in a **for-profit** family business or farm for 15 hours or more per week

b. What was the name of this person's employer, business, agency, or branch of the Armed Forces?

_____

Changes to the Internet Instrument:

In the current ACS internet instrument, Active Duty U.S. Armed Forces member is included as one of the response options for the initial Class of Worker question. In contrast, in the current paper questionnaire, an Active Duty member of the Armed Forces should first check the Federal Government option and then check the Active Duty checkbox that accompanies the Employer Name question. Otherwise, the current internet version of the Class of Worker question appears similar to the mail question. The test version of the question for the internet instrument groups the response options and gives each group a heading, such as in the mail mode. In addition, as with the mail form question and response category, the wording was revised for clarity in both the Class of Worker question and the Employer Name question, the question series was renumbered as a single question with subparts, and the instructional text reworded for clarity.

The control and test versions for the internet instrument differ in essentially the same way the paper questionnaires do. The test version has distinct groupings and headings, while the control version does not. However, in both the control and test versions of the internet instrument, Active Duty is included as a response option in the first part of the question. Furthermore, in both treatments, what the respondent sees for the Employer Name question depends on whether they checked the Active Duty option in the Class of Worker question. If they did not check this option, they see "What was the name of <(Name)'s,/your> company, business, or other employer?" If they did check the Active Duty option, they see "Which branch of the Armed Forces <does/do> <(Name)/you> work for?" followed by five checkboxes for control – U.S. Army, U.S. Navy, U.S. Air Force, U.S. Marines, and U.S. Coast Guard, and two additional categories for test – U.S. Public Health Service and National Oceanic and Atmospheric Administration. (For the internet question format, see Appendix A.)

Changes to the Computer-Assisted Telephone Instrument (CATI):

In CATI, Class of Worker is asked using an unfolding structure in both control and test, with the first question having five general categories (Private, Government, Active Duty, Self-Employed, and Unpaid Family Worker). The test version now includes a follow-up question for the Worked Without Pay in a for-Profit Family Business or Farm category, which asks the respondent "Did <(Name)/you> work without pay in this for-profit family business or farm for 15 hours or more per week?" If the response to this question is "No," then the person is not asked the rest of the Industry and Occupation series. The general question is then followed by a question on the specific type of category, except for Active Duty, which is followed by a version of Employer Name question. As with the internet mode, the test instrument has two more branches of Armed Forces than control. Additionally, the test version utilizes a fill in the instructional text, based on the response to when the person last worked, which is expected to help shorten the introduction for interviewers and decrease reference period confusion for respondents. Like the mail mode, the instructional text is also modified for simplicity. Finally, persons for whom the respondent answers the first, general part of the question, but does not answer the second, more specific part are classified as "unspecified" within the general category. (For the CATI question format, see Appendix B.)

Changes to the CAPI Instrument:

In the test version of the CAPI instrument, the Class of Worker response categories on the flashcard shown to respondents were worded and grouped with headings the same way as in the test version of the mail and internet modes. In the control version, these categories on the CAPI flashcard are not worded the same way as any of the other control version instruments as it contains a category for the U.S. Armed Forces (see Figure B-2 in Appendix B). As with the internet and CATI modes, the test instrument has two more branches of Armed Forces than control. Additionally, like CATI, the test version utilizes a fill in the instructional text, based on the response to when the person last worked, which is expected to help shorten the introduction for interviewers and decrease reference period confusion for respondents. The instructional text is modified for simplicity. (For the CAPI question format, see Appendix B.)

1.4   Research Questions

The following research questions were formulated to guide the analyses of the Class of Worker question. These analyses assessed how the test version of the questions performed compared to the control version in the following ways: how often the respondents answered the question, the consistency and accuracy of the responses, and how the responses affected the resulting estimates.

This research addresses the following questions:

1. *Are the control and test missing data rates the same? We also answer this question by data collection mode – internet, mail, CATI, and CAPI.*

2. *Are the control and test response distributions consistent? If not, for which response categories are the differences in the control and test proportion estimates statistically significant? We also answer this question by data collection mode – internet, mail, CATI, and CAPI.*

3. *How do the control and test estimates of Unpaid Family Workers compare with the 2015 Current Population Survey (CPS) estimates?*

4. *How do the control and test estimates of the three employee categories (Federal, State, and Local) compare with estimates from the 2015 Annual Survey of Public Employment and Payroll (ASPEP)?*

5. *How do the control and test estimates of Self-Employed (Incorporated and Non-incorporated) Workers compare with estimates from the 2012 Survey of Business Owners (SBO)?*

6. *Is response reliability the same for the control and test versions of the Class of Worker question? We also answer this question by data collection mode – internet, mail, CATI, and CAPI.*

7. *In the paper (mail) mode, are there more multiple responses to the Class of Worker question in test than in control?*

8. *How often is the coding output value of Class of Worker changed by coding clerks from the coding input value?[10]*

9. *In the telephone interviewing (CATI) mode, are the control and test rates of "unspecified" responses the same?*

10. *Is the reporting of self-employment income and wages income consistent with the Class of Worker response?*

11. *Do those persons reported as Unpaid Family Workers work at least 15 hours per week? How often is this not the case?*

2 METHODOLOGY

2.1 Sample Design

The 2016 ACS Content Test consisted of a nationally representative sample of 70,000 residential addresses in the United States, independent of the production ACS sample. The Content Test sample universe did not include GQs, nor did it include housing units in Alaska, Hawaii, or Puerto Rico.[11] The sample design for the Content Test was largely based on the ACS production sample design with some modifications to better meet the test objectives.[12] The modifications included adding an additional level of stratification by stratifying addresses into high and low self-response areas, oversampling addresses from low self-response areas to ensure equal response from both strata, and sampling units as pairs.[13] The high and low self-response strata were defined based on ACS self-response rates at the tract level. Sampled pairs were formed by first systematically sampling an address within the defined sampling stratum and then pairing that address with the address listed next in the geographically sorted list. Note that the pair was likely not neighboring addresses. One member of the pair was randomly assigned to receive the control version of the question and the other member was assigned to receive the test version of the question, thus resulting in a sample of 35,000 control cases and 35,000 test cases.

---

[10] An error in creation of the coding input used during the coding process resulted in clerical coders seeing a different Class of Worker category than the original respondent selected category. As a result, we could not conduct an analysis to address research question 8.

[11] Alaska and Hawaii were excluded for cost reasons. GQs and Puerto Rico were excluded because the sample sizes required to produce reliable estimates would be overly large and burdensome, as well as costly.

[12] The ACS production sample design is described in Chapter 4 of the ACS Design and Methodology report (U.S. Census Bureau, 2014).

[13] Tracts with the highest response rate based on data from the 2013 and 2014 ACS were assigned to the high response stratum in such a way that 75 percent of the housing units in the population (based on 2010 Census estimates) were in the high response areas; all other tracts were designated in the low response strata. Self-response rates were used as a proxy for overall cooperation. Oversampling in low response areas helps to mitigate larger variances due to CAPI subsampling. This stratification at the tract level was successfully used in previous ACS Content Tests, as well as the ACS Voluntary Test in 2003.

As in the production ACS, if efforts to obtain a response by mail or telephone were unsuccessful, attempts were made to interview in person a sample of the remaining nonresponding addresses (see Section 2.2 Data Collection for more details). Addresses were sampled at a rate of 1-in-3, with some exceptions that were sampled at a higher rate.[14] For the Content Test, the development of workload estimates for CATI and CAPI did not take into account the oversampling of low response areas. This oversampling resulted in a higher than expected workload for CATI and CAPI and therefore required more budget than was allocated. To address this issue, the CAPI sampling rate for the Content Test was adjusted to meet the budget constraint.

2.2 Data Collection

The field test occurred in parallel with the data collection activities for the March 2016 ACS production panel, using the same basic data collection protocol as production ACS with a few differences as noted below. The data collection protocol consisted of three main data collection operations: 1) a six-week mailout period, during which the majority of internet and mailback responses were received; 2) a one-month CATI period for nonresponse follow-up; and 3) a one-month CAPI period for a sample of the remaining nonresponse. Internet and mailback responses were accepted until three days after the end of the CAPI month.

As indicated earlier, housing units included in the Content Test sample were randomly assigned to a control or test version of the questions. CATI interviewers were not assigned specific cases; rather, they worked the next available case to be called and therefore conducted interviews for both control and test cases. CAPI interviewers were assigned Content Test cases based on their geographic proximity to the cases and therefore could also conduct both control and test cases.

The ACS Content Test's data collection protocol differed from the production ACS in a few significant ways. The Content Test analysis did not include data collection via the Telephone Questionnaire Assistance (TQA) program since those who responded via TQA used the ACS production TQA instrument. The Content Test excluded the telephone Failed Edit Follow-Up (FEFU) operation.[15] Furthermore, the Content Test had an additional telephone reinterview operation used to measure response reliability. We refer to this telephone reinterview component as the Content Follow-Up, or CFU. The CFU is described in more detail in Section 2.3.

---

[14] The ACS production sample design for CAPI follow-up is described in Chapter 4, Section 4.4 of the ACS Design and Methodology report (U.S. Census Bureau, 2014).

[15] In ACS production, paper questionnaires with an indication that there are more than five people in the household or questions about the number of people in the household, and self-response returns that are identified as being vacant or a business or lacking minimal data are included in FEFU. FEFU interviewers call these households to obtain any information the respondent did not provide.

ACS production provides Spanish-language versions of the internet, CATI, and CAPI instruments, and callers to the TQA number can request to respond in Spanish, Russian, Vietnamese, Korean, or Chinese. The Content Test had Spanish-language automated instruments; however, there were no paper versions of the Content Test questionnaires in Spanish.[16] Any case in the Content Test sample that completed a Spanish-language internet, CATI, or CAPI response was included in analysis. However, if a case sampled for the Content Test called TQA to complete an interview in Spanish or any other language, the production interview was conducted and the response was excluded from the Content Test analysis. This was due to the low volume of non-English language cases and the operational complexity of translating and implementing several language instruments for the Content Test. CFU interviews for the Content Test were conducted in either Spanish or English. The practical need to limit the language response options for Content Test respondents is a limitation to the research, as some respondents self-selected out of the test.

2.3   Content Follow-Up

For housing units that completed the original interview, a CFU telephone reinterview was also conducted to measure response error.[17] A comparison of the original interview responses and the CFU reinterview responses was used to answer research questions about response error and response reliability.

A CFU reinterview was attempted with every household that completed an original interview for which there was a telephone number. A reinterview was conducted no sooner than two weeks (14 calendar days) after the original interview. Once the case was sent to CFU, it was to be completed within three weeks. This timing balanced two competing interests: (1) conducting the reinterview as soon as possible after the original interview to minimize changes in truth between the two interviews, and (2) not making the two interviews so close together that the respondents were simply recalling their previous answers. Interviewers made two call attempts to interview the household member who originally responded, but if that was not possible, the CFU reinterview was conducted with any other eligible household member (15 years or older).

---

[16] In the 2014 ACS, respondents requested 1,238 Spanish paper questionnaires, of which 769 were mailed back. From that information, we projected that fewer than 25 Spanish questionnaires would be requested in the Content Test.

[17] Throughout this report, the "original interview" refers to responses completed via paper questionnaire, internet, CATI, or CAPI.

The CFU asked basic demographic questions and a subset of housing and detailed person questions that included all of the topics being tested, with the exception of Telephone Service, and any questions necessary for context and interview flow to set up the questions being tested.[18] All CFU questions were asked in the reinterview, regardless of whether or not a particular question was answered in the original interview. Because the CFU interview was conducted via telephone, the wording of the questions in CFU followed the same format as the CATI nonresponse interviews. Housing units assigned to the control version of the questions in the original interview were asked the control versions of the question in CFU; housing units assigned to the test version of the questions in the original interview were asked the test version of the questions in CFU. The only exception was for retirement, survivor, and disability income, for which a different set of questions was asked in CFU.[19]

2.4 Analysis Metrics

This section describes the metrics used to assess the revised version of the question, which includes the item missing data rate, response distributions, comparisons to benchmarks, response error, and other metrics. This section also describes the methodology used to calculate unit response rates and standard errors for the test.

All Content Test data were analyzed without imputation due to our interest in how question changes or differences between versions of new questions affected "raw" responses, not the final edited variables. Some editing of responses was done for analysis purposes, such as collapsing response categories or modes together or calculating a person's age based on his or her date of birth.

All estimates from the ACS Content Test were weighted. Analysis involving data from the original interviews used the final weights that take into account the initial probability of selection (the base weight) and CAPI subsampling. For analysis involving data from the CFU interviews, the final weights were adjusted for CFU nonresponse to create CFU final weights.

The significance level for all hypothesis tests is $\alpha = 0.1$. When conducting numerous comparisons between the control and test treatments, there is a concern about incorrectly rejecting a hypothesis that is actually true (a "false positive" or Type I error). The overall Type I error rate is called the familywise error rate and is the probability of making one or more Type I errors among all hypotheses tested simultaneously. When adjusting for multiple comparisons, the Holm-Bonferroni method was used (Holm, 1979).

---

[18] Because the CFU interview was conducted via telephone the Telephone Service question was not asked. We assume that CFU respondents have telephone service.

[19] Refer to the 2016 ACS Content Test report on Retirement Income for a discussion on CFU questions for survivor, disability, and retirement income.

2.4.1   Unit Response Rates and Demographic Profile of Responding Households

The unit response rate is generally defined as the proportion of sample addresses eligible to respond that provided a complete or sufficient partial response. Unit response rates from the original interview are an important measure to look at when considering the analyses in this report that compare responses between the control and test versions of the survey questionnaire. We expected that the unit response rates for both treatments would be similar and we compared them to verify this assumption.

For both control and test treatments, we calculated the overall unit response rate (all modes of data collection combined) and unit response rates by mode: internet, mail, CATI, and CAPI. We also calculated the total self-response rate by combining internet and mail modes together. Some Content Test analyses focused on the different data collection modes for topic-specific evaluations, thus we felt it was important to include each mode in the response rates section. In addition to those rates, we calculated the response rates for high and low response areas because analysis for some Content Test topics was done by high and low response areas. Using the Census Bureau's Planning Database (U.S. Census Bureau, 2016a), we defined these areas at the tract level based on the low response score.

The universe for the overall unit response rates consists of all addresses in the initial sample (70,000 addresses) that were eligible to respond to the survey. Some examples of addresses ineligible for the survey were a demolished home, a home under construction, a house or trailer that was relocated, or an address determined to be a permanent business or storage facility. The universe for self-response (internet and mail) rates consists of all mailable addresses that were eligible to respond to the survey. The universe for the CATI response rate consists of all nonrespondents at the end of the mailout month from the initial survey sample that were eligible to respond to the survey and for whom we possessed a telephone number. The universe for the CAPI response rates consists of a subsample of all remaining nonrespondents (after CATI) from the initial sample that were eligible to respond to the survey. Any nonresponding addresses that were sampled out of CAPI were not included in any of the response rate calculations.

We also calculated the CFU interview unit response rate overall and by mode of data collection of the original interview and compared the control and test treatments because response error analysis (discussed in Section 2.4.5.) relies upon CFU interview data. Statistical differences between CFU response rates for control and test treatments will not be taken as evidence that one version is better than the other. For the CFU response rates, the universe for each mode consists of housing units that responded to the original questionnaire in the given mode (internet, mail, CATI, or CAPI) and were eligible for the CFU interview. We expected the response rates to be similar between treatments; however, we calculated the rates to verify that assumption.

Another important measure to look at in comparing experimental treatments is the demographic profile of the responding households in each treatment. The Content Test sample was designed with the intention of having respondents in both control and test treatments exhibit similar distributions of socioeconomic and demographic characteristics. Similar distributions allow us to compare the treatments and conclude that any differences are due to the experimental treatment instead of underlying demographic differences. Thus, we analyzed distributions for data from the following response categories: *age, sex, educational attainment,* and *tenure.* The topics of *race, Hispanic origin,* and *relationship* are also typically used for demographic analysis; however, those questions were modified as part of the Content Test, so we could not include them in the demographic profile. Additionally, we calculated *average household size* and the *language of response* for the original interview.[20]

For response distributions, we used chi-square tests of independence to determine statistical differences between control and test treatments. If the distributions were significantly different, we performed additional testing on the differences for each response category. To control for the overall Type I error rate for a set of hypotheses tested simultaneously, we performed multiple-comparison procedures with the Holm-Bonferroni method (Holm, 1979). A family for our response distribution analysis was the set of p-values for the overall characteristic categories (*age, sex, educational attainment,* and *tenure*) and the set of p-values for a characteristic's response categories if the response distributions were found to have statistically significant differences. To determine statistical differences for *average household size* and the *language of response* of the original interview we performed two-tailed hypothesis tests.

For all response-related calculations mentioned in this section, addresses that were either sampled out of the CAPI data collection operation or that were deemed ineligible for the survey were not included in any of the universes for calculations. Unmailable addresses were also excluded from the self-response universe. For all unit response rate estimates, differences, and demographic response analysis, we used replicate base weights adjusted for CAPI sampling (but not adjusted for CFU nonresponse).

2.4.2   Item Missing Data Rates

Respondents leave items blank for a variety of reasons including not understanding the question (clarity), their unwillingness to answer a question as presented (sensitivity), and their lack of knowledge of the data needed to answer the question. The item missing data rate (for a given item) is the proportion of eligible units, housing units for household-level items or persons for person-level items, for which a required response (based on skip patterns) is missing.

---

[20] Language of response analysis excludes paper questionnaire returns because there was only an English questionnaire.

The percent of eligible persons who did not provide a valid response to Class of Worker in the control treatment was compared with the corresponding percent from the test treatment. For the Class of Worker question, cases with multiple marks (only possible in the mail mode) were treated as missing, as were responses of "Don't Know" or "Refused." In the control treatment, if the Active Duty checkbox was selected in the Employer Name question without a response to the main Class of Worker question, this was treated as a response for this analysis, in order to make the control and test categories comparable.

We tested the statistical significance of differences between the control and test rates using two-tailed t-tests.

2.4.3   Response Distributions

Comparing the response distributions between the control version of a question and the test version of a question allowed us to assess whether the question change affected the resulting estimates. Comparisons were made using Rao-Scott chi-squared tests (Rao & Scott, 1987) for distribution and two-tailed t-tests for single categories when the corresponding distributions were found to be statistically different.

Proportion estimates were calculated as:

$$\text{Category proportion} = \frac{\text{weighted count of valid responses in category}}{\text{weighted count of all valid responses}}$$

2.4.4   Benchmarks

For the Class of Worker question, we compared estimates from both control and test treatments to data from the 2015 Current Population Survey Annual Social and Economic Supplement (CPS ASEC).[21] This comparison allowed us to tell whether our results differed from another reliable resource.[22]

---

[21] For more information on the 2015 CPS Annual Social and Economic Supplement (ASEC), see https://www2.census.gov/programs-surveys/cps/techdocs/cpsmar15.pdf.

[22] Initially, it was our plan to compare estimates of the 2015 Annual Survey of Public Employment and Payroll (ASPEP) and the 2012 Survey of Business Owners (SBO) with the control and test treatments. However, due to methodological differences later discovered between the ACS Content Test and these surveys, we did not compare our results with the ASPEP or SBO. For more information on the 2015 Annual Survey of Public Employment and Payroll (ASPEP) http://www.census.gov/govs/apes/. For more information on the 2012 Survey of Business Owners, see http://census.gov/programs-surveys/sbo/technical-documentation/methodology.html.

The national distributions of Unpaid Family Workers from the control and test treatments were compared with the corresponding estimates from the 2015 CPS ASEC. Note that for the benchmark estimates, state-level estimates for Alaska and Hawaii were subtracted from the national estimates in order to be comparable with the ACS Content Test data, which we did not sample in these two states. The CPS reference period is fixed, and is typically the week including the 12th of the month, with interviews being conducted the following week (typically the week including the 19th of the month). ACS responses are collected at times that vary throughout the month and year.

The ACS data collection methodology is substantially different from the CPS ASEC, which is conducted by CATI or CAPI. Additionally, the ACS is mandatory and therefore response at the unit and item level is higher in the ACS than in the CPS ASEC (U.S. Census Bureau, 2016b). The ACS samples 3.5 million addresses each year; around 290,000 addresses each month. The CPS ASEC annual sample size is about 150,000 housing units. About 54,000 of those housing units result in interviews and contain approximately 106,000 persons 15 years old and over. The ACS universe includes people living in housing units and in institutionalized and non-institutionalized GQs. The CPS universe is the civilian noninstitutionalized population of the United States living in housing units and members of the Armed Forces living in civilian housing units on a military base or in a household not on a military base. The Armed Forces members, however, are not asked the monthly labor force questions, while in the ACS they are asked these questions. In addition, the CPS ASEC is supplemented with a sample of Hispanic households identified the previous November. This results in the addition of about 6,500 households (5,500 interviewed).

2.4.5   Response Error

Response error occurs for a variety of reasons, such as flaws in the survey design, misunderstanding of the questions, misreporting by respondents, or interviewer effects. There are two components of response error: response bias and simple response variance. Response bias is the degree to which respondents consistently answer a question incorrectly. Simple response variance is the degree to which respondents answer a question inconsistently. A question has good response reliability if respondents tend to answer the question consistently. Re-asking the same question of the same respondent (or housing unit) allows us to measure response variance.

We measured simple response variance by comparing valid responses to the CFU reinterview with valid responses to the corresponding original interview.[23] The Census Bureau has frequently used content reinterview surveys to measure simple response variance for large demographic data collection efforts, including the 2010 ACS Content Test, and the 1990, 2000, and 2010 decennial censuses (Dusch & Meier, 2012).

---

[23] A majority of the CFU interviews were conducted with the same respondent as the original interview (see the Limitations section for more information).

The following measures were used to evaluate consistency:

- Gross difference rate (GDR)
- Index of inconsistency (IOI)
- L-fold index of inconsistency (IOI$_L$)

The first two measures – GDR and IOI were calculated for individual response categories. The L-fold index of inconsistency was calculated for questions that had three or more mutually exclusive response categories, as a measure of overall reliability for the question.

The GDR, and subsequently the simple response variance, are calculated using the following table and formula.

Table 1: Interview and Reinterview Counts for Each Response Category Used for Calculating the Gross Difference Rate and Index of Inconsistency

	Original Interview "Yes"	Original Interview "No"	Reinterview Totals
CFU Reinterview "Yes"	a	b	**a + b**
CFU Reinterview "No"	c	d	**c + d**
Original Interview Totals	**a + c**	**b + d**	**n**

Where a, b, c, d, and n are defined as follows:

a = weighted count of units in the category of interest for both the original interview and reinterview
b = weighted count of units NOT in the category of interest for the original interview, but in the category for the reinterview
c = weighted count of units in the category of interest for the original interview, but NOT in the category for the reinterview
d = weighted count of units NOT in the category of interest for either the original interview or the reinterview
n = total units in the universe = a + b + c + d.

The GDR for a specific response category is the percent of inconsistent answers between the original interview and the reinterview (CFU). We calculate the GDR for a response category as

$$GDR = \frac{(b + c)}{n} \times 100$$

Statistical significance between the GDR for a specific response category between the control and test treatments is determined using a two-tailed t-test.

In order to define the IOI, we must first discuss the variance of a category proportion estimate. If we are interested in the true proportion of a total population that is in a certain category, we can use the proportion of a survey sample in that category as an estimate. Under certain reasonable assumptions, it can be shown that the total variance of this proportion estimate is the sum of two components, sampling variance (SV) and simple response variance (SRV). It can also be shown that an unbiased estimate of SRV is half of the GDR for the category (Flanagan, 1996).

SV is the part of total variance resulting from the differences among all the possible samples of size *n* one might have selected. SRV is the part of total variance resulting from the aggregation of response error across all sample units. If the responses for all sample units were perfectly consistent, then SRV would be zero, and the total variance would be due entirely to SV. As the name suggests, the IOI is a measure of how much of the total variance is due to inconsistency in responses, as measured by SRV and is calculated as:

$$IOI = \frac{n(b + c)}{(a + c)(c + d) + (a + b)(b + d)} \times 100$$

Per the Census Bureau's general rule, index values of less than 20 percent indicate low inconsistency, 20 to 50 percent indicate moderate inconsistency, and over 50 percent indicate high inconsistency.

An IOI is computed for each response category and an overall index of inconsistency, called the L-fold index of inconsistency, is reported for the entire distribution. The L-fold index is a weighted average of the individual indexes computed for each response category.

When the sample size is small, the reliability estimates are unstable. Therefore, we do not report the IOI and GDR values for categories with a small sample size, as determined by the following formulas: $2a + b + c < 40$ or $2d + b + c < 40$, where a, b, c, and d are unweighted counts as shown in Table 1 above (see Flanagan 1996, p. 15).

The measures of response error assume that those characteristics in question did not change between the original interview and the CFU interview. To the extent that this assumption is incorrect, we assume that it is incorrect at similar rates between the control and test treatments.

A limitation on assessing the reliability of the Class of Worker question is that the time frame of the question is "last week." The time frame for the original response will therefore always be different from the time frame for the CFU response. This could reasonably lead to a different answer between responses. We assume however that any inconsistency in responses would occur at the same rate in the control version as in the test version.

2.4.6 Analysis Specific to Class of Worker

Analyses specific to Class of Worker include comparisons between the test and control treatment of multiple responses in the mail mode, evaluating "unspecified" responses in the CATI mode, changes to coding output values made by coding clerks, consistency with write-in responses to the industry question, and consistency checks regarding income and wages, and hours worked (for Unpaid Family Workers).

Multiples Responses:
The Class of Worker question asks respondents to mark only one box, but some respondents in the mail mode incorrectly mark more than box. Automated questionnaires mitigate this error by requiring that only one box be selected. We compared the rate at which eligible mail mode respondents marked multiple response categories between control and test. Since the Active Duty checkbox is part of the Employer Name question and therefore a separate item on the control questionnaire, we did not count it as a multiple response when there were two responses and one was the Active Duty checkbox. In the test questionnaire, if a respondent marked Active Duty and another category, it counted as a multiple response.[24] We used a two-tailed t-test to check the significance of any difference between control and test.

Unspecified Category:
In CATI, Class of Worker has an unfolding structure with the first question having five general categories (Private, Government, Active Duty, Self-Employed, Unpaid Family Worker). The general question is followed by a question on the specific type of work within that category, except for Active Duty, which is followed by a version of the Employer Name question. The 'unspecified' rate is the rate of those answering the first Class of Worker question (Private, Government, Self-Employed, or Unpaid Family Worker) but not the Class of Worker follow-up question. We evaluate the unspecified rate as an additional check on the quality of the data. The higher the unspecified rate, the lower the quality of the response. Rates were calculated separately for each category.

Changes to Coding Output:
Clerical coders may change the Class of Worker response during the process of coding the Industry and Occupation responses. This could potentially result in a difference between the coding input and output for unedited Class of Worker data. Clerks change Class of Worker to correct the data or to change the value from an unspecified to a specified Class of Worker category. We intended to examine the percentage of valid Class of Worker responses that were changed, as well as the rate of unspecified Class of Worker responses after coding as another check of the data quality. However, this analysis was not performed due to an error in the creation of the input coding file used by the clerical coders.

---

[24] In other analyses performed for this topic, we treated the Active Duty checkbox on the Employer Name question of the control version as the equivalent to the Active Duty category in the test version of the Class of Worker question to make the control and test response categories mostly comparable.

Consistency Checks – Income:
The final two research questions check that the test version of Class of Worker is collecting consistent, "correct" content.[25] We tested the differences between the control and test content consistency rates using two-tailed t-tests. The first consistency check involves whether or not self-employment income and salary wage income are consistent with Class of Worker. Self-Employed Not Incorporated Workers are expected to have non-zero self-employment income. Unpaid Family Workers are expected to have no earned income. All other workers (Private sector, Government, Self-Employed Incorporated) are expected to have non-zero wage and salary income. All three groups could have non-zero income from other sources due to the difference in reference jobs and earned income sources. In this comparison, blanks were considered the same as zero income. Only respondents who had worked in the past 12 months were included in this check.

It is important to note, however, that income data is collected for all jobs worked in the past 12 months, while Class of Worker data are collected for the primary job last week. If the person did not work last week, the Class of Worker data are collected for the last job the respondent had in the last 5 years. For example, if a respondent had two jobs, we collect information for only one job, the primary job, from "last week." For the income questions, the respondents are asked to provide self-employment income and/or salary wage income for all jobs this person worked in the last 12 months. While the reference periods between income and the job reported are not exactly the same, this is a useful check since post-coding editing processes handle inconsistencies using assumptions between income and job reported. We tested the statistical significance of differences between the control and test using two-tailed t-tests.

Consistency Check – Unpaid Family Workers:
The second content consistency check is for Unpaid Family Workers. This is particularly important because Unpaid Family Workers are the only exception to the definition of employment as paid work. Unpaid work-like activities are otherwise out of universe. Unpaid Family Workers are defined as people who worked without pay for 15 hours or more per week on a farm or other for-profit business operated by a relative. Thus, we examined the percentage of Unpaid Family Workers whose usual hours worked per week is 15 hours or more, as well as the percentage who appear to meet the full definition (i.e., work 15 hours or more per week and have zero wage or self-employment income). Additionally, in the unfolding structure of the CATI version, people recorded as an Unpaid Family Worker in the first Class of Worker question but do not work 15 or more hours per week are considered out of universe and not asked the rest of the Industry and Occupation series. We examined the percentage of cases in this scenario, but removed them from the universe for the rest of the analyses. We tested the statistical significance of differences between the control and test using two-tailed t-tests.

---

[25] Assuming people are reporting Class of Worker, Industry and Occupation, and Labor Force data for the same job or at least equally misreporting for two jobs on test and control.

2.4.7   Standard Error Calculations

We estimated the variances of the estimates using the Successive Differences Replication (SDR) method with replicate weights, the standard method used in the ACS (see U.S. Census Bureau, 2014a, Chapter 12). We calculated the variance for each rate and difference using the formula below.

The standard error of the estimate (X_0) is the square root of the variance:

$$\text{Var}(X_0) = \frac{4}{80} \sum_{r=1}^{80} (X_r - X_0)^2$$

where:
  X_0 = the estimate calculated using the full sample,
  X_r = the estimate calculated for replicate r.

3   DECISION CRITERIA FOR CLASS OF WORKER

Before fielding the 2016 ACS Content Test, we identified which of the metrics would be given higher importance in determining which version of the question would be recommended for inclusion in the ACS moving forward. The following table identifies the research questions and associated metrics in priority order.

Table 2. Decision Criteria

Research Questions	Decision Criteria, in order of priority[26]
4	The item missing data rate for the test version should be the same or lower than the control version, for all modes.
6	The response reliability should be the same or higher for the test version as on the control version, for all modes.
10 and 11	The proportion of people with consistent Class of Worker and Income and Usual Hours Worked per Week should be the same or higher for test than for control.
7	The proportion of people with multiple responses (mail mode only) should be the same or lower on test than on control.
8 and 9	The rate that clerks change the Class of Worker coding input value should be the same or lower for test than control. The proportion of "unspecified" responses (CATI mode only) should be the same or lower for test than control.
1, 2, 3, and 5	The distributions between the test and control versions should have minimal to no differences. In addition, differences in estimates between the test and benchmark data sources should be the same or lower than differences in estimates between the control and benchmark data sources.

4   LIMITATIONS

CATI and CAPI interviewers were assigned control and test treatment cases, as well as production cases. The potential risk of this approach is the introduction of a cross-contamination or carry-over effect due to the same interviewer administering multiple versions of the same question item. Interviewers are trained to read the questions verbatim to minimize this risk, but there still exists the possibility that an interviewer may deviate from the scripted wording. This could potentially mask a treatment effect.

Interviews were only conducted in English and Spanish. Respondents who needed language assistance in another language were not able to participate in the test. Additionally, the 2016 ACS Content Test was not conducted in Alaska, Hawaii, or Puerto Rico. Any conclusions drawn from this test may not apply to these areas or populations.

For statistical analysis specific to the mail mode, there may be bias in the results because of unexplained response rate differences between the control and test treatments.

We were not able to conduct demographic analysis by relationship status, race, or ethnicity because these topics were tested as part of the Content Test.

---

[26] The following decision criterion was moved to the *2016 American Community Survey Content Test Evaluation Report: Industry and Occupation*: "The proportion of people with consistent Industry and Class of Worker should be the same or higher for test than for control."

The CFU reinterview was not conducted in the same mode of data collection for households that responded by internet, by mail, or by CAPI in the original interview since CFU interviews were only administered using a CATI mode of data collection. As a result, the data quality measures derived from the reinterview may include some bias due to the differences in mode of data collection.

To be eligible for a CFU reinterview, respondents needed to either provide a telephone number in the original Content Test interview or have a telephone number available to the Census Bureau through reverse address look up. As a result, 2,284 of the responding households (11.8 percent with a standard error of 0.2) from the original control interviews and 2,402 of the responding households (12.4 percent with a standard error of 0.2) from the original test interviews were not eligible for the CFU reinterview. The difference between the control and test treatments is statistically significant (p-value=0.06).

Although we reinterviewed the same person who responded in the original interview when possible, we interviewed a different member of the household in the CFU for 7.5 percent (standard error of 0.4) of the CFU cases for the control treatment and 8.4 percent (standard error of 0.5) of the CFU cases for the test treatment.[27] The difference between the test and control treatments is not statistically significant (p-value=0.3). This means that differences in results between the original interview and the CFU interview for these cases could be due in part to having different people answering the questions. However, those changes were not statistically significant between the control and test treatments and should not impact the conclusions drawn from the reinterview.

The Content Test does not include the production weighting adjustments for seasonal variations in ACS response patterns, nonresponse bias, and under-coverage bias. As a result, any estimates derived from the Content Test data do not provide the same level of inference as the production ACS and cannot be compared with the production estimates.

In developing initial workload estimates for CATI and CAPI, we did not take into account the fact that we oversampled low response areas as part of the Content Test sample design. Therefore, workload and budget estimates were too low. In order to stay within budget, the CAPI workload was subsampled more than originally planned. This caused an increase in the variances for the analysis metrics used.

An error in addressing and assembling the materials for the 2016 ACS Content Test caused some Content Test cases to be mailed production ACS questionnaires instead of Content Test questionnaires. There were 49 of these cases that returned completed questionnaires, and they were all from the test treatment. These cases were excluded from the analysis. Given the small number of cases affected by this error, there is very little effect on the results.

---

[27] This is based on comparing the first name of the respondent between the original interview and the CFU interview. Due to a data issue we were not able to use the full name to compare.

Questionnaire returns were expected to be processed and keyed within two weeks of receipt. Unfortunately, a check-in and keying backlog prevented this requirement from being met, thereby delaying eligible cases from being sent to CFU on a schedule similar to the other modes. Additionally, the control treatment questionnaires were processed more quickly in keying than the test treatment questionnaires resulting in a longer delay for test mail cases to be eligible for CFU. On average, it took 18 days for control cases to become eligible for CFU; it took 20 days for test cases. The difference is statistically significant. This has the potential to impact the response reliability results.

The Class of Worker question was recoded to make the control and test categories comparable. In addition to changes to category wording and question format, the test treatment also contains a specific category for Active Duty. The Active Duty category is not listed as an option of Class of Worker in the control treatment, rather it is asked as a checkbox under the Employer Name question of "For whom did this person work?" To make the treatments comparable, if the Active Duty checkbox was selected without a response to the main Class of Worker question, this was treated as the Class of Worker response for this analysis. When the Active Duty checkbox was selected along with a response to the main Class of Worker question was selected, the response to the main Class of Worker question was treated as the Class of Worker response for this analysis.

An error in the creation of the coding input file used during the coding process resulted in clerical coders seeing a different Class of Worker category than the original respondent selected category. As a result, an analysis could not be conducted to address research question 8: *How often is the coding output value of Class of Worker changed by coding clerks from the coding input value?* This error did not affect the other Class of Worker specific analyses. The other analyses used a separate Class of Worker variable, which was unaffected by the error.

There is a limitation with the benchmark data, which precluded us from performing this analysis. The data from ASPEP and SBO could only be obtained as point estimates and not a proportional distribution of all workers 15 years and older. The ASPEP only collects data on civilian employees of Federal Government agencies (except the Central Intelligence Agency, National Security Agency, and Defense Intelligence Agency), all agencies of the 50 state governments, and 90,690 local governments (i.e., counties, municipalities, townships, special districts, and school districts) including the District of Columbia (U.S. Census Bureau, 2015d). The SBO survey sample consists of companies or firms operating during the survey year with receipts of $1,000 or more and not classified in one of the following industries: Crop and Animal Production, Rail Transportation, Postal Service, Monetary Authorities-Central Bank, Funds, Trusts, and Other Financial Vehicles, Religious, Grant making, Civic, Professional, and Similar Organizations, Private Households, or Public Administration. Since the Content Test does not include the production weighting adjustment, a comparable point estimate could not be developed for statistical testing purposes (U.S. Census Bureau, 2015c).

5 RESEARCH QUESTIONS AND RESULTS

This section presents the results from the analyses of the 2016 ACS Content Test data for Class of Worker. An analysis of unit response rates is presented first followed by topic-specific analyses. For the topic-specific analyses, each research question is restated, followed by corresponding data and a brief summary of the results.

5.1 Unit Response Rates and Demographic Profile of Responding Households

This section presents results for unit response rates for both control and test treatments for the original Content Test interview and for the CFU interview. It also provides results of a comparison of socioeconomic and demographic characteristics of respondents in both control and test treatments.

5.1.1 Unit Response Rates for the Original Content Test Interview

The unit response rate is generally defined as the proportion of sample addresses eligible to respond that provided a complete or sufficient partial response. We did not expect the unit response rates to differ between treatments. This is important because the number of unit responses should also affect the number of item responses we receive for analyses done on specific questions on the survey. Similar item response universe sizes allow us to compare the treatments and conclude that any differences are due to the experimental treatment instead of differences in the populations sampled for each treatment.

Table 3 shows the unit response rates for the original interview for each mode of data collection (internet, mail, CATI, and CAPI), all modes combined, and both self-response modes (internet and mail combined) for the control and test treatments. When looking at the overall unit response rate (all modes combined) the difference between control (93.5 percent) and test (93.5 percent) is less than 0.1 percentage points and is not statistically significant.

Table 3. Original Interview Unit Response Rates for Control and Test Treatments, Overall and by Mode

Mode	Test Interviews	Test Percent	Control Interviews	Control Percent	Test minus Control	P-Value
All Modes	19,400	93.5 (0.3)	19,455	93.5 (0.3)	<0.1 (0.4)	0.98
Self-Response	13,131	52.9 (0.5)	13,284	53.7 (0.5)	-0.8 (0.6)	0.23
Internet	8,168	34.4 (0.4)	8,112	34.1 (0.4)	0.4 (0.6)	0.49
Mail	4,963	18.4 (0.3)	5,172	19.6 (0.3)	-1.2 (0.5)	0.01*
CATI	872	8.7 (0.4)	880	9.2 (0.4)	-0.4 (0.6)	0.44
CAPI	5,397	83.5 (0.7)	5,291	83.6 (0.6)	<0.1 (0.9)	0.96

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test at the α=0.1 level. The weighted response rates account for initial sample design as well as CAPI subsampling.

When analyzing the unit response rates by mode of data collection, the only modal comparison that shows a statistically significant difference is the mail response rate. The control treatment had a higher mail response (19.6 percent) than the test treatment (18.4 percent) by 1.2 percentage points. As a result of this difference, we looked at how mail responses differed in the high and low response areas. Table 4 shows the mail response rates for both treatments in high and low response areas.[28] The difference in mail response rates appears to be driven by the difference of rates in the high response areas.

It is possible that the difference in the mail response rates between control and test is related to the content changes made to the test questions. There are some test questions that could be perceived as being too sensitive by some respondents (such as the test question relating to same-sex relationships) and some test questions that could be perceived to be too burdensome by some respondents (such as the new race questions with added race categories). In the automated modes (internet, CATI, and CAPI) there is a higher likelihood of obtaining a sufficient partial response (obtaining enough information to be deemed a response for calculations before the respondent stops answering questions) than in the mail mode. If a respondent is offended by the questionnaire or feels that the questions are too burdensome, they may just throw the questionnaire away, and not respond by mail. This could be a possible explanation for the unit response rate being lower for test than control in the mail mode.

We note that differences between overall and total self-response response rates were not statistically significant. As most analysis was conducted at this level, we are confident the response rates were sufficient to conduct topic-specific comparisons between the control and test treatments and that there are no underlying response rate concerns that would impact those findings.

Table 4. Mail Response Rates by Designated High (HRA) and Low (LRA) Response Areas

	Test Interviews	Test Percent	Control Interviews	Control Percent	Test minus Control	P-Value
HRA	2,082	20.0 (0.4)	2,224	21.5 (0.4)	-1.5 (0.6)	0.02*
LRA	2,881	13.8 (0.3)	2,948	14.1 (0.3)	-0.3 (0.4)	0.43
Difference		6.2 (0.5)		7.4 (0.4)	-1.1 (0.7)	0.11

Source: U.S. Census Bureau, 2016 American Community Survey Content Test

Note: Minor additive discrepancies are due to rounding. Standard errors are in parentheses. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test at the α=0.1 level. The weighted response rates account for initial sample design as well as CAPI subsampling.

---

[28] Table C1 (including all modes) can be found in Appendix C.

5.1.2   Unit Response Rates for the Content Follow-Up Interview

Table 5 shows the unit response rates for the CFU interview by mode of data collection of the original interview and for all modes combined, for control and test treatments. Overall, the differences in CFU response rates between the treatments are not statistically. The rate at which CAPI respondents from the original interview responded to the CFU interview is lower for test (34.8 percent) than for control (37.7 percent) by 2.9 percentage points. While the protocols for conducting CAPI and CFU were the same between the test and control treatments, we could not account for personal interactions that occur in these modes between the respondent and interviewer. This can influence response rates. We do not believe that the difference suggests any underlying CFU response issues that would negatively affect topic specific response reliability analysis for comparing the two treatments.

Table 5. Content Follow-Up Interview Unit Response Rates for Control and Test Treatments, Overall and by Mode of Original Interview

Original Interview Mode	Test Interviews	Test Percent	Control Interviews	Control Percent	Test minus Control	P-Value
All Modes	7,867	44.8 (0.5)	7,903	45.7 (0.6)	-0.8 (0.8)	0.30
Internet	4,078	51.9 (0.6)	4,045	52.5 (0.7)	-0.6 (0.8)	0.49
Mail	2,202	46.4 (0.9)	2,197	44.2 (0.9)	2.1 (1.3)	0.11
CATI	369	48.9 (1.9)	399	51.5 (2.5)	-2.5 (2.9)	0.39
CAPI	1,218	34.8 (1.2)	1,262	37.7 (1.1)	-2.9 (1.6)	0.07*

Source: U.S. Census Bureau, 2016 American Community Survey Content Test

Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test at the α=0.1 level.

6   DEMOGRAPHIC AND SOCIOECONOMIC PROFILE OF RESPONDING HOUSEHOLDS

One of the underlying assumptions of our analyses in this report is that the sample for the Content Test was selected in such a way that responses from both treatments would be comparable. We did not expect the demographics of the responding households for control and test treatments to differ. To test this assumption, we calculated distributions for respondent data for the following response categories: *age, sex, educational attainment,* and *tenure.*[29] The response distribution calculations can be found in Table 6. Items with missing data were not included in the calculations. After adjusting for multiple comparisons, none of the differences in the categorical response distributions shown below is statistically significant.

---

[29] We were not able to conduct demographic analysis by relationship status, race, or ethnicity because these topics were tested as part of the Content Test.

Table 6. Response Distributions: Test versus Control Treatment

Item	Test Percent	Control Percent	Adjusted P-Value
AGE	(n=43,236)	(n=43,325)	0.34
Under 5 years old	5.7 (0.2)	6.1 (0.2)	
5 to 17 years old	17.8 (0.3)	17.6 (0.3)	
18 to 24 years old	8.6 (0.3)	8.1 (0.3)	
25 to 44 years old	25.1 (0.3)	26.2 (0.3)	
45 to 64 years old	26.8 (0.4)	26.6 (0.4)	
65 years old or older	16.0 (0.3)	15.4 (0.3)	
SEX	(n=43,374)	(n=43,456)	1.00
Male	48.8 (0.3)	49.1 (0.3)	
Female	51.2 (0.3)	50.9 (0.3)	
EDUCATIONAL ATTAINMENT#	(n=27,482)	(n=27,801)	1.00
No schooling completed	1.3 (0.1)	1.2 (0.1)	
Nursery to 11th grade	8.1 (0.3)	8.0 (0.3)	
12th grade (no diploma)	1.7 (0.1)	1.6 (0.1)	
High school diploma	21.7 (0.4)	22.3 (0.4)	
GED† or alternative credential	3.5 (0.2)	3.6 (0.2)	
Some college	21.0 (0.4)	20.2 (0.4)	
Associate's degree	8.8 (0.3)	9.1 (0.3)	
Bachelor's degree	20.9 (0.4)	20.3 (0.4)	
Advanced degree	13.1 (0.3)	13.7 (0.3)	
TENURE	(n=17,190)	(n=17,236)	1.00
Owned with a mortgage	43.1 (0.6)	43.2 (0.5)	
Owned free and clear	21.1 (0.4)	21.2 (0.4)	
Rented	33.8 (0.6)	34.0 (0.5)	
Occupied without payment of rent	1.9 (0.2)	1.7 (0.1)	

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
#For ages 25 and older
†General Educational Development
Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding.
Significance testing done at the α=0.1 level. P-values have been adjusted for multiple comparisons using the Holm-Bonferroni method.

We also analyzed two other demographic characteristics shown by the responses from the survey: *average household size* and *language of response.* The results for the remaining demographic analyses can be found in Table 7 and Table 8.

Table 7. Comparison of Average Household Size

	Test (n=17,608)	Control (n=17,694)	Test minus Control	P-value
Average Household Size (Number of People)	2.51 (<0.1)	2.52 (<0.1)	>-0.01 (<0.1)	0.76

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: Standard errors are shown in parentheses. Significance was tested based on a two-tailed t-test at the α=0.1 level.

Table 8. Comparison of Language of Response

Language of Response	Test Percent (n=17,608)	Control Percent (n=17,694)	Test minus Control	P-value
English	96.1 (0.2)	96.2 (0.2)	<0.1 (0.3)	0.52
Spanish	2.7 (0.2)	2.6 (0.2)	<0.1 (0.2)	0.39
Undetermined	1.2 (0.1)	1.2 (0.1)	<0.1 (0.2)	0.62

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. Significance was tested based on a two-tailed t-test at the α=0.1 level.

The Content Test was available in two languages, English and Spanish, for all modes except the mail mode. However, the language of response variable was missing for some responses, so we created a category called "undetermined" to account for those cases.

There are no detectable differences between control and test for *average household size* or *language of response*. There are also no detectable differences for any of the response distributions that we calculated. As a result of these analyses, it appears that respondents in both treatments do exhibit comparable demographic characteristics since none of the resulting findings is significant, which verifies our assumption of demographic similarity between treatments.

In the following sections, we present the results of the comparisons of test and control treatments conducted to evaluate the impact of the changes and modifications to the Class of Worker question. Unless otherwise stated, the comparisons between control and test treatments were based on a universe of people 15 years and older who worked in the last five years.

6.1   Item Missing Data Rates

This section addresses research question 1: *Are the control and test missing data rates the same?*

We expected the item missing data rates to be the same or lower for the test treatment. The results shown in Table 9 indicate that overall and for all modes of data collection, except mail, there were no significant differences between the item missing data rates for the control and test treatments. In the mail mode, the test treatment had a significantly higher item missing data rate (14.0 percent versus 10.4 percent).

If the respondent incorrectly answered the Class of Worker question with more than one response (i.e., multiple marks — only possible in the mail mode), we considered the response invalid and treated it as a nonresponse. Table 18 reveals that for the mail mode the test treatment had significantly more multiple mark responses than the control treatment (4.4 percent versus 0.6 percent, respectively). It is important to note that during the coding of the regular production cases, clerical coders assign a valid Class of Worker category to missing values when possible.

Table 9. Class of Worker Item Missing Data Rates

Mode	Test Sample Size	Test Percent	Control Sample Size	Control Percent	Test minus Control	P-Value
Overall	22,712	5.2 (0.2)	22,973	4.9 (0.2)	0.3 (0.3)	0.32
Internet	11,950	4.6 (0.3)	11,860	4.6 (0.3)	>-0.1 (0.4)	0.97
Mail	4,811	14.0 (0.6)	5,126	10.4 (0.5)	3.6 (0.8)	<0.01*
CATI	889	2.6 (0.6)	869	2.1 (0.6)	0.5 (0.9)	0.54
CAPI	5,062	2.0 (0.3)	5,118	2.7 (0.4)	-0.7 (0.5)	0.13

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test at the α=0.1 level.

6.2 Response Distributions

This section addresses research question 2: *Are the control and test response distributions consistent? If not, for which response categories are the differences in the control and test proportion estimates statistically significant?*

We compared the response distributions for the test and control treatments for the Class of Worker question using a Rao-Scott chi-squared test. Table 10 presents the results of the overall comparison and Tables 11, 12, and 13 present the results of the comparison by mode — internet, mail, and CATI/CAPI combined (to ensure sufficiently large cell sizes), respectively.

The p-value of 0.02 indicates that the test and control response distributions are not consistent. A comparison of the individual response categories found that the proportion of responses in the Private Not-For-Profit Worker category was significantly higher for the test treatment (9.0 percent versus 7.9 percent).

Table 10. Response Distributions: Overall

Category	Test Percent (n = 21,083)	Control Percent (n = 21,458)	Test minus Control	Adjusted P-Value
Private For-Profit Workers	65.7 (0.5)	66.4 (0.5)	-0.7 (0.7)	1.00
Private Not-For-Profit Workers	9.0 (0.3)	7.9 (0.3)	1.1 (0.4)	0.04*
Local Government Workers	7.3 (0.2)	6.8 (0.2)	0.5 (0.3)	1.00
State Government Workers	4.3 (0.2)	4.3 (0.2)	<0.1 (0.3)	1.00
Federal Government Workers	2.8 (0.2)	3.1 (0.2)	-0.3 (0.3)	1.00
Self-Employed Not Incorporated Workers	6.3 (0.3)	6.7 (0.2)	-0.4 (0.4)	1.00
Self-Employed Incorporated Workers	4.0 (0.2)	4.0 (0.2)	>-0.1 (0.3)	1.00
Unpaid Family Workers	0.4 (0.1)	0.5 (0.1)	-0.2 (0.1)	0.81
Unspecified Private Workers	0.2 (0.1)	0.1 (<0.1)	<0.1 (0.1)	1.00
Unspecified Government Workers	<0.1 (<0.1)	<0.1 (<0.1)	<0.1(<0.1)	1.00
Unspecified Self-Employed Workers	<0.1 (<0.1)	0.1 (<0.1)	>-0.1 (<0.1)	1.00
Total	100.0	100.0		

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: χ^2:21.3, p-value=0.02
Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test at the α=0.1 level. P-values have been adjusted for multiple comparisons using the Holm-Bonferroni method.

The internet (Table 11) and mail (Table 12) modes also show inconsistencies between the response distributions for the test and control treatments. In both modes, the proportion of Unpaid Family Workers is lower for the test treatment (0.3 percentage points lower for internet and 0.6 percentage points lower for mail). We suspect the lower proportion of Unpaid Family Workers is due to the added wording to clarify the definition of the category. The added wording clarifies to respondents that Unpaid Family Workers must work at least 15 hours or more in a for-profit business or farm. The lower estimate for Unpaid Family Workers in the test treatment is probably more accurate due to this clarification.

In the internet mode, the estimate for Local Government Workers for the test treatment was 1.4 percentage points higher than the control treatment. A possible reason for the higher percentage was the change to the examples listed next to the category (see Figure 1). The test treatment listed the example of "county school district," which might have made it clear that county school district workers should be counted as Local Government Workers.

Table 11. Response Distributions: Internet Mode

Category	Test Percent (n= 11,314)	Control Percent (n=11,202)	Test minus Control	Adjusted P-Value
Private For-Profit Workers	61.2 (0.5)	62.8 (0.5)	-1.6 (0.8)	0.27
Private Not-For-Profit Workers	10.3 (0.3)	10.3 (0.4)	<0.1 (0.5)	1.00
Local Government Workers	8.9 (0.3)	7.5 (0.3)	1.4 (0.4)	<0.01*
State Government Workers	4.7 (0.3)	5.2 (0.3)	-0.5 (0.4)	0.70
Federal Government Workers	3.5 (0.2)	3.1 (0.2)	0.4 (0.3)	0.70
Self-Employed Not Incorporated Workers	6.3 (0.3)	6.2 (0.3)	<0.1 (0.4)	1.00
Self-Employed Incorporated Workers	4.7 (0.3)	4.2 (0.2)	0.5 (0.4)	0.70
Unpaid Family Workers	0.3 (0.1)	0.7 (0.1)	-0.3 (0.1)	0.05*
Total	100.0	100.0		

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: χ^2:22.6, p-value<0.01
Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test at the α=0.1 level. P-values have been adjusted for multiple comparisons using the Holm-Bonferroni method.

Table 12. Response Distribution: Mail Mode

Category	Test Percent (n=3,958)	Control Percent (n=4,449)	Test minus Control	Adjusted P-Value
Private For-Profit Workers	65.2 (1.0)	65.4 (0.9)	-0.1 (1.3)	1.00
Private Not-For-Profit Workers	9.5 (0.6)	8.3 (0.5)	1.2 (0.9)	1.00
Local Government Workers	8.1 (0.5)	7.2 (0.5)	0.9 (0.7)	1.00
State Government Workers	4.1 (0.4)	4.1 (0.4)	<0.1 (0.5)	1.00
Federal Government Workers	2.3 (0.3)	2.8 (0.3)	-0.5 (0.4)	1.00
Self-Employed Not Incorporated Workers	6.3 (0.6)	7.1 (0.5)	-0.8 (0.8)	1.00
Self-Employed Incorporated Workers	4.1 (0.4)	4.2 (0.4)	-0.1 (0.6)	1.00
Unpaid Family Workers	0.4 (0.1)	1.0 (0.2)	-0.6 (0.2)	0.06*
Total	100.0	100.0		

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: χ^2:13.1, p-value=0.07
Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test at the α=0.1 level. P-values have been adjusted for multiple comparisons using the Holm-Bonferroni method.

In Table 13, the CATI and CAPI results are combined to ensure sufficiently large cell sizes. The response proportion for the Private Not-For-Profit category was 2.4 percentage points higher in the test treatment than in the control.

Table 13. Response Distribution: CATI/CAPI

Category	Test Percent (n=5,811)	Control Percent (n=5,807)	Test minus Control	Adjusted P-Value
Private For-Profit Workers	71.1 (0.9)	71.0 (1.0)	0.1 (1.5)	1.00
Private Not-For-Profit Workers	7.2 (0.6)	4.8 (0.4)	2.4 (0.8)	0.02*
Local Government Workers	5.2 (0.4)	5.9 (0.5)	-0.7 (0.6)	1.00
State Government Workers	3.8 (0.4)	3.2 (0.5)	0.6 (0.6)	1.00
Federal Government Workers	2.2 (0.3)	3.3 (0.4)	-1.1 (0.5)	0.35
Self-Employed Not Incorporated Workers	6.3 (0.5)	7.1 (0.5)	-0.8 (0.7)	1.00
Self-Employed Incorporated Workers	3.2 (0.4)	3.8 (0.4)	-0.6 (0.5)	1.00
Unpaid Family Workers	0.4 (0.1)	0.2 (0.1)	0.2 (0.2)	1.00
Unspecified Private Workers	0.5 (0.1)	0.4 (0.1)	0.1 (0.2)	1.00
Unspecified Government Workers	0.1 (0.1)	<0.1 (<0.1)	0.1 (0.1)	1.00
Unspecified Self-Employed Workers	0.1 (<0.1)	0.1 (0.1)	-0.1 (0.1)	1.00
Total	100.0	100.0		

Source: U.S. Census Bureau, 2016 American Community Survey Content Test.
Note: χ^2:30.9, p-value<0.01
Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test at the α=0.1 level. P-values have been adjusted for multiple comparisons using the Holm-Bonferroni method.

6.3 Benchmarks

This section addresses research question 3: *How do the control and test estimates of Unpaid Family Workers compare with the 2015 Current Population Survey (CPS) estimate?*

Table 14 shows the proportion estimates for Unpaid Family Worker for the test and control treatments and for the 2015 CPS ASEC. The estimates of Unpaid Family Workers for the test and control treatments fall outside of the CPS ASEC's confidence interval (at the 90 percent confidence level).

Table 14. Benchmarking of Unpaid Family Workers

Category	Test Percent (n=21,083)	Control Percent (n=21,458)	CPS ASEC Percent[†] (n = 92,730)
Unpaid Family Workers	0.4 (0.1)	0.5 (0.1)	0.1 (<0.1)

Source: U.S. Census Bureau, 2016 American Community Survey Content Test, 2015 Current Population Survey - Annual Social and Economic Supplement
[†] CPS ASEC estimate excludes Alaska and Hawaii from the national estimate in effort to increase the comparability to the 2016 ACS Content Test.
Note: Standard errors are shown in parentheses.

6.4 Response Error

This section addresses research question 6: *Is response reliability the same for the control and test versions of the Class of Worker question?*

For housing units that responded to the original Content Test and for which we had a telephone number, a CFU telephone reinterview was conducted to measure response error using the GDR and IOI metrics. Overall (see Table 15), and in the internet mode (see Table D-1 in Appendix D), the GDR for Unpaid Family Workers in the control treatment was 0.3 percentage points higher than that of the test. It appears that the added wording to clarify the definition of Unpaid Family Workers (see Figure 2) resulted in a lower rate of inconsistent answers and a more reliable estimate for the test treatment. For the remaining modes (mail, CATI, and CAPI), there were no significant results (see Table D-2 through D-4 in Appendix D).

Table 15. Gross Difference Rates (GDR): Overall

Category	Test GDR Percent (n=8,391)	Control GDR Percent (n=8,801)	Test minus Control	P-Value
Private For-Profit Workers	12.1 (0.5)	11.8 (0.5)	0.3 (0.7)	0.63
Private Not-For-Profit Workers	7.2 (0.6)	6.2 (0.4)	1.0 (0.7)	0.13
Local Government Workers	2.8 (0.2)	2.9 (0.3)	-0.1 (0.4)	0.84
State Government Workers	3.0 (0.3)	2.7 (0.3)	0.3 (0.4)	0.41
Federal Government Workers	0.6 (0.1)	0.6 (0.1)	0.1 (0.2)	0.59
Self-Employed Not Incorporated Workers	4.0 (0.3)	4.3 (0.3)	-0.3 (0.5)	0.59
Self-Employed Incorporated Workers	2.8 (0.2)	2.7 (0.3)	0.1 (0.3)	0.76
Unpaid Family Workers	0.3 (0.1)	0.6 (0.1)	-0.3 (0.1)	0.01*

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test at the $\alpha=0.1$ level.

The IOI is the proportion of the total variance of a proportion estimate that is due to simple response variance. If the estimate of the index is less than 20, the response variance is low. If the estimate of the index is between 20 and 50, the response variance is moderate. If the estimate of the index is greater than 50, the response variance is high. There were no significant differences in the IOI between the test and control treatments for any of the response categories examined in Table 16.

The simple response variance for Unpaid Family Workers is high in both treatments, overall (see Table 16), and in the self-response modes (internet and mail) (see Tables D-5 and D-6, respectively in Appendix D). In the computer-assisted interview modes (CATI and CAPI) (see Tables D-7 and D-8, respectively), the simple response variance for Unpaid Family Workers was moderate in both treatments.

In the internet mode, two response categories, Self-Employed Not Incorporated and Self-Employed Incorporated Workers, had significantly higher IOI values in the test version than the control (35.2 percent versus 28.2 percent and 47.4 percent versus 28.2 percent, respectively).

Table 16. Index of Inconsistency (IOI): Overall

Category	Test IOI Percent (n=8,391)	Control IOI Percent (n=8,801)	Test minus Control	P-Value
Private For-Profit Workers	26.0 (1.1)	25.3 (1.1)	0.7 (1.4)	0.60
Private Not-For-Profit Workers	37.4 (2.3)	35.4 (2.0)	2.0 (3.0)	0.51
Local Government Workers	22.5 (1.7)	22.6 (2.0)	-0.1 (3.0)	0.98
State Government Workers	31.6 (3.2)	28.8 (2.5)	2.8 (4.2)	0.51
Federal Government Workers	10.6 (1.8)	8.5 (1.8)	2.1 (2.7)	0.43
Self-Employed Not Incorporated Workers	35.1 (2.8)	36.1 (2.5)	-1.0 (4.0)	0.80
Self-Employed Incorporated Workers	42.0 (3.5)	38.2 (3.9)	3.8 (4.8)	0.42
Unpaid Family Workers	74.0 (10.8)	78.6 (7.9)	-4.6 (12.7)	0.72

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. Significance was tested based on a two-tailed t-test at the α=0.1 level.

The L-fold IOI provides an overall response reliability of a question that has multiple mutually exclusive response categories. Table 17 indicates no significant L-fold IOI results overall, or in any mode of data collection.

Table 17. Index of Inconsistency (IOI): L-Fold

Category	Test Sample Size	Test IOI Percent	Control Sample Size	Control IOI Percent	Test minus Control	P-Value
Overall	8,391	31.0 (1.1)	8,801	30.4 (1.0)	0.6 (1.5)	0.68
Internet	5,336	25.3 (1.0)	5,361	23.4 (1.0)	1.9 (1.3)	0.15
Mail	1,425	27.9 (1.9)	1,643	28.6 (1.8)	-0.7 (2.6)	0.78
CATI	281	30.0 (4.0)	290	24.5 (4.3)	5.4 (6.4)	0.40
CAPI	1,349	41.9 (2.8)	1,507	43.2 (2.7)	-1.2 (3.7)	0.74

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. Significance was tested based on a two-tailed t-test at the α=0.1 level.

6.5 Results for Analysis Specific to Class of Worker

This section addresses research question 7: *In the mail mode, are there more multiple responses to the Class of Worker question in test than in control?*

The respondent is instructed to mark only one response category for the Class of Worker question, not multiple categories. In the internet, CATI, and CAPI modes, we have built-in checks that do not allow respondents to select more than one response. However, in the mail mode where respondents complete a paper questionnaire, respondents can incorrectly select multiple categories.

Table 18 presents results on the percent of eligible respondents marking multiple boxes in the test and control treatments for the mail mode. The expectation was that the test treatment would result in the same or fewer multiple mark responses than control. However, the test treatment had significantly more multiple mark responses (3.8 percentage points higher).

36

In the test version of the Class of Worker question, response categories were grouped under the three general headings: Private Sector Employee, Government Employee, and Self-Employed or Other. In the control version, the response categories were listed in a single group with no headings (See Figures 1 and 2).

Further analysis of multiple mark responses in the test treatment showed the about 75.0 percent of respondents who marked more than one response category first selected the category of Private For-Profit under the Private Sector Employee subheading. Table E-1 in Appendix E displays the distributions of the second category selected after respondents marked Private For-Profit in the test treatment. It shows that the largest categories of the second selection were Self-Employed Incorporated Workers (30.8 percent) and Self-Employed Not Incorporated Workers (29.4 percent). For the majority of these cases, we believe the multiple mark responses can be addressed during the Industry and Occupation coding and the editing process and should therefore not impact the final data quality.

Table 18. Percent of Eligible Persons with Multiple Boxes Marked (Mail Mode Only)

Mode	Test Percent (n=4,811)	Control Percent (n=5,126)	Test minus Control	P-Value
Mail	4.4 (0.4)	0.6 (0.1)	3.8 (0.4)	<0.01*

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test at the α=0.1 level.

This section addresses research question 9: *In the telephone interviewing (CATI) mode, are the control and test rates of "unspecified" responses the same?*

In CATI, an "unspecified" response indicates that a respondent answered the general Class of Worker question (worked for a Private Company or Organization, or Government, or was Self-Employed), but did not provide a specific type of Private Company or Organization employee (For-Profit or Not-For-Profit), Government employee (Local, State, or Federal), or Self-Employment (Incorporated or Not Incorporated) during the follow-up question of the general category.[30] The difference in the percentage of "unspecified" responses between the test and control treatments was not statistically significant for any of these three categories (see Table 19).

---

[30] See Figure B-1 and B-3 in the Appendix B for CATI/CFU Class of Worker question format and order for test and control treatments.

Table 19. Percent of "Unspecified" Class of Worker Responses

Category	Test Sample Size	Test Percent	Control Sample Size	Control Percent	Test minus Control	P-Value
Unspecified Private Workers	584	6.8 (1.6)	590	6.3 (1.4)	0.5 (2.1)	0.82
Unspecified Government Workers	142	0.5 (0.4)	143	0.8 (0.5)	-0.3 (0.5)	0.60
Unspecified Self-Employed Workers	128	4.4 (2.3)	104	0.6 (0.4)	-3.8 (2.4)	0.11

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. Significance was tested based on a two-tailed t-test at the α=0.1 level.

This section addresses research question 10: *Is the reporting of self-employment income and wages income consistent with the Class of Worker response?*

The comparisons between control and test treatments were based on the universe of people 15 years and older who worked last week. The editing process for Class of Worker, Industry and Occupation uses assumptions based on income to address inconsistencies. An income consistency check with Class of Worker is a useful analysis to understand how valid those assumptions remain.

Table 20 contains information on the percentage of Self-Employed Not Incorporated Workers who also reported having non-zero self-employment income (i.e., positive or negative income). We expect respondents who select the category of Self-Employed Not Incorporated to report self-employment income. For both the control and test treatment, about 59.0 percent of respondents who selected Self-Employed Not Incorporated reported self-employed income – about 40 percent did not. However, it is important to note, that income data is collected for all jobs worked in the past 12 months, while Class of Worker data are collected for the primary job last week. The difference between the results for the test and control treatments was not significant.

Table 20. Percent of Self-Employed Not Incorporated Workers Who Reported Non-Zero Self-Employment Income Amounts

Category	Test Percent (n=1,194)	Control Percent (n=1,332)	Test minus Control	P-Value
Self-Employed Not Incorporated Workers	59.7 (1.9)	58.8 (2.0)	0.9 (2.5)	0.71

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. Significance was tested based on a two-tailed t-test at the α=0.1 level.

Table 21 presents the percentage of private workers that reported having non-zero wages (i.e., positive or negative wages) for the test and control treatments. The expectation is that a respondent working as a Private For-Profit, Private Not-For-Profit, or Self-Employed Incorporated Worker would also report a wage income amount. About 95.0 percent of them did in both treatments. The difference between these results was not significant.

Table 21. Percent of Private Workers Who Reported Non-Zero Wage Income Amounts

Category	Test Percent (n=17,683)	Control Percent (n=17,926)	Test minus Control	P-Value
Private Workers	94.7 (0.2)	94.9 (0.2)	-0.2 (0.3)	0.57

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. Significance was tested based on a two-tailed t-test at the α=0.1 level.

Table 22 displays the percent of Unpaid Family Workers who reported non-zero (i.e., positive or negative) wage or self-employment income amounts. By definition, an Unpaid Family Worker should not receive any earnings (wages and/or self-employment income) for their work in a family business or farm. About 36.6 percent of these respondents in the test version and about 52.8 percent in the control version reported inconsistently. The difference between these results was not significant. However, respondents who select the category of Unpaid Family Workers often have wage, self-employment income, or both. The wage and self-employment income may reflect any earnings from all jobs held during the 12 months prior to the interview. The Class of Worker status reflects the job or business held the week prior to the 2016 ACS Content Test interview. As a result, some of the inconsistency between Unpaid Family Workers also reporting non-zero wages or income might be the result in the differences in the reference periods between the questions.

Table 22. Percent of Unpaid Family Workers Who Reported Non-Zero Wage or Self-Employment Income Amounts

Category	Test Percent (n=61)	Control Percent (n=87)	Test minus Control	P-Value
Unpaid Family Workers	36.6 (9.2)	52.8 (6.9)	-16.2 (11.8)	0.17

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. Significance was tested based on a two-tailed t-test at the α=0.1 level.

This section addresses research question 11: *Do those persons reported as Unpaid Family Workers work at least 15 hours per week? How often is this not the case?*

Table 23 presents the percentage of respondents who reported themselves as Unpaid Family Workers and worked at least 15 hours per week for test and control treatments. By the Bureau of Labor Statistics definition, to be classified as an Unpaid Family Worker, a respondent must have worked without pay for at least 15 hours in a family business or farm. We anticipated that the addition of more precise wording for Unpaid Family Workers in the test treatment would result in a higher proportion of Unpaid Family Workers that worked at least 15 hours per week. This modification resulted in a significantly higher proportion in the test treatment (23.7 percentage points higher).

Table 23. Percent of Unpaid Family Workers Whose Usual Hours Worked per Week is 15 Hours or More

Mode	Test Sample Size	Test Percent	Control Sample Size	Control Percent	Test minus Control	P-Value
Overall	76	63.6 (7.9)	135	39.9 (5.6)	23.7 (10.5)	0.02*
Internet	42	45.3 (11.2)	83	36.2 (6.4)	9.2 (13.8)	0.51
Mail	-	-	-	-	-	-
CATI	-	-	-	-	-	-
CAPI	-	-	-	-	-	-

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test at the α=0.1 level. An entry of '-' in a cell indicates that no or too few observations were available to meet statistical standards for reliability.

Table 24 indicates that the percent of Unpaid Family Workers whose usual hours worked per week is 15 hours or more and have zero wage or self-employment income is relatively low in both the test and control treatments (16.2 percent and 18.5 percent, respectively). A possible explanation for the low consistency is the difference in the reference period of Class of Worker and income. As stated previously, Unpaid Family Workers may have non-zero wage and self-employment income. The income questions reflect earnings held for all jobs in the prior 12 months before the survey was conducted, while the category of Unpaid Family Worker reflects the person's main employment or business held the week prior to the survey. When the person has more than one job, we collect information on the job where more hours were worked. The difference between the results for the test and control treatments was not significant overall or by the internet mode of data collection.

Table 24. Percent of Unpaid Family Workers Whose Usual Hours Worked per Week is 15 Hours or More and Have Zero Wage or Self-Employment Income

Mode	Test Sample Size	Test Percent	Control Sample Size	Control Percent	Test minus Control	P-Value
Overall	76	16.2 (6.5)	135	18.5 (4.5)	-2.3 (8.3)	0.78
Internet	42	14.6 (6.8)	83	15.7 (5.6)	-1.1 (9.4)	0.91
Mail	-	-	-	-	-	-
CATI	-	-	-	-	-	-
CAPI	-	-	-	-	-	-

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test at the α=0.1 level. An entry of '-' in a cell indicates that no or too few observations were available to meet statistical standards for reliability.

7   CONCLUSIONS AND RECOMMENDATIONS

This report presents the results of proposed modifications to the Class of Worker question (for the 2016 ACS Content Test). Class of Worker categorizes people according to the type of ownership of the employing organization. The Class of Worker question seeks to determine whether an employed person is salaried or self-employed, and if the person works in the private sector or in government (i.e., the public sector). The motivation for modifying this question was to clarify the intent of the question and the response categories while also improving the question layout.

Based on the decision criteria, the majority of the analyses provided evidence to support the implementation of the test version of the Class of Worker question over the control version. As expected, the rate of Unpaid Family Workers working at least 15 hours or more was higher for the test. We expected this difference to be significant because the test version provided a clearer definition of this category. In addition, the overall response error for Unpaid Family Workers was significantly lower in the test than the control treatment. However, the proportional estimates for Unpaid Family Worker were not significantly different between the test and control treatment. Furthermore, the wage and income consistency checks, and the proportion of "unspecified" responses (CATI mode only) were not significantly different between test and control treatments.

Our highest priority and primary decision criteria concerned item missing data rates. We expected item missing data rates for the test version to be the same or lower than the control version, for all modes. There were no significant differences in the item missing data rates between the treatments overall and across all modes except mail. In the mail mode, the item missing data rate was significantly higher in the test treatment. A possible explanation for the higher item missing data rate in mail mode is the significantly higher rate of multiple mark responses in the test treatment, which were treated as missing values. Multiple mark responses are only possible in the mail mode, as the survey instruments in the other modes do not allow respondents to mark multiple responses.

Further analysis into multiple mark responses showed the majority of these respondents were selecting Private For-Profit and a second category of Self-Employed Incorporated or Self-Employed Not Incorporated. This pattern provides insight into how we could address multiple marks in our coding and editing process if the test treatment of the question is implemented. The majority of multiple mark respondents are selecting a specific combination of logical response categories rather than selecting contradictory response categories, such as Private For-Profit along with a Government employee category. Ideally, we would like to reduce the number of multiple marks in the mail mode and future research should examine this issue. However, for now we feel confident we can address multiple marks in our editing process.

Adding Active Duty as a Class of Worker category and removing it from the Employer Name question appears to reduce respondent confusion as indicated by the lower rate of response error for Military industries in the *2016 American Community Survey Content Test Evaluation Report: Industry and Occupation.*

The modified instructions and wording of the Class of Worker categories in the test treatment is generally an improvement over the control treatment, with the exception of multiple mark responses discussed above. The layout and wording was preferred by respondents during cognitive testing and is similar in format to other federal surveys, such as the National Survey of College Graduates (NCSG). Our expectation is that if implemented, the changes to the Class of Worker question and proposed modifications for the Industry and Occupation questions will improve the overall quality of the employment data collected by this series of questions. Any significant differences between the treatments were minor, suggesting future continuity in data reliability and consistency, and no break in series should the test version be implemented. The recommendation of the Industry and Occupation Statistics Branch is to move forward with the implementation of the test version of the Class of Worker question.

8   ACKNOWLEDGEMENTS

9   REFERENCES

Dusch, G., & Meier, F. (2012). *2010 Census Content Reinterview Survey Evaluation Report,* U.S. Census Bureau, June 13, 2012. Retrieved May 17, 2016, from http://www.census.gov/2010census/pdf/2010_Census_Content_Reinterview_Survey_Evaluation_Report.pdf

Flanagan, P. (1996). *Survey Quality & Response Variance* (Unpublished Internal Document). U.S. Census Bureau. Demographic Statistical Methods Division. Quality Assurance and Evaluation Branch.

Holder, Kelly A. & Raglin D. (2007). Evaluation Report Covering Employment Status. (2006 American Community Survey Content Test Report P.6.a.). Retrieved July 1, 2017, from United States Census Bureau Official Web Site: https://www.census.gov/content/dam/Census/library/working-papers/2007/acs/2007_Holder_01.pdf.

Holm, S. (1979). "A Simple Sequentially Rejective Multiple Test Procedure," Scandinavian Journal of Statistics, Vol. 6, No. 2: 65-70. Retrieved on January 31, 2017, from https://www.jstor.org/stable/4615733?seq=1#page_scan_tab_contents.

Kirk, M. (1996). Summary Results on 1996 National Content Survey for Industry, Occupation, and Class of Worker [Unpublished]. U.S. Census Bureau.

Martínez, A., Montalvo, A., & Oliver, B. (2017). *2016 American Community Survey Content Test Evaluation Report: Industry and Occupation.* U.S. Census Bureau.

Polivka, A., & Rothgeb, J. (1993). Redesigning the CPS Questionnaire. *Monthly Labor Review, 116*(9), 10-28.

Polivka, A., & Miller, S. (1998). The CPS after the Redesign: Refocusing the Economic Lens. *Labor Statistics Measurement Issues,* 249-289. Retrieved on July 15, 2015, from http://www.nber.org/chapters/c8362.pdf.

Raglin, D. (2014). *American Community Survey Fiscal Year 2014 Content Review Interviewer Survey Results.* U.S. Census Bureau.

Rao, J. N. K., & Scott, A. J. (1987). "On Simple Adjustments to Chi-Square Tests with Sample Survey Data," The Annals of Statistics, Vol. 15, No. 1, 385-397. Retrieved on January 31, 2017, from http://projecteuclid.org/euclid.aos/1176350273

Stapleton, M., & Steiger, D. (2015). *Cognitive Testing of the 2016 American Community Survey Content Test Items: Summary Report for Round 1 and Round 2 Interviews.* Westat, Rockville, Maryland, January 2015.

Steiger, D., Anderson, J., Folz, J., Leonard, M., & Stapleton, M. (2015). *Cognitive Testing of the 2016 American Community Survey Content Test Items: Briefing Report for Round 3 Interviews.* Westat, Rockville, Maryland, June, 2015.

U.S. Bureau of Labor Statistics. (2015). *Labor Force Statistics from the Current Population Survey*. Retrieved June 25, 2015, from U.S. Bureau of Labor Statistics Official Web Site: http://www.bls.gov/cps/

U.S. Bureau of Labor Statistics. (2015). *Definition: Unpaid Family Workers.* Retrieved February 23, 2017, from U.S. Bureau of Labor Statistics Official Web Site: https://www.bls.gov/bls/glossary.htm#U

U.S. Census Bureau. (2014a). *American Community Survey Design and Methodology (January 2014).* Retrieved February 1, 2017, from United States Census Bureau Official Web Site: http://www.census.gov/programs-surveys/acs/methodology/design-and-methodology.html

U.S. Census Bureau. (2014b, October 28). Health Insurance and Class of Worker Debriefing with CATI Interviewers. Hagerstown, Maryland.

U.S. Census Bureau. (2015a, February 3). Health Insurance and Class of Worker Debriefing with CAPI Field Representatives. Richmond, Virginia.

U.S. Census Bureau. (2015b). *American Community Survey, 1996 to 2015 Questionnaires.* Retrieved July 13, 2015, from American Community Survey Questionnaire Archive: http://www.census.gov/programs-surveys/acs/methodology/questionnaire-archive.html

U.S. Census Bureau. (2015c). *Survey of Business Owners - Main*. Retrieved June 25, 2015, from United States Census Bureau Official Web Site: https://census.gov/programs-surveys/sbo.html

U.S. Census Bureau. (2015d). *Government Employment & Payroll*. Retrieved June 25, 2015, from United States Census Bureau Official Web Site: http://www.census.gov/govs/apes/index.html

U.S. Census Bureau. (2015e). *2015 CPS Annual Social and Economic Supplement (ASEC) Technical Documentation*. Retrieved on February 23, 2017, from United States Census Bureau Official Web Site: https://www2.census.gov/programs-surveys/cps/techdocs/cpsmar15.pdf.

U.S. Census Bureau. (2016a). *2015 Planning Database Tract Data* [Data file]. Retrieved on January 31, 2017, from United States Census Bureau Official Web Site: http://www.census.gov/research/data/planning_database/2015/

U.S. Census Bureau. (2016b). *Fact Sheet: Differences between the American Community Survey (ACS) and the Annual Social and Economic Supplement to the Current Population Survey (CPS ASEC)*. Retrieved on January 31, 2017, from United States Census Bureau Official Web Site: https://www.census.gov/topics/income-poverty/poverty/guidance/data-sources/acs-vs-cps.html

U.S. Census Bureau. (2016c). Table B24080, Sex by Class of Worker for the Civilian Employed Population 16 years and over, ACS 1-year estimates 2011-2015. Retrieved on January 31, 2017, from United States Census Bureau Official Web Site: https://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml.

U.S. Census Bureau. (2016d). Current Population Survey Class of Worker Data, 2011-2015. Retrieved on January 31, 2017, from Data Ferrett on United States Census Bureau Official Web Site: https://dataferrett.census.gov/.

Appendix A: Internet Versions of the Control and Test Questions

Figure A-1. Internet Control Version of the Class of Worker Question

The next series of questions are about the type of business (Name) worked for and the type of work that (he/she/he or she) did.
Describe clearly (name's) chief job activity or business last week. If (name) had more than one job, describe the one at which (he/she/he or she) worked the most hours. If (name) had no job or business last week, give information for (his/her/his or her) last job or business.

Was (name) –
- 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?
- an ACTIVE DUTY U.S. Armed Forces member?
- a Federal GOVERNMENT employee (excluding active duty military)?
- 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?

What was the name of (Name)'s company, business, or other employer?
_____

Which branch of the Armed Forces does (Name) work for?
- U.S. Army
- U.S. Navy
- U.S. Air Force
- U.S. Marine Corps
- U.S. Coast Guard

Figure A-2. Internet Test Version of the Class of Worker Question

COWA [40]
40. DESCRIPTION OF EMPLOYMENT
The next series of questions is about the type of employment (Name) had last week.

If (Name) had more than one job, describe the one at which the most hours were worked.

If (Name) did not work last week, describe the most recent employment in the past five years.

a. Which one of the following best describes (Name)'s employment last week or the most recent employment in the past 5 years?

PRIVATE SECTOR EMPLOYEE
○   **For-profit** company or organization
○   **Non-profit** organization (including tax-exempt and charitable organizations)

GOVERNMENT EMPLOYEE
○   **Local government** (for example: city or county school district)
○   **State government** (including state colleges/universities)
○   **Active duty** U.S. Armed Forces or Commissioned Corps
○   **Federal government** civilian employee

SELF-EMPLOYED OR OTHER
○   **Owner of non-incorporated** business, professional practice, or farm
○   **Owner of incorporated** business, professional practice, or farm
○   Worked **without pay** in a **for-profit** family business or farm for 15 hours or more per week

----

INW2 (if COWA not = 5 "Active duty…") [40b]
b. What was the name of (Name)'s employer, business, or agency?

| _____ | *[60 characters]*

INMIL (if COWA= 5 "Active duty...") [40b]
b. Which branch of the Armed Forces or Commissioned Corps did (Name) work for?

○   U.S. Army
○   U.S. Navy
○   U.S. Air Force
○   U.S. Marine Corps
○   U.S. Coast Guard
○   U.S. Public Health Service
○   National Oceanic and Atmospheric Administration (NOAA)

Appendix B: CATI/CFU and CAPI Versions of the Control and Test Questions

Figure B-1. CATI/CFU Control Version of the Class of Worker Question

The next series of questions are about the type of business {Fill 1: <Name>/ you} worked-for and the type of work that {Fill 3: he/ she/<Name>/you} did....(If {Fill 7: <Name>/ you} had more than 1 job, describe the one at which the most hours were worked. If {Fill 8: <Name>/ you} did not work last week, give information for the last job or business in the past five years.)

Let's start with the first question. I am going to read 5 categories. Please pick the one that best describes who   {Fill 5: he/ she/<Name>/ you} worked for - a private organization or company, government, the US Armed Forces (active duty), self-employed, or working without pay in a family business.

<1> Private organization or company
<2> Government
<3> US Armed Forces (active duty)
<4> Self-employed
<5> Working without pay in a family business

Was this a non-profit organization or a for-profit company?

        <1> Non-profit organization
        <2> For-profit company

Was this for Local, State, or the Federal Government?

        <1> Local
        <2> State
        <3> Federal

Was this self-employment incorporated or not incorporated?

        <1> Incorporated
        <2> Not incorporated

What was the name of <(Name)'s/your> employer, business, or agency?

_____

    Which branch of the Armed Forces <does (Name)/do you> work for?
    <1> U.S. Army
    <2> U.S. Navy
    <3> U.S. Air Force
    <4> U.S. Marine Corps
    <5> U.S. Coast Guard

Figure B-2. CAPI Control Version of the Class of Worker Question

The next series of questions are about the type of business {Fill 1: <Name>/ you} worked for and the type of work that {Fill 3: he/ she/<Name>/ you} did…. (If {Fill 7: <Name>/ you} had more than 1 job, describe the one at which the most hours were worked. If {Fill 8: <Name>/ you} did not work last week, give information for the last job or business in the past five years.)

Let's start with the first question. Using Card H, please pick the category that best describes who {Fill 5: he/ she/<Name>/ you} worked for.

<1> An employee of a PRIVATE FOR-PROFIT company or business, or of an individual for wages, salary, or commissions?
<2> An employee of a PRIVATE NOT-FOR-PROFIT, tax-exempt, or charitable organization?
<3> A local GOVERNMENT employee (city, county, [Fill 9: municipio,] etc.)?
<4> A state GOVERNMENT employee?
<5> An active duty U.S. Armed Forces member?
<6> A federal GOVERNMENT employee (excluding active duty military)?
<7> SELF-EMPLOYED in own NOT INCORPORATED business, professional practice, or farm?
<8> SELF-EMPLOYED in own INCORPORATED business, professional practice, or farm?
<9> Working WITHOUT PAY in family business or farm?

What was the name of <(Name)'s/your> employer, business, or agency?
_____

Which branch of the Armed Forces <does (Name)/do you> work for?
  <1> U.S. Army
  <2> U.S. Navy
  <3> U.S. Air Force
  <4> U.S. Marine Corps
  <5> U.S. Coast Guard

Figure B-3. CATI/CFU Test Version of the Class of Worker Question

The next series of questions is about the type of employment {Fill 1: <Name>/ you} had {Fill 2: last week/most recently in the past 5 years}.

(If {Fill 1: <Name>/ you} had more than one job, describe the one at which the most hours were worked.

I am going to read 5 categories. Please choose the one that best describes {Fill 3: <Name>'s/your>} employment - a private organization or company, government, active duty U.S. Armed Forces or Commissioned Corps, self-employed, or worked without pay in a for-profit family business or farm.

> <1> Private company or organization
> <2> Government
> <3> Active duty U.S. Armed Forces or Commissioned Corps
> <4> Self-employed
> <5> Working without pay in a for-profit family business or farm

Did {Fill 1: <Name>/you} work for a for-profit company or non-profit organization?

> <1> For-profit company
> <2> Non-profit organization

Did {Fill 1: <Name>/you} work for a local, state, or federal government?

> <1> Local
> <2> State
> <3> Federal

Was {Fill 1: <Name>'s/your} self-employed business, professional practice, or farm incorporated or not incorporated?

> <1> Incorporated
> <2> Not incorporated

Did {Fill 1: <Name>/you} work without pay in this for-profit family business or farm for 15 hours or more per week?

> <1> Yes
> <2> No

What was the name of <(Name)'s/your> employer, business, or agency?

_____

Which branch of the Armed Forces or Commissioned Corps did <(Name)/you> work for?
<1> U.S. Army
<2> U.S. Navy
<3> U.S. Air Force
<4> U.S. Marine Corps
<5> U.S. Coast Guard
<6> U.S. Public Health Service
<7> National Oceanic and Atmospheric Administration (NOAA)

Figure B-4. CAPI Test Version of the Class of Worker Question

The next series of questions is about the type of employment {Fill 1: <Name>/ you} had {Fill 2: last week/most recently in the past 5 years}.

(If {Fill 1: <Name>/ you} had more than one job, describe the one at which the most hours were worked.

Let's start with the first question. Using Card H, which one of the following best describes {Fill 3: <Name>'s/your} employment?

<1> For-profit company or organization [PRIVATE SECTOR EMPLOYEE]
<2> Non-profit organization (including tax-exempt and charitable organizations) [PRIVATE SECTOR EMPLOYEE]
<3> Local government [GOVERNMENT EMPLOYEE]
<4> State government (including state colleges/universities) [GOVERNMENT EMPLOYEE]
<5> Active duty U.S. Armed Forces or Commissioned Corps [GOVERNMENT EMPLOYEE]
<6> Federal government civilian employee [GOVERNMENT EMPLOYEE]
<7> Owner of non-incorporated business, professional practice, or farm [SELF-EMPLOYED]
<8> Owner of incorporated business, professional practice, or farm [SELF- EMPLOYED]
<9> Worked without pay in a for-profit family business or farm for 15 hours or more per week

What was the name of <(Name)'s/your> employer, business, or agency?
_____

Which branch of the Armed Forces or Commissioned Corps did <(Name)/you> work for?
<1> U.S. Army
<2> U.S. Navy
<3> U.S. Air Force
<4> U.S. Marine Corps
<5> U.S. Coast Guard
<6> U.S. Public Health Service
<7> National Oceanic and Atmospheric Administration (NOAA)

Appendix C: Unit Response Rates by Designated High and Low Response Areas

Table C-1. Unit Response Rates by Designated High (HRA) and Low (LRA) Response Areas

Mode	Test Interviews	Test Percent	Control Interviews	Control Percent	Test minus Control	P-Value
Total Response	19,400		19,455			
HRA	7,556	94.3 (0.4)	7,608	94.5 (0.3)	-0.2 (0.6)	0.72
LRA	11,844	91.5 (0.3)	11,847	91.0 (0.3)	0.5 (0.5)	0.29
Difference		2.7 (0.5)		3.5 (0.5)	-0.7 (0.7)	0.33
Self-Response	13,131		13,284			
HRA	6,201	59.7 (0.7)	6,272	60.6 (0.7)	-0.9 (0.9)	0.31
LRA	6,930	33.2 (0.4)	7,012	33.6 (0.4)	-0.4 (0.6)	0.55
Difference		26.5 (0.8)		27.0 (0.8)	-0.5 (1.2)	0.66
Internet	8,168		8,112			
HRA	4,119	39.6 (0.6)	4,048	39.1 (0.6)	0.5 (0.8)	0.51
LRA	4,049	19.4 (0.3)	4,064	19.5 (0.3)	0.1 (0.4)	0.87
Difference		20.2 (0.6)		19.6 (0.7)	0.6 (0.9)	0.52
Mail	4,963		5,172			
HRA	2,082	20.0 (0.4)	2,224	21.5 (0.4)	-1.5 (0.6)	0.02*
LRA	2,881	13.8 (0.3)	2,948	14.1 (0.3)	-0.3 (0.4)	0.43
Difference		6.2 (0.5)		7.4 (0.4)	-1.1 (0.7)	0.11
CATI	872		880			
HRA	296	9.0 (0.5)	301	9.6 (0.6)	-0.6 (0.8)	0.44
LRA	576	7.9 (0.4)	579	8.0 (0.3)	-0.1 (0.5)	0.85
Difference		1.1 (0.6)		1.6 (0.7)	-0.5 (0.9)	0.58
CAPI	5,397		5,291			
HRA	1,059	82.2 (1.0)	1,035	82.7 (0.9)	-0.5 (1.3)	0.69
LRA	4,338	85.8 (0.5)	4,256	85.0 (0.4)	0.8 (0.7)	0.23
Difference		-3.7 (1.1)		-2.3 (1.0)	-1.3 (1.5)	0.36

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: Standard errors are in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test at the α=0.1 level. The weighted response rates account for initial sample design as well as CAPI subsampling.

Appendix D: Response Reliability Tables by Mode

Table D-1. Gross Difference Rates (GDR): Internet

Category	Test GDR Percent (n=5,336)	Control GDR Percent (n=5,361)	Test minus Control	P-Value
Private For-Profit Workers	9.7 (0.5)	8.8 (0.5)	0.9 (0.8)	0.23
Private Not-For-Profit Workers	5.6 (0.4)	5.4 (0.4)	0.2 (0.5)	0.74
Local Government Workers	3.1 (0.3)	2.6 (0.3)	0.5 (0.4)	0.21
State Government Workers	2.6 (0.3)	2.7 (0.3)	-0.1 (0.3)	0.81
Federal Government Workers	0.7 (0.1)	0.6 (0.1)	0.1 (0.2)	0.57
Self-Employed Not Incorporated Workers	3.8 (0.3)	3.2 (0.3)	0.6 (0.4)	0.16
Self-Employed Incorporated Workers	3.1 (0.3)	2.6 (0.3)	0.5 (0.4)	0.25
Unpaid Family Workers	0.3 (0.1)	0.6 (0.1)	-0.3 (0.2)	0.09*

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test at the α=0.1 level.

Table D-2. Gross Difference Rates (GDR): Mail

Category	Test GDR Percent (n=1,425)	Control GDR Percent (n=1,643)	Test minus Control	P-Value
Private For-Profit Workers	10.5 (0.8)	10.6 (0.9)	-0.1 (1.2)	0.94
Private Not-For-Profit Workers	6.2 (0.6)	5.9 (0.7)	0.3 (0.9)	0.75
Local Government Workers	3.1 (0.5)	3.0 (0.6)	0.1 (0.8)	0.91
State Government Workers	2.5 (0.6)	2.6 (0.5)	>-0.1 (0.8)	0.96
Federal Government Workers	0.5 (0.2)	1.0 (0.3)	-0.6 (0.4)	0.10†
Self-Employed Not Incorporated Workers	3.2 (0.6)	4.4 (0.6)	-1.2 (0.8)	0.13
Self-Employed Incorporated Workers	3.4 (0.5)	2.3 (0.5)	1.1 (0.7)	0.12
Unpaid Family Workers	0.1 (0.1)	0.6 (0.3)	-0.5 (0.3)	0.10†

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
† These P-Values are rounded values to the nearest tenth. The actual p-values are larger than 0.10.
Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. Significance was tested based on a two-tailed t-test at the α=0.1 level.

Table D-3. Gross Difference Rates (GDR): CATI

Category	Test GDR Percent (n=281)	Control GDR Percent (n=290)	Test minus Control	P-Value
Private For-Profit Workers	11.8 (2.3)	8.2 (2.2)	3.6 (3.4)	0.29
Private Not-For-Profit Workers	6.6 (1.7)	4.8 (1.5)	1.8 (2.3)	0.44
Local Government Workers	2.7 (1.1)	4.6 (1.7)	-1.9 (2.1)	0.37
State Government Workers	3.1 (1.3)	2.6 (1.2)	0.4 (1.8)	0.80
Federal Government Workers	1.8 (1.1)	1.2 (0.8)	0.6 (1.5)	0.69
Self-Employed Not Incorporated Workers	2.4 (1.1)	2.3 (1.1)	0.1 (1.6)	0.97
Self-Employed Incorporated Workers	3.8 (1.5)	2.5 (1.4)	1.3 (2.1)	0.52
Unpaid Family Workers	0.8 (0.7)	0.8 (0.8)	>-0.1 (1.1)	0.98

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. Significance was tested based on a two-tailed t-test at the α=0.1 level.

Table D-4. Gross Difference Rates (GDR): CAPI

Category	Test GDR Percent (n=1,349)	Control GDR Percent (n=1,507)	Test minus Control	P-Value
Private For-Profit Workers	16.2 (1.3)	16.9 (1.4)	-0.7 (1.6)	0.68
Private Not-For-Profit Workers	9.9 (1.4)	7.5 (1.0)	2.5 (1.7)	0.16
Local Government Workers	2.3 (0.5)	3.1 (0.6)	-0.9 (0.8)	0.29
State Government Workers	3.6 (0.8)	2.6 (0.5)	1.0 (0.9)	0.28
Federal Government Workers	0.6 (0.2)	0.3 (0.2)	0.3 (0.3)	0.29
Self-Employed Not Incorporated Workers	4.8 (0.9)	5.9 (0.8)	-1.2 (1.3)	0.36
Self-Employed Incorporated Workers	2.0 (0.5)	2.9 (0.5)	-0.9 (0.7)	0.18
Unpaid Family Workers	0.2 (0.1)	0.5 (0.3)	-0.3 (0.3)	0.26

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. Significance was tested based on a two-tailed t-test at the α=0.1 level.

Table D-5. Index of Inconsistency (IOI): Internet

Category	Test IOI Percent (n=5,336)	Control IOI Percent (n=5,361)	Test minus Control	P-Value
Private For-Profit Workers	20.3 (1.1)	18.4 (1.1)	2.0 (1.6)	0.21
Private Not-For-Profit Workers	27.1 (1.8)	26.7 (1.9)	0.4 (2.6)	0.88
Local Government Workers	21.0 (1.8)	18.5 (2.1)	2.5 (2.8)	0.38
State Government Workers	24.7 (2.4)	24.3 (2.4)	0.5 (2.9)	0.87
Federal Government Workers	9.4 (1.8)	9.1 (2.1)	0.3 (2.8)	0.92
Self-Employed Not Incorporated Workers	35.2 (2.7)	28.2 (2.6)	7.0 (3.6)	0.05*
Self-Employed Incorporated Workers	47.4 (4.0)	36.1 (4.2)	11.4 (5.8)	0.05*
Unpaid Family Workers	72.9 (15.3)	77.8 (11.3)	-4.9 (19.0)	0.80

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test at the α=0.1 level.

Table D-6. Index of Inconsistency (IOI): Mail

Category	Test IOI Percent (n=1,425)	Control IOI Percent (n=1,643)	Test minus Control	P-Value
Private For-Profit Workers	22.6 (1.8)	22.8 (2.0)	-0.2 (2.5)	0.94
Private Not-For-Profit Workers	34.3 (3.8)	34.2 (3.8)	0.2 (5.2)	0.97
Local Government Workers	24.8 (4.0)	20.5 (4.1)	4.3 (6.2)	0.49
State Government Workers	22.7 (5.5)	29.2 (5.4)	-6.4 (7.8)	0.41
Federal Government Workers	9.4 (4.2)	21.1 (6.7)	-11.7 (7.4)	0.11
Self-Employed Not Incorporated Workers	24.5 (4.4)	33.1 (4.4)	-8.6 (6.2)	0.16
Self-Employed Incorporated Workers	50.9 (7.2)	35.9 (7.5)	15.0 (9.9)	0.13
Unpaid Family Workers	100.0 (0.0)	71.7 (20.8)	28.3 (20.8)	0.17

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding.
Significance was tested based on a two-tailed t-test at the α=0.1 level.

Table D-7. Index of Inconsistency (IOI): CATI

Category	Test IOI Percent (n=281)	Control IOI Percent (n=290)	Test minus Control	P-Value
Private For-Profit Workers	23.6 (4.6)	16.7(4.5)	7.0 (6.8)	0.31
Private Not-For-Profit Workers	36.6 (9.1)	34.3 (11.1)	2.3 (15.0)	0.88
Local Government Workers	18.5 (8.3)	-29.6 (11.7)	-11.1 (15.0)	0.46
State Government Workers	34.9 (14.1)	44.9 (17.9)	-10.0 (20.9)	0.63
Federal Government Workers	32.6 (19.6)	12.3 (9.3)	20.4 (22.3)	0.36
Self-Employed Not Incorporated Workers	15.8 (6.4)	11.9 (6.2)	3.9 (9.2)	0.67
Self-Employed Incorporated Workers	26.6 (11.1)	39.9 (19.8)	-13.3 (23.5)	0.57
Unpaid Family Workers	33.7 (0.3)	27.8 (27.3)	5.9 (27.3)	0.83

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. Significance was tested based on a two-tailed t-test at the α=0.1 level.

Table D-8. Index of Inconsistency (IOI): CAPI

Category	Test IOI Percent (n=1,349)	Control IOI Percent (n=1,507)	Test minus Control	P-Value
Private For-Profit Workers	36.7 (2.6)	37.9 (3.0)	-1.3 (3.6)	0.72
Private Not-For-Profit Workers	56.0 (6.0)	55.9 (6.5)	0.1 (9.3)	0.99
Local Government Workers	25.7 (5.4)	31.8 (5.5)	-6.2 (8.8)	0.48
State Government Workers	51.8 (10.7)	39.4 (8.8)	12.4 (13.6)	0.36
Federal Government Workers	12.2 (4.5)	3.9 (2.4)	8.3 (5.4)	0.13
Self-Employed Not Incorporated Workers	40.4 (7.0)	51.1 (5.7)	-10.7 (9.5)	0.26
Self-Employed Incorporated Workers	32.5 (7.3)	42.2 (9.1)	-9.7 (11.5)	0.40
Unpaid Family Workers	-	-	-	-

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. Significance was tested based on a two-tailed t-test at the α=0.1 level. An entry of '-' in a cell indicates that no or too few observations were available to meet statistical standards for reliability.

Appendix E: Multiple Mark Responses in Test Treatment Tables for Mail Mode

Table E-1. Multiple Mark Responses – Second Marked Response after Selecting Private For-Profit Workers in Test Treatment

Category	Test Treatment (n=168)
Private Not-For-Profit Workers	7.1 (2.3)
Local Government Workers	16.8 (3.4)
State Government Workers	6.7 (2.5)
Active duty U.S. Armed Forces or Commissioned Corps	-
Federal Government Workers	6.4 (2.5)
Self-Employed Not Incorporated Workers	29.4 (5.1)
Self-Employed Incorporated Workers	30.8 (4.9)
Unpaid Family Workers	-
Total	100.0

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. An entry of '-' in a cell indicates that no or too few observations were available to meet statistical standards for reliability.