2022 American Community Survey Content Test Evaluation Report: Exploratory Analysis for Industry and Occupation and Class of Worker [1]

Anthony Martínez, Lindsay Longsine, and Michael Risley

U.S. Census Bureau

SEHSD Working Paper Number 2023-33

December 2023

EXECUTIVE SUMMARY

The U.S. Census Bureau conducted the 2022 American Community Survey (ACS) Content Test, from September through December of 2022. The 2022 ACS Content Test tested the wording, format, and placement of proposed new ACS questions and proposed revisions of current ACS questions for potential inclusion in the ACS data collection instruments. The tested questions came from 10 topics. The industry & occupation (I&O) and class of worker (COW) questions were not tested; however, changes to the Labor Force question may affect the results of this question series. This report presents the exploratory analysis of the data collected for the I&O and COW questions during the 2022 Content Test.

In preparation for the 2022 Content Test, the Census Bureau, in consultation with the Office of Management and Budget (OMB) and the Interagency Council on Statistical Policy Subcommittee on the ACS, determined which proposals solicited from over 25 federal agencies would be tested in 2022. Approved proposals for new content or changes to existing content were tested according to the ACS content change process, which includes cognitive testing and field testing.

The 2022 ACS Content Test consisted of a nationally representative sample of 120,000 housing unit addresses, excluding Puerto Rico, Alaska, and Hawaii. The sample, which was independent of production ACS, was divided evenly among three treatments, a Control treatment and two test treatments (referred to as the Test Version 1 and Test Version 2 treatments).

Like production ACS, the data collection for the 2022 ACS Content Test was conducted in two phases: a self-response phase, which lasted up to nine weeks, followed by a nonresponse followup phase, conducted via Computer-Assisted Personal Interviewing (CAPI). The CAPI operation lasted about one month. For households where we received a response in the original Content Test interview, a Content Follow-Up telephone reinterview was conducted to measure response error.

Significant changes were tested on the Labor Force and Income series of questions, including a change in the reference period. The Census Bureau is researching the use administrative records as a data source in the future for Income; this change in reference period would align ACS data to better match administrative records data usage. The COW and I&O series of questions are preceded by the Labor Force section and immediately followed by the Income series of questions. The COW and I&O section instructs respondents to report the type of employment "last week or the most recent employment in the past 5 years." Labor Force (weeks worked and usual hours worked questions) and Income tested a changed to their reference period from a rolling "past 12 months" to the previous calendar year in both the test treatments. Additionally, instructional changes for the Labor Force questions were tested.

The purpose of this report is to examine any impact on COW and I&O out of concern that a fixed calendar reference period throughout the employment and income sections could contribute to respondent confusion and misreporting their class of work, industry, and occupation status. The I&O and COW questions were not modified for the Content Test. The ACS publishes annual count estimates and median earnings estimates for full-time year-round workers by COW and I&O that would be impacted by this changing referencing period. Analysis on median earnings by occupation is included in the 2022 American Community Survey Content Test Evaluation Report: Income (Posey et al. 2023).

This research was guided by several questions concerning item missing data rates and differences in response distributions between treatments.

Item missing data rates:

- Overall, both test treatments produced higher missing data rates for the class of worker, industry, and occupation section of questions as a whole, suggesting that respondents skipped the entire section.
- Missing data rates and statistically significant differences varied by the treatment and mode at the individual question level. For example, missing data rates were higher for class of worker, industry, and occupation in the Test Version 1 treatment in the internet mode and overall. Overall, the Test Version 2 treatment only had a significantly higher missing data rates for class of worker.

Response distributions:

- There was no statistically significant difference in the distribution of full-time, year-round workers by North American Industry Classification System (NAICS) industry sectors between the Control and two test treatments.
- Although the chi-square test indicated that the distributions were marginally significantly different between the Test and Control treatments for the Standard Occupational Classification (SOC) major groups, the test can be unreliable for a large number of categories. Results from the subsequent two-sided t-tests showed no significant differences between individual SOC major groups between the Control and Test Version 1 treatments.
- There was no statistically significant difference in the distribution of full-time, year-round workers by SOC major groups between the Control and Test Version 2 treatments.

The use of administrative data for ACS Labor Force and Income estimates will undergo further testing and analysis. The results of this report will not be used in the recommendation of changing the reference period for Labor Force or Income, but it serves to inform data users about its

potential impact. If implemented, results suggest the change in reference period to a calendar year is not expected to lead to a significant difference in the distribution of full-time year-round workers among the NAICS industry sectors or the SOC major occupations based on the Content Test results for the Test Version 1 or Test Version 2 treatments.

1   BACKGROUND

The U.S. Census Bureau conducted the 2022 American Community Survey (ACS) Content Test from September to December of 2022. The 2022 ACS Content Test tested the wording, format, and placement of proposed new ACS questions and proposed revisions of current ACS questions for potential inclusion in the ACS data collection instruments. The questions came from these ten ACS topics, three of which, Sewer, Electric Vehicles, and Solar Panels were new:

- Household Roster
- Sewer
- Electric Vehicles
- Solar Panels
- Supplemental Nutrition Assistance Program (SNAP)
- Educational Attainment
- Health Insurance Coverage
- Disability
- Labor Force
- Income

The industry, occupation, and class of worker questions appeared directly after the Labor Force questions and directly before the Income questions. Significant changes were tested on the Labor Force series of questions, including a change in the reference period. The section following Labor Force was the industry and occupation (I&O) and class of worker (COW) series of questions, which used different reference periods based on employment in the last week or in the last five years. While the I&O and COW questions were not modified for the Content Test, the Census Bureau publishes estimates of full-time, year-round workers by class of worker, industry, and occupation, so we analyzed Content Test data to see if there was impact on the resulting estimates.

1.1 Proposals for New and Revised ACS Questions

In June 2018, the Census Bureau solicited proposals for new or revised ACS content from over 25 federal agencies. For new questions, the proposals explained why these data were needed and why other data sources that provide similar information were not sufficient. Proposals for new content were reviewed to ensure that the requests met a statutory or regulatory need for data at small geographic levels or for small populations.

The Census Bureau, in consultation with the Office of Management and Budget (OMB) and the Interagency Council on Statistical Policy Subcommittee on the ACS, determined which proposals moved forward. Approved proposals for new content or changes to current content were tested via the ACS content change process. This process includes cognitive testing and field

testing. An interagency team consisting of Census Bureau staff and representatives from other federal agencies participated in development and testing activities.

In accordance with OMB's Standards and Guidelines for Statistical Surveys (OMB, 2006) and the Census Bureau's Statistical Quality Standards (U.S. Census Bureau, 2022a), the Census Bureau conducted cognitive interviewing to pretest survey questions prior to field testing or implementing the questions in production.

1.2 Cognitive Testing

For the 2022 ACS Content Test, the Census Bureau contracted with Research Triangle Institute (RTI) International to conduct three rounds of cognitive testing.[2] Cognitive interviews were conducted virtually, in English and Spanish.[3] In the first round of cognitive testing, each topic tested one or two versions of the question. Based on the results of the first round, wording modifications to the questions were made and one or two versions per topic were tested in the second round. The interagency team used the results of both rounds of cognitive testing to recommend question content for the field test. For more information on the cognitive testing procedures and results from rounds one and two, see RTI International (2022a).

The third round of cognitive testing was conducted in Puerto Rico and in Group Quarters (GQ), as the 2022 ACS Content Test did not include field testing in these areas. Cognitive interviews in Puerto Rico were conducted in Spanish; GQ cognitive interviews were conducted in English. For more information on the cognitive testing procedures and results from the third round, see RTI International (2022b).

The I&O and COW series of questions were included in cognitive testing performed by RTI International. However, there were no wording modifications nor differing versions to I&O and COW questions between the cognitive testing rounds. The I&O and COW series of questions wording remained consistent with the wording used in the current production ACS questionnaire. During cognitive testing, interviewers did not probe respondents for further details on I&O and COW content. Lastly, the responses to the I&O and COW content were not included in any cognitive testing analyses.

Three topics included in the cognitive testing were not included in the field test: Homeowners Association or Condominium Fees, Home Heating Fuel, and Means of Transportation to Work. For the most part, the changes to these questions are expected to either impact a small population or result in a small change in the data that would not be detectable in the Content

---

[2] For each test topic, subcommittees were formed to develop question wording and research requirements for cognitive testing. The subcommittees included representation from the Census Bureau and other federal agencies.

[3] Cognitive testing interviews were conducted virtually due to the COVID-19 pandemic. Interviews were attempted by videoconferencing first and were moved to phone interviews if there were technical problems with Skype or MS Teams.

Test. The subject matter experts recommended that cognitive testing was sufficient for these questions and that field testing was not necessary; the Interagency Council on Statistical Policy Subcommittee on the ACS agreed with this recommendation. Content changes for these topics will be implemented in production ACS in 2024.

1.3 I&O and COW in the 2022 ACS Content Test

1.3.1 Question Content

Figure 1 is an image of the I&O and COW questions on the ACS paper questionnaire. The questions remained the same on all the Content Test questionnaires. Class of worker is obtained from question "a" in the series of questions. Responses from questions "b," "c," and "d" are used to determine industry while questions "e," and "f" are used to determine occupation.[4]

---

[4] Much of the analysis in this report used the coded responses, which are assigned based on the write-in responses. The coding process assigns one of 270 2017 census industry categories and one of 570 2018 census occupation categories, including Military. For more information on Census industry and occupation codes, see https://www.census.gov/topics/employment/industry-occupation/guidance.html /.

Figure 1. Industry and Occupation and Classof Worker (Questions) Paper

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?

c. What kind of business or industry was this? *Include the main activity, product, or service provided at the location where employed. (For example: elementary school, residential construction)*

d. Was this mainly – *Mark (X) ONE box.*

☐ manufacturing?

☐ wholesale trade?

☐ retail trade?

☐ other (agriculture, construction, service, government, etc.)?

e. What was this person's main occupation? *(For example: 4th grade teacher, entry-level plumber)*

f. Describe this person's most important activities or duties. *(For example: instruct and evaluate students and create lesson plans, assemble and install pipe sections and review building plans for work details)*

The Content Test mail questionnaire contained three versions of the Labor Force series of questions, the Control version (Figure 2) and Test Version 1 and Test Version 2 (Figure 3). Images of those also appear below. The Labor Force series of questions underwent several changes. Of primary concern for the I&O and COW section, is the change in reference period. For the weeks worked and usual hours worked questions, the reference period changed from a rolling "past 12 months" in the control version to a fixed calendar year in both test versions. A full detailed description of all the Labor Force question changes can be found in the [2022 American Community Survey Content Test Evaluation Report: Labor Force](#) (Mendez-Smith et al. 2023).

Figure 2. Control Version of the Labor Force Questions (Paper)

Figure 3. Test Version 1 (Left) and Test Version 2 (Right) of the Labor Force Questions (Paper)

When did this person last work for pay, even for a few days?

☐ Within the past 12 months

☐ 1 to 5 years ago

☐ Over 5 years ago or never worked → *SKIP to question 44*

In 2021, did this person work for pay, even for a few days?

☐ Yes

☐ No → *SKIP to question 43*

NOTE: For question 41a and b, include as WORK:

✓ all jobs for pay
✓ paid vacation
✓ paid sick leave
✓ military service

a. In 2021 (52 weeks), did this person work EVERY week? *Remember to include paid vacation and paid sick leave as work.*

☐ Yes → *SKIP to question 42*

☐ No

b. In 2021 (52 weeks), how many WEEKS did this person work for at least one day? *Include weeks when this person only worked for a few hours.*

Weeks

☐☐

In 2021, for the weeks worked, how many HOURS did this person usually work each WEEK? *Include all jobs for pay and military service.*

Usual hours worked each WEEK

☐☐☐

When did this person last work for pay, even for a few days?

☐ Within the past 12 months

☐ 1 to 5 years ago

☐ Over 5 years ago or never worked → *SKIP to question 44*

In 2021, did this person work for pay, even for a few days?

☐ Yes

☐ No → *SKIP to question 43*

a. In 2021 (52 weeks), did this person work EVERY week? *Count paid vacation, paid sick leave, and military service as work. Include all jobs for pay.*

☐ Yes → *SKIP to question 42*

☐ No

b. In 2021 (52 weeks), how many WEEKS did this person work for at least one day? *Include weeks when this person only worked for a few hours. Include all jobs for pay. Count paid vacation, paid sick leave, and military service as work.*

Weeks

☐☐

In 2021, for the weeks worked, how many HOURS did this person usually work each WEEK? *Include all jobs for pay and military service.*

Usual hours worked each WEEK

☐☐☐

1.3.2 Research Questions

The questions examined for this research are presented below. Control refers to the I&O and COW questions that were on the questionnaire with the Control version of the Labor Force questions. Similarly, Test Version 1 and Test Version 2 refer to the I&O and COW questions that accompanied those versions of the Labor Force questions, respectively.[5]

RQ1. Is there a difference in the item missing data rates for class of worker between the Control and Test treatments?

---

[5] The research questions were numbered 23 through 28 under the Additional Analyses for Class of Worker, Industry, and Occupation Question Series in the Labor Force section of the Research and Analysis Plan for the Content Test.

RQ2. Is there a difference in the item missing data rates for industry between the Control and test treatments?

RQ3. Is there a difference in the item missing data rates for occupation between the Control and test treatments?

RQ4. Is there a difference in the section missing data rates for all part of the class of worker, industry, and occupation series of questions between the Control and test treatments?

RQ5. Are the distributions of full-time, year-round workers by North American Industry Classification System (NAICS) industry sectors different between test treatments and the Control treatment?

RQ6. Are the distributions of full-time, year-round workers by major Standard Occupational Classification (SOC) System groups different between test treatments and the Control treatment?

2 METHODOLOGY

2.1 Sample Design

The 2022 ACS Content Test consisted of a national sample of roughly 120,000 housing unit addresses, excluding Puerto Rico, Alaska, and Hawaii (due to cost constraints, only stateside housing units were included). The sample was independent of the ACS production sample; however, the sample design for the Content Test was largely based on the ACS production sample design, with some modifications to meet the test objectives. The ACS production sample design is described in Chapter 4 of the ACS and Puerto Rico Community Survey (PRCS) Design and Methodology report (U.S. Census Bureau, 2022b).

The sample design modifications included stratifying addresses into high and low self-response areas, oversampling addresses from the low self-response areas to ensure equal response from both strata, and selecting an initial sample of addresses, followed by a nearest neighbor method for selecting the remaining addresses for sample. The high and low self-response strata were defined based on ACS self-response rates from the 2018 and 2019 panels at the tract level.

In the sample selection process, we selected an initial sample of 40,000 addresses, then selected the two nearest neighbors for each initially selected address. If possible, we selected nearest neighbors that were in both the same content test sampling stratum as well as the same state, county, and sub-county area as the initially selected address. In total, three samples were selected, one for the Control treatment and two for the two test treatments. These three treatments are shown in Table 1.

The Control treatment contained production questions and questions from the three new topics: Solar Panels, Electric Vehicles, and Sewer. The Test treatment contained a Test version question for all topics except Household Roster. Two of the new topics, Solar Panels and Sewer, only had one version of the test question; therefore, the same question was asked in the Control and test treatments. The other new topic, Electric Vehicles, had two versions; one was asked in the Control and Roster Test treatments and the other in the Test treatment. The I&O and COW series of questions used the production question wording across all 3 treatments of the Content Test.

The primary purpose of the Roster Test treatment was to test the household roster test question separately since changes in the amount and types of people included in the household could impact the results of person-level topics. Therefore, the analyses for Test Version 2 of the Health Insurance Coverage, Labor Force, and Income questions could have been impacted by these changes. However, it was determined that the additional information gained from testing an additional version of the topics in the Roster Test treatment was worth the risk.[6]

Table 1. Questions by Treatment

Topic	Control Treatment	Test Treatment	Roster Test Treatment
Household Roster	Production	Production	Test Version
Solar Panels	Test Version	Test Version	Test Version
Electric Vehicles	Test Version 1	Test Version 2	Test Version 1
Sewer	Test Version	Test Version	Test Version
Educational Attainment	Production	Test Version	Production
Health Insurance Coverage	Production	Test Version 1	Test Version 2
Disability	Production	Test Version	Production
SNAP	Production	Test Version	Test Version[†]
Labor Force	Production	Test Version 1	Test Version 2
Income	Production	Test Version 1	Test Version 2

† The SNAP Test Version was in both test treatments to align with Labor Force and Income that also had a reference period change to the previous calendar year.

---

[6] We examined differences in key household and person characteristics among the Control and Roster Test treatments to explore any indication of bias in the Health Insurance Coverage, Labor Force, and Income analyses. See Spiers et al. (2023) for more information.

2.2 Data Collection

The 2022 ACS Content Test occurred in parallel with data collection activities for the September 2022 ACS production panel. Data collection for production ACS data consists of two main phases: an approximately two-month self-response data collection phase and a one-month follow-up phase.

During the self-response phase, addresses in sample are asked to self-respond by internet or mail. The Census Bureau sends addresses in sample up to five mailings to encourage self-response. This operation is followed by a one-month Computer-Assisted Personal Interviewing (CAPI) operation, where Census Bureau field representatives attempt to complete a survey for a sub-sample of the remaining nonresponding addresses.

The following data collection protocols for the 2022 ACS Content Test remained the same as production ACS:

- Data were collected using the self-response modes of internet (in English and Spanish) and paper questionnaires for the first and second month of data collection.
- In the third month of data collection, a sub-sample of nonresponding addresses were selected for CAPI.
- During CAPI, Census Bureau field representatives conducted interviews in person and over the phone.
- Self-response via internet or paper was accepted throughout the three-month data collection period.

The following data collection protocols for the 2022 ACS Content Test differed from production ACS:

- There were no paper versions of the 2022 ACS Content Test questionnaires in Spanish.[7]
- If respondents called Telephone Questionnaire Assistance (TQA) and opted to complete the survey over the phone, the interviewers conducted the survey using the production ACS questionnaire.[8] Since the TQA interviews did not include test questions, they were excluded from the analysis of the 2022 ACS Content Test.
- The 2022 ACS Content Test did not include the Telephone Failed-Edit Follow-Up (FEFU) operation. In production, this operation follows up on households that provided

---

[7] In 2019, 412 Spanish questionnaires were mailed back out of all mailable cases. Based upon this rate, we projected that only 8 Spanish questionnaires would be mailed back in the 2022 Content Test, which would not be cost-effective.

[8] The interviewer did not know which treatment the caller was in and therefore administered the production questionnaire. In 2019, less than one percent (0.6%) of cases responded by TQA and had no other response in a different mode. Based upon this rate, we projected about 744 TQA-only responses would be excluded from the 2022 ACS Content Test analysis.

incomplete information on the form or reported more than five people on the roster of a paper questionnaire.[9]

- The 2022 ACS Content Test used a telephone reinterview component to measure response reliability or response bias (depending upon the ACS topic). This telephone reinterview operation is discussed in Section 2.3 below.

For detailed information about ACS data collection procedures, consult the ACS and PRCS Design and Methodology Report (U.S. Census Bureau, 2022b).

2.3 Content Follow-Up Operation

To measure response reliability or response bias, a Content Follow-Up (CFU) reinterview was attempted with every household with an original Content Test interview that met the CFU eligibility requirements. Among the requirements were that the household must be occupied, and the household must have a valid telephone number. See the CFU requirements document for the complete list of eligibility requirements (Spiers, 2021a).

The CFU data collection instrument included the questions being tested for the 2022 ACS Content Test and some production ACS questions for context.

The I&O and COW questions were not part of the CFU reinterview.

2.4 Analysis Metrics

The sample addresses for the Control and test treatments were selected in a manner so that their response propensities and response distributions (on particular characteristics) would be the same. Similar distributions allow us to conclude that any difference in the metrics used to analyze I&O and COW is attributable to the changes made to the Labor Force questions. We tested these unit-level assumptions in both the original interview and the CFU interview. See Section 2.4.1 for details. The metrics that we used to evaluate I&O and COW are presented in Section 2.4.2.

For the 2022 ACS Content Test, typical production ACS edits were not made because the primary concern of this test was how changes to existing questions and differences between versions of new questions affected the unaltered responses provided directly by respondents. For this reason, responses were not imputed either. A few edits were applied to the non-topic

---

[9] The information obtained from the FEFU improves accuracy in a production environment but confounds the evaluation of respondent behavior in the Content Test environment. For paper questionnaires, where the household size is six or more (up to 12), we only collected name, age, and sex of these additional persons, but not detailed information as we do in the FEFU operation for ACS production.

data, such as calculating a person's age based on his or her date of birth, but such edits were minimal. [10]

All estimates from the ACS Content Test were weighted. The final content test weights took into account the initial probability of selection (the base weight) and CAPI sub-sampling. The weights used in the CFU analysis also included an adjustment for CFU non-response.[11]

Comparisons between the Control and test versions were conducted using a two-tailed t-test at the α=0.1 level of significance. The Content Test sample size was chosen to provide enough statistical power (0.80) to detect a difference in the gross difference rates (measuring differences in adds and deletes from the household roster) of at least two percentage points between the Control and Roster Test groups for the Household Roster question.[12] In statistical tests involving multiple comparisons, we controlled for the overall Type I error rate by adjusting the resulting p-values using the Hochberg method (Hochberg, 1988).[13]

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, 2022b, Chapter 12). We calculated the variance for each rate and difference using the formula below. The standard error of an estimate (X_0) is the square root of the variance:

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

where:

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

2.4.1 Unit-Level Analysis

The unit response rate is important, as it provides an indication of the quality of the survey data. As part of our analysis, we examined unit-level (i.e., address-level) responses for the Control and test treatments in the original interviews and CFU reinterviews. These results are

---

[10] This only refers to edits made to the data sets before analysis. During the analysis phase, additional edits, such as collapsing categories, were made based on the needs of the individual question.

[11] The Content Test weight creation process does not include all the steps followed in the ACS, including the noninterview adjustment for the original interview and calibration to housing unit and population controls (see U.S. Census Bureau, 2022b, Chapter 11). For more information on the 2022 Content Test weighting procedure, see Risley and Oliver (2022) and Keathley (2022).

[12] See Section 2.4.2.4 for the definition of Gross Difference Rate.

[13] Use the MULTTEST Procedure in SAS®.

provided in a separate report (Spiers et al., 2023) since I&O and COW questions were not part of the CFU reinterview.[14]

2.4.2 Topic-Level Analysis

To evaluate whether the changes to the Labor Force questions affected the data for I&O and/or COW we calculated a variety of metrics, presented in Sections 2.4.2.1 through 2.4.2.6.

2.4.2.1 Item Missing Data Rates

To measure nonresponse to the I&O and COW questions, we calculated its item missing data rate, the proportion of eligible persons for which a required response is missing. A high item missing data rate can be indicative of a question that lacks clarity, is sensitive, or is simply too difficult to answer.

The universe for calculating item missing data rates for the I&O and COW questions was persons 15 years and over who (1) did any work last week, (2) worked within the past 12 months, or (3) worked 1-5 years ago.

The industry questions were: "What was the name of this person's employer, business, agency, or branch of the Armed Forces?" and "What kind of business or industry was this?" We considered a response to be missing only if both industry write-in questions were blank.

For industry, we calculated and compared item missing data rates for Test Version 1 vs. Control and Test Version 2 vs. Control. The rates were calculated for each data collection mode separately and for all modes combined.

The occupation questions were: "What was this person's main occupation?" and "Describe the person's main activities or duties." We considered a response to be missing only if both occupation write-in questions were blank.

For occupation, we calculated and compared item missing data rates for Test Version 1 vs. Control and Test Version 2 vs. Control. The rates were calculated for each data collection mode separately and for all modes combined.

The COW question was, "Which one of the following best describes this person's employment last week or the most recent employment in the past 5 years? Mark ONE box." For the purposes of this analysis, we counted mail mode responses where multiple (two or more) class of worker categories were selected (checked) as responses, even though the variable is blanked in normal ACS processing to create the unedited data file.

---

[14] As part of the 2022 ACS Content Test, we analyzed respondent burden. The result of this analysis is contained in Virgile et al. (2023).

For the COW question, we calculated and compared item missing rates for Test Version 1 vs. Control and Test Version 2 vs. Control. The rates were calculated for each data collection mode separately and for all modes combined.

We considered the entire section to be "missing" if there were no valid answers for any of the questions in the section. We calculated and compared the rates for Test Version 1 vs. Control and Test Version 2 vs. Control. The rates were calculated for each data collection mode separately and for all modes combined.

We compared item missing data rates via two-tailed t-tests.

2.4.2.2 Response Distributions

We compared each pair of distributions of full-time, year-round workers among the NAICS industry sectors, Control versus each test treatment. The NAICS industry sectors are defined by the 2017 North American Industry Classification System. The 27 industry sectors and their corresponding census industry codes are indicated in Table 2. Census Industry Codes Corresponding to NAICS Industry Codes below.

Table 2. Census Industry Codes Corresponding to NAICS Industry Codes

Description of NAICS Industry Sector	Range of Census Industry Codes
Agriculture, forestry, fishing and hunting	0170-0290
Mining quarrying, and oil and gas extraction	0370-0490
Construction	0770
Manufacturing	1070-3990
Wholesale trade	4070-4590
Retail trade	4670-5790
Transportation and warehousing	6070-6390
Utilities	0570-0690
Information	6470-6780
Finance and insurance	6870-6992
Real estate and rental and leasing	7071-7190
Professional, scientific, and technical services	7270-7490
Management of companies and enterprises	7570
Administrative and support and waste management services	7580-7790
Educational services	7860-7890
Health care and social assistance	7970-8470
Arts, entertainment, and recreation	8561-8590
Accommodation and food services	8660-8690
Other public services, except public administration	8770-9290
Public administration	9370-9590
Military	9670-9870

We compared each pair of distributions of full-time, year-round workers among the Standard Occupational Classification (SOC) major groups, Control versus each test treatment. The SOC major groups are defined by the 2018 Standard Occupational Classification Manual. The 23 occupation groups and their corresponding census occupation codes are indicated in Table 3.

Table 3. Census Occupation Codes Corresponding to SOC Major Groups

Description of SOC Major Group	Range of Census Occupation Codes
Management occupations	0010-0440
Business and financial operations occupations	0500-0960
Computer and mathematical occupations	1005-1240
Architecture and engineering occupations	1305-1560
Life, physical, and social science occupations	1600-1980
Community and social services occupations	2001-2060
Legal occupations	2100-2180
Education, training, and library occupations	2205-2555
Arts, design, entertainment, sports, and media occupations	2600-2970
Healthcare practitioner and technical occupations	3000-3550
Healthcare support occupations	3601-3655
Protective service occupations	3700-3960
Food preparation and serving related occupations	4000-4160
Building and grounds cleaning and maintenance occupations	4200-4255
Personal care and service occupations	4330-4655
Sales and related occupations	4700-4965
Office and administrative support occupations	5000-5940
Farming, fishing, and forestry occupations	6005-6130
Construction and extraction occupations	6200-6950
Installation, maintenance, and repair occupations	7000-7640
Production occupations	7700-8990
Transportation and materials moving occupations	9005-9760
Military specific occupations	9800-9830

We calculated the response distributions as the proportion of valid responses in a category to all valid responses. Comparisons were made using a Rao-Scott chi-square test that checks for a significant difference between two sample distributions (Rao & Scott, 1987). If the chi-square test indicated a significant difference between the Control and test distributions, we tested for significant differences in the individual category proportions using two-tailed t-tests.

3 DECISION CRITERIA

Before the field test, a team of subject matter experts identified and prioritized which of the research questions presented in Section 1.3.4 would determine which version of the Labor Force questions had minimal impact on the reporting for the I&O and COW questions. The decision criteria are presented in Table 4.

Table 4. Decision Criteria for Industry and Occupation and Class of Worker

Priority	Research Questions	Decision Criteria
1	1, 2, 3, 4	Item missing data rates: We hoped to see no difference or a decrease in item nonresponse.
2	5, 6	Response distributions: We hoped to see no difference in distributions of full-time, year-round workers for NAICS sectors and SOC major groups.

4 ASSUMPTIONS AND LIMITATIONS

4.1 Assumptions

- The sample addresses for the Control and test treatments were selected in a manner so that their response propensities and response distributions would be the same. This assumption of homogeneity allows us to conclude that any difference between treatments is attributable to differences in wording and format. See Section 5 for more details.

- There was no difference between treatments in mail delivery timing or subsequent response time. The treatments have the same sample size and use the same postal sort and mailout procedures. Previous research indicated that postal procedures alone could cause a difference in response rates at a given point in time between experimental treatments of different sizes, with response for the smaller treatments lagging (Heimel, 2016).

4.2 Limitations

- GQs were not included in the sample for the 2022 ACS Content Test. The results of the Content Test may not extend to GQ populations.

- Housing units from Alaska, Hawaii, and Puerto Rico were not included in the sample for the 2022 ACS Content Test. The results of the Content Test may not extend to the housing unit population in these areas.

- The paper questionnaire was only available in English and was not available in Spanish like in production. The Content Test results related to the English paper questionnaire may not extend to Spanish paper questionnaire.

- For paper questionnaires, where the household size was six or more (up to 12), we only collected name, age, and sex of these additional persons. Detailed information for these persons in ACS production are collected in the FEFU operation. We did not include the FEFU operation because the information collected from it improves accuracy and could confound respondent behavior in the Content Test environment.

- We did not have response data for some partial internet responses (179 cases) due to a server issue. These cases were excluded from the analyses.

- TQA responses were excluded from the analysis of the 2022 ACS Content Test response data because survey responses completed via the TQA operation were only conducted using the ACS production data collection instrument.

- CAPI interviewers were assigned 2022 ACS Content Test cases as well as regular production cases. The potential risk of this approach is the introduction of a cross-contamination or carry-over effect among Control and test treatments and production due to the same interviewer administering multiple versions of the same question item (despite their training to read questions verbatim).

- Due to budget constraints, the CAPI workload could not exceed 28,000 housing units. This workload was less than what was subsampled originally because we over-sampled addresses in low response areas. Limiting the CAPI workload caused an increase in the variances for the analysis metrics used.

- The 2022 ACS Content Test did not include the production weighting adjustments for unit nonresponse or population controls which are designed to minimize nonresponse and under-coverage bias. As a result, any estimates derived from the Content Test data did not provide the same level of inference as the production ACS and cannot be compared to production estimates.

5 RESULTS

This section of the report presents the results of various metrics used to evaluate the effects of changes to the Labor Force questions on the I&O and COW questions. The comparisons presented assume homogeneity of the response distributions for the three treatments, prior to the field test. We tested this assumption via unit-level (i.e., address level) analyses. The results are presented in (Spiers et al., 2023).

Original Interview

The overall unit response rates were not significantly different between the treatments, nor were the response rate portions by mode. When looking at response rates within high and low response areas, a couple of modal comparisons were significant, but these results did not appear in the overall comparisons.

Additionally, when examining demographic and socioeconomic distributions, none of the response distributions were significantly different between treatments. When looking at distributions among self-responses and CAPI responses, only the distribution for race among CAPI responses for the Control and Test treatments was significantly different. This distribution difference showed up in the Other Race Only category.

We are confident there were no statistically significant differences that would impact original interview comparisons between treatments for the COW and I&O questions.

For more information about the unit-level analyses, see Spiers et al. (2023).

Respondent Burden Analysis

Measures of respondent burden were also analyzed for the 2022 Content Test. Metrics of interest were completion times, help screen access rates, breakoff rates, and form completeness.

The internet breakoff rate for the Test treatment was higher than the Control treatment. The internet breakoff rate for the Roster treatment was also higher than the Control treatment. The form completeness rate in the internet mode for the Control treatment in the detailed persons section of the ACS was higher than both Test and Roster.

For more information on the respondent burden analysis see Virgile et al. (2023).

5.1 Item Missing Data Rate Results for I&O and COW

RQ1. Is there a difference in the item missing data rates for class of worker between the Control and test treatments?

Table 5 provides the item missing data rates of the COW question for the Control and Test Version 1 treatments of the Content Test. We calculated the rates overall and by mode. We compared the item missing data rates using a two-sided t-test.

Table 5. Item Missing Data Rates for the COW question - Control vs Test Version 1

Mode	Test Version 1 Percent	Control Percent	Difference	Adjusted P-value
Overall	8.1 (0.3)	6.8 (0.3)	1.3 (0.4)	<0.01*
Internet	8.7 (0.3)	6.7 (0.3)	2.0 (0.4)	<0.01*
Mail	15.7 (1.1)	16.5 (1.1)	-0.9 (1.5)	0.63
CAPI	2.3 (0.4)	2.1 (0.3)	0.3 (0.5)	0.63

Source: U.S. Census Bureau, 2022 American Community Survey Content Test | DRB No. CBDRB-FY23-ACSO003-B0072
Note: Minor additive discrepancies are due to rounding. Standard errors are in parentheses. Significance was tested using a two-tailed t-test at the α=0.1 level. An asterisk (*) indicates a statistically significant result. P-values have been adjusted for multiple comparisons using the Hochberg method.

The Test Version 1 had higher item missing data rates than Control in the internet mode and overall for the COW question.

Table 6 provides the item missing data rates of the COW question for the Control and Test Version 2 treatments of the Content Test.

Table 6. Item Missing Data Rates for the COW question - Control vs Test Version 2

Mode	Test Version 2 Percent	Control Percent	Difference	Adjusted P-value
Overall	7.6 (0.3)	6.8 (0.3)	0.8 (0.4)	0.09*
Internet	8.4 (0.3)	6.7 (0.3)	1.6 (0.4)	<0.01*
Mail	14.3 (0.9)	16.5 (1.1)	-2.2 (1.3)	0.16
CAPI	2.0 (0.3)	2.1 (0.3)	-0.1 (0.4)	0.82

Source: U.S. Census Bureau, 2022 American Community Survey Content Test | DRB No. CBDRB-FY23-ACSO003-B0072
Note: Minor additive discrepancies are due to rounding. Standard errors are in parentheses. Significance was tested using a two-tailed t-test at the α=0.1 level. An asterisk (*) indicates a statistically significant result. P-values have been adjusted for multiple comparisons using the Hochberg method.

The item missing data rate of the COW question was higher for Test Version 2 than Control in the internet mode and overall.

RQ2. Is there a difference in the item missing data rates for industry between the Control and test treatments?

Table 7 provides the item missing data rates of the industry questions for the Control and Test Version 1 treatments. As previously explained, we considered a response to be missing only if both industry write-in questions were blank.

Table 7. Item Missing Data Rates for the Industry questions – Control vs Test Version 1

Mode	Test Version 1 Percent	Control Percent	Difference	Adjusted P-value
Overall	8.0 (0.3)	7.2 (0.3)	0.8 (0.4)	0.07*
Internet	8.8 (0.4)	7.6 (0.3)	1.3 (0.5)	0.02*
Mail	12.7 (1.1)	13.1 (1.1)	-0.4 (1.5)	0.89
CAPI	2.9 (0.4)	2.8 (0.4)	0.1 (0.6)	0.89

Source: U.S. Census Bureau, 2022 American Community Survey Content Test | DRB No. CBDRB-FY23-ACSO003-B0072
Note: Minor additive discrepancies are due to rounding. Standard errors are in parentheses. Significance was tested using a two-tailed t-test at the α=0.1 level. An asterisk (*) indicates a statistically significant result. P-values have been adjusted for multiple comparisons using the Hochberg method.

The item missing data rate of the industry questions was higher for Test Version 1 than Control in the internet mode and overall.

Table 8 provides the item missing data rates of the industry questions for the Control and Test Version 2 treatments.

Table 8. Item Missing Data Rates for the Industry questions - Control vs Test Version 2

Mode	Test Version 2 Percent	Control Percent	Difference	Adjusted P-value
Overall	7.9 (0.3)	7.2 (0.3)	0.7 (0.4)	0.15
Internet	9.0 (0.4)	7.6 (0.3)	1.5 (0.5)	0.01*
Mail	11.0 (0.8)	13.1 (1.1)	-2.2 (1.3)	0.21
CAPI	3.0 (0.4)	2.8 (0.4)	0.2 (0.6)	0.68

Source: U.S. Census Bureau, 2022 American Community Survey Content Test | DRB No. CBDRB-FY23-ACSO003-B0072
Note: Minor additive discrepancies are due to rounding. Standard errors are in parentheses. Significance was tested using a two-tailed t-test at the α=0.1 level. An asterisk (*) indicates a statistically significant result. P-values have been adjusted for multiple comparisons using the Hochberg method.

The item missing data rate of the industry questions was higher for Test Version 2 than Control in the internet mode.

RQ3. Is there a difference in the item missing data rates for occupation between the Control and test treatments?

Table 9 provides the item missing data rates of the occupation questions for the Control and Test Version 1 treatments. As previously explained, we considered a response to be missing only if both occupation write-in questions were blank.

Table 9. Item Missing Data Rates for the Occupation questions - Control vs Test Version 1

Mode	Test Version 1 Percent	Control Percent	Difference	Adjusted P-value
Overall	9.5 (0.3)	8.6 (0.3)	0.9 (0.4)	0.09*
Internet	11.1 (0.4)	9.6 (0.4)	1.5 (0.5)	0.01*
Mail	9.6 (0.7)	11.3 (1.0)	-1.7 (1.2)	0.30
CAPI	4.2 (0.5)	3.9 (0.5)	0.3 (0.7)	0.70

Source: U.S. Census Bureau, 2022 American Community Survey Content Test | DRB No. CBDRB-FY23-ACSO003-B0072
Note: Minor additive discrepancies are due to rounding. Standard errors are in parentheses. Significance was tested using a two-tailed t-test at the α=0.1 level. An asterisk (*) indicates a statistically significant result. P-values have been adjusted for multiple comparisons using the Hochberg method.

The item missing data rate of the occupation questions was higher for Test Version 1 than Control in the internet mode and overall.

Table 10 provides the item missing data rates of the occupation questions for the Control and Test Version 2 treatments.

Table 10. Item Missing Data Rates for the Occupation questions - Control vs Test Version 2

Mode	Test Version 2 Percent	Control Percent	Difference	Adjusted P-value
Overall	9.1 (0.3)	8.6 (0.3)	0.5 (0.5)	0.51
Internet	10.8 (0.4)	9.6 (0.4)	1.1 (0.6)	0.16
Mail	9.8 (0.7)	11.3 (1.0)	-1.5 (1.1)	0.49
CAPI	3.9 (0.5)	3.9 (0.5)	-0.1 (0.7)	0.91

Source: U.S. Census Bureau, 2022 American Community Survey Content Test | DRB No. CBDRB-FY23-ACSO003-B0072
Note: Minor additive discrepancies are due to rounding. Standard errors are in parentheses. Significance was tested using a two-tailed t-test at the α=0.1 level. An asterisk (*) indicates a statistically significant result. P-values have been adjusted for multiple comparisons using the Hochberg method.

There was no statistically significant difference in the item missing data rates of the occupation questions between Test Version 2 and Control.

RQ4. Is there a difference in the section missing data rates for all parts of the class of worker, industry, and occupation series of questions between the Control and test treatments?

Viewing the series COW and I&O as an entire section is necessary as these topics are interrelated. Responses to any of these questions is used by clerical coders to determine or to correct class of worker category and/or the best code for industry and occupation. For example, if COW is missing, a clerical coder can determine the appropriate COW category through information provided for industry in the employer name and kind of business questions. This analysis considered the entire section missing when no information was provided for any of COW or I&O questions.

Table 11 provides the missing data rates of all parts of the COW and I&O section of the ACS for the Control and Test Version 1 treatments.

Table 11. Section Missing Data Rates - Control vs Test Version 1

Mode	Test Version 1 Percent	Control Percent	Difference	Adjusted P-value
Overall	5.4 (0.2)	4.5 (0.2)	0.9 (0.3)	<0.01*
Internet	6.5 (0.3)	5.0 (0.2)	1.5 (0.3)	<0.01*
Mail	5.9 (0.6)	7.7 (0.8)	-1.9 (1.0)	0.12
CAPI	1.6 (0.3)	1.4 (0.3)	0.2 (0.4)	0.60

Source: U.S. Census Bureau, 2022 American Community Survey Content Test | DRB No. CBDRB-FY23-ACSO003-B0072
Note: Minor additive discrepancies are due to rounding. Standard errors are in parentheses. Significance was tested using a two-tailed t-test at the $\alpha=0.1$ level. An asterisk (*) indicates a statistically significant result. P-values have been adjusted for multiple comparisons using the Hochberg method.

The section missing data rate was higher for Test Version 1 than Control in the internet mode and overall.

Table 12 provides the missing data rates of all parts of the COW and I&O section of the ACS for the Control and Test Version 2 treatments.

Table 12. Section Missing Data Rates - Control vs Test Version 2

Mode	Test Version 2 Percent	Control Percent	Difference	Adjusted P-value
Overall	5.2 (0.2)	4.5 (0.2)	0.7 (0.3)	0.05*
Internet	6.3 (0.3)	5.0 (0.2)	1.3 (0.3)	<0.01*
Mail	6.2 (0.6)	7.7 (0.8)	-1.5 (1.0)	0.24
CAPI	1.4 (0.3)	1.4 (0.3)	<0.1 (0.4)	0.94

Source: U.S. Census Bureau, 2022 American Community Survey Content Test | DRB No. CBDRB-FY23-ACSO003-B0072
Note: Minor additive discrepancies are due to rounding. Standard errors are in parentheses. Significance was tested using a two-tailed t-test at the $\alpha=0.1$ level. An asterisk (*) indicates a statistically significant result. P-values have been adjusted for multiple comparisons using the Hochberg method.

The section missing data rate was higher for Test Version 2 than Control in the internet mode and overall.

5.2 Response Distribution Results for I &O and COW

RQ5. Are the distributions of full-time, year-round workers by North American Industry Classification System (NAICS) industry sectors different between test treatments and the Control Treatment?

Table 13 shows the distribution of full-time, year-round workers by NAICS industry sectors for the Control and Test Version 1 treatments. We compared the distributions using a chi-square test.

Table 13. Distribution of full-time, year-round workers by NAICS industry sectors – Control vs Test Version 1

Industry Category	Test Version 1 Percent	Control Percent	Chi-square	P-value
Agriculture, forestry, fishing and hunting	0.7 (0.1)	1.0 (0.2)	25.8	0.17
Mining quarrying, and oil and gas extraction	0.5 (0.1)	0.4 (0.1)		
Construction	7.5 (0.4)	7.0 (0.4)		
Manufacturing	12.2 (0.5)	12.3 (0.5)		
Wholesale trade	2.4 (0.2)	2.6 (0.2)		
Retail trade	9.5 (0.5)	8.7 (0.5)		
Transportation and warehousing	4.6 (0.3)	4.9 (0.3)		
Utilities	1.1 (0.1)	1.1 (0.1)		
Information	2.1 (0.2)	2.8 (0.2)		
Finance and insurance	6.0 (0.3)	6.6 (0.4)		
Real estate and rental and leasing	2.0 (0.2)	2.0 (0.2)		
Professional, scientific, and technical services	10.2 (0.4)	10.2 (0.4)		
Management of companies and enterprises	0.1 (0.1)	0.1 (<0.1)		
Administrative and support and waste management	3.3 (0.2)	3.1 (0.3)		
Educational services	8.6 (0.4)	8.1 (0.4)		
Health care and social assistance	14.3 (0.5)	13.5 (0.5)		
Arts, entertainment, and recreation	1.3 (0.2)	1.2 (0.1)		
Accommodation and food services	3.3 (0.2)	4.0 (0.3)		
Other public services, except public administration	4.0 (0.3)	3.7 (0.3)		
Public administration	5.5 (0.4)	5.3 (0.3)		
Military	0.7 (0.1)	1.0 (0.2)		

Source: U.S. Census Bureau, 2022 American Community Survey Content Test | DRB No. CBDRB-FY23-ACSO003-B0072
Note: Minor additive discrepancies are due to rounding. Standard errors are in parentheses. Significance was tested using a chi-square test at the α=0.1 level.

There was no statistically significant difference in the distribution of full-time, year-round workers by NAICS industry sectors between the Control and Test Version 1 treatments.

Table 14 shows the distribution of full-time, year-round workers by NAICS industry sectors for the Control and Test Version 2 treatments.

Table 14. Distribution of full-time, year-round workers by NAICS industry sectors – Control vs Test Version 2

Industry Category	Test Version 2 Percent	Control Percent	Chi-square	P-value
Agriculture, forestry, fishing and hunting	1.0 (0.1)	1.0 (0.2)	19.3	0.51
Mining quarrying, and oil and gas extraction	0.3 (0.1)	0.4 (0.1)		
Construction	7.7 (0.4)	7.0 (0.4)		
Manufacturing	11.0 (0.4)	12.3 (0.5)		
Wholesale trade	2.8 (0.2)	2.6 (0.2)		
Retail trade	8.7 (0.4)	8.7 (0.5)		
Transportation and warehousing	5.1 (0.3)	4.9 (0.3)		
Utilities	1.1 (0.1)	1.1 (0.1)		
Information	2.6 (0.2)	2.8 (0.2)		
Finance and insurance	6.5 (0.3)	6.6 (0.4)		
Real estate and rental and leasing	2.0 (0.2)	2.0 (0.2)		
Professional, scientific, and technical services	9.4 (0.4)	10.2 (0.4)		
Management of companies and enterprises	0.1 (<0.1)	0.1 (<0.1)		
Administrative and support and waste management	3.4 (0.2)	3.1 (0.3)		
Educational services	8.5 (0.4)	8.1 (0.4)		
Health care and social assistance	14.4 (0.5)	13.5 (0.5)		
Arts, entertainment, and recreation	1.2 (0.1)	1.2 (0.1)		
Accommodation and food services	3.2 (0.2)	4.0 (0.3)		
Other public services, except public administration	4.3 (0.3)	3.7 (0.3)		
Public administration	6.0 (0.3)	5.3 (0.3)		
Military	0.8 (0.1)	1.0 (0.2)		

Source: U.S. Census Bureau, 2022 American Community Survey Content Test | DRB No. CBDRB-FY23-ACSO003-B0072
Note: Minor additive discrepancies are due to rounding. Standard errors are in parentheses. Significance was tested using a chi-square test at the $\alpha=0.1$ level.

There was no statistically significant difference in the distribution of full-time, year-round workers by NAICS industry sectors between the Control and Test Version 2.

RQ6. Are the distributions of full-time, year-round workers by major Standard Occupational Classification (SOC) System groups different between test treatments and the Control treatment?

Table 15 shows the distribution of full-time, year-round workers by SOC major groups for the Control and Test Version 1 treatments.

Table 15. Distribution of full-time, year-round workers by SOC major groups – Control vs Test Version 1

Occupation Category	Test Version 1 Percent	Control Percent	Chi-square	P-value
Management occupations	15.3 (0.5)	16.1 (0.5)	33.3	0.06*
Business and financial operations occupations	7.6 (0.4)	8.2 (0.4)		
Computer and mathematical occupations	5.5 (0.3)	5.0 (0.3)		
Architecture and engineering occupations	2.9 (0.2)	3.0 (0.2)		
Life, physical, and social science occupations	1.5 (0.2)	1.1 (0.1)		
Community and social services occupations	2.0 (0.2)	1.6 (0.2)		
Legal occupations	1.4 (0.1)	1.6 (0.2)		
Education, training, and library occupations	5.4 (0.3)	5.5 (0.3)		
Arts, design, entertainment, sports, and media	1.4 (0.1)	2.0 (0.2)		
Healthcare practitioner and technical occupations	7.2 (0.3)	6.5 (0.3)		
Healthcare support occupations	2.5 (0.2)	2.1 (0.2)		
Protective service occupations	2.0 (0.2)	1.9 (0.2)		
Food preparation and serving related occupations	2.1 (0.2)	2.6 (0.2)		
Building and grounds cleaning and maintenance	2.1 (0.2)	2.4 (0.2)		
Personal care and service occupations	1.2 (0.1)	1.5 (0.2)		
Sales and related occupations	8.2 (0.4)	8.3 (0.4)		
Office and administrative support occupations	10.5 (0.4)	10.0 (0.4)		
Farming, fishing, and forestry occupations	0.3 (0.1)	0.2 (0.1)		
Construction and extraction occupations	5.2 (0.3)	4.7 (0.3)		
Installation, maintenance, and repair occupations	3.5 (0.2)	3.4 (0.3)		
Production occupations	5.7 (0.3)	5.3 (0.3)		
Transportation and material moving occupations	6.1 (0.3)	6.8 (0.5)		
Military specific occupations	0.3 (0.1)	0.4 (0.2)		

Source: U.S. Census Bureau, 2022 American Community Survey Content Test | DRB No. CBDRB-FY23-ACSO003-B0072

Note: Minor additive discrepancies are due to rounding. Standard errors are in parentheses. Significance was tested using a chi-square test at the $\alpha=0.1$ level. An asterisk (*) indicates a statistically significant result.

The chi-square test indicated that the distributions are significantly different, however, this type of test can be unreliable for items with a large number of categories. We also compared the distributions using two-sided t-tests.

Table 16 shows the t-test comparisons of full-time, year-round workers by SOC major groups for the Control and Test Version 1 treatments.

Table 16. T-tests for full-time, year-round workers by SOC major groups – Control vs Test Version 1

Occupation Category	Test Version 1 Percent	Control Percent	Difference	Adjusted P-value
Management occupations	14.8 (0.5)	15.5 (0.5)	-0.7 (0.7)	0.92
Business and financial operations occupations	7.3 (0.3)	7.8 (0.4)	-0.6 (0.5)	0.92
Computer and mathematical occupations	5.3 (0.3)	4.8 (0.3)	0.5 (0.4)	0.92
Architecture and engineering occupations	2.8 (0.2)	2.9 (0.2)	<0.1 (0.3)	0.92
Life, physical, and social science occupations	1.5 (0.2)	1.1 (0.1)	0.4 (0.2)	0.69
Community and social services occupations	1.9 (0.2)	1.6 (0.2)	0.3 (0.2)	0.92
Legal occupations	1.3 (0.1)	1.5 (0.2)	-0.2 (0.2)	0.92
Education, training, and library occupations	5.2 (0.3)	5.3 (0.3)	-0.1 (0.4)	0.92
Arts, design, entertainment, sports, and media	1.3 (0.1)	1.9 (0.2)	-0.6 (0.2)	0.34
Healthcare practitioner and technical occupations	7.0 (0.3)	6.3 (0.3)	0.7 (0.4)	0.92
Healthcare support occupations	2.5 (0.2)	2.0 (0.2)	0.5 (0.2)	0.92
Protective service occupations	1.9 (0.2)	1.8 (0.2)	0.1 (0.2)	0.92
Food preparation and serving related occupations	2.0 (0.2)	2.5 (0.2)	-0.5 (0.3)	0.92
Building and grounds cleaning and maintenance	2.0 (0.2)	2.3 (0.2)	-0.3 (0.3)	0.92
Personal care and service occupations	1.2 (0.1)	1.5 (0.2)	-0.3 (0.2)	0.92
Sales and related occupations	7.9 (0.4)	8.0 (0.4)	-0.1 (0.5)	0.92
Office and administrative support occupations	10.1 (0.4)	9.6 (0.4)	0.5 (0.6)	0.92
Farming, fishing, and forestry occupations	0.3 (0.1)	0.2 (0.1)	0.1 (0.1)	0.92
Construction and extraction occupations	5.0 (0.3)	4.5 (0.3)	0.5 (0.4)	0.92
Installation, maintenance, and repair occupations	3.4 (0.2)	3.2 (0.3)	0.2 (0.3)	0.92
Production occupations	5.5 (0.3)	5.1 (0.3)	0.4 (0.4)	0.92
Transportation and material moving occupations	5.9 (0.3)	6.5 (0.4)	-0.6 (0.5)	0.92
Military specific occupations	0.3 (0.1)	0.3 (0.1)	-0.1 (0.2)	0.92

Source: U.S. Census Bureau, 2022 American Community Survey Content Test | DRB No. CBDRB-FY23-ACSO003-B0072
Note: Minor additive discrepancies are due to rounding. Standard errors are in parentheses. Significance was tested using a two-tailed t-test at the $\alpha=0.1$ level. P-values have been adjusted for multiple comparisons using the Hochberg method.

The t-tests indicate there was no statistically significant difference for any category in the distribution of full-time, year-round workers by SOC major groups between the Control and Test Version 1 treatments.

Table 17 shows the distribution of full-time, year-round workers by SOC major groups for the Control and Test Version 2 treatments.

Table 17. Distribution of full-time, year-round workers by SOC major groups – Control vs Test Version 2

Occupation Category	Test Version 2 Percent	Control Percent	Chi-square	P-value
Management occupations	16.2 (0.5)	16.1 (0.5)	13.5	0.92
Business and financial operations occupations	7.6 (0.4)	8.2 (0.4)		
Computer and mathematical occupations	5.0 (0.3)	5.0 (0.3)		
Architecture and engineering occupations	3.0 (0.2)	3.0 (0.2)		
Life, physical, and social science occupations	1.4 (0.2)	1.1 (0.1)		
Community and social services occupations	1.9 (0.2)	1.6 (0.2)		
Legal occupations	1.4 (0.1)	1.6 (0.2)		
Education, training, and library occupations	5.3 (0.3)	5.5 (0.3)		
Arts, design, entertainment, sports, and media	1.7 (0.2)	2.0 (0.2)		
Healthcare practitioner and technical occupations	7.1 (0.3)	6.5 (0.3)		
Healthcare support occupations	2.7 (0.2)	2.1 (0.2)		
Protective service occupations	2.1 (0.2)	1.9 (0.2)		
Food preparation and serving related occupations	2.2 (0.2)	2.6 (0.2)		
Building and grounds cleaning and maintenance	2.1 (0.2)	2.4 (0.2)		
Personal care and service occupations	1.3 (0.1)	1.5 (0.2)		
Sales and related occupations	8.4 (0.4)	8.3 (0.4)		
Office and administrative support occupations	9.9 (0.4)	10.0 (0.4)		
Farming, fishing, and forestry occupations	0.2 (0.1)	0.2 (0.1)		
Construction and extraction occupations	4.7 (0.3)	4.7 (0.3)		
Installation, maintenance, and repair occupations	3.3 (0.2)	3.4 (0.3)		
Production occupations	5.4 (0.3)	5.3 (0.3)		
Transportation and material moving occupations	6.7 (0.3)	6.8 (0.5)		
Military specific occupations	0.4 (0.1)	0.4 (0.2)		

Source: U.S. Census Bureau, 2022 American Community Survey Content Test | DRB No. CBDRB-FY23-ACSO003-B0072

Note: Minor additive discrepancies are due to rounding. Standard errors are in parentheses. Significance was tested using a chi-square test at the α=0.1 level.

There was no statistically significant difference in the distribution of full-time, year-round workers by SOC major groups between the Control and Test Version 2 treatments.

6 CONCLUSIONS

The Industry and Occupation Statistics Branch (IOSB) requested to have content included in the 2022 Content Test. It was important to examine any impact that a fixed calendar reference period in the employment and income sections could have on respondent confusion and misreporting of their class of work, industry, or occupation.

Substantial changes were tested on the Labor Force and Income series of questions, including a change in the reference period. The COW and I&O series of questions are preceded by the Labor Force section and immediately followed by the Income series of questions. The COW and I&O section instructs respondents to report the type of employment "last week or the most recent employment in the past 5 years." Labor Force (weeks worked and usual hours worked questions) and Income tested a changed to their reference period from a rolling "past 12 months" to the previous calendar year in both test treatments. This change in reference period, if implemented, would align ACS data to better match administrative records data (the Census Bureau is currently researching the use administrative records as a data source in the future).

This reference period modification would affect the annual count estimates and median earnings estimates of full-time, year-round workers by COW and I&O. This exploratory analysis provides an initial assessment of the overall impact if a reference period change is implemented. Analysis on the median earnings by occupation is included in the 2022 American Community Survey Content Test Evaluation Report: Income (Posey et al. 2023).

The primary decision criteria concerned item missing data rates. We expected the item missing data rates for the test treatments to be no different from or lower than the Control treatment, for all modes.

For class of worker, the Test Version 1 and Test Version 2 treatments each had a higher overall missing data rate when individually compared with the Control treatment. By mode, the Test Version 1 and Test Version 2 treatments had a statistically significant higher missing rate within internet mode as well.

The Test Version 1 treatment also had a significantly higher data missing rate for industry, overall and within internet mode, than the Control treatment. For the Test Version 2 treatment, overall, there was not a significant difference in the item missing data rate compared with the Control treatment. However, when examined by mode, the Test Version 2 treatment did have a significantly higher item missing data rate within internet mode.

The item missing data rate for occupation was significantly higher for the Test Version 1 treatment, overall and for internet, compared with the Control treatment. Conversely, there was

no significant difference in the item missing data rates for occupation between Test Version 2 and Control treatments.

This analysis also looked at the item missing data rates of the COW and I&O section of questions. Information provided in any of these questions is used by clerical coders to determine the correct class of worker category and the best code for industry and occupation. For example, if COW is missing, a clerical coder can determine the appropriate COW category through information provided for industry. This analysis considered the entire section missing when no information was provided for any of COW or I&O questions.

The section missing data rate was higher for Test Version 1 and for Test Version 2 when compared with the Control treatment. In each instance, the missing data rate was higher overall and for the internet mode.

We also analyzed the response distribution of full-time, year-round workers by NAICS industry sectors and by SOC major occupation groupings. The criteria for the response distributions was to see no difference in distributions of full-time, year-round workers.

For industry, there were no differences in the distribution of full-time, year-round workers by NAICS industry sectors between the Control and Test Version 1 or Control and Test Version 2 treatments.

The chi-square test indicated the full-time year-round distributions for the SOC major groups were marginally significantly different between the Test Version 1 and Control treatments. The subsequent two-tailed t-tests showed no statistically significant differences between the individual occupation major groups. For the Test Version 2 treatment, there was no statistically significant difference in the SOC response distribution when compared with the Control treatment.

The use of administrative data for ACS Labor Force and Income estimates will undergo further testing and analysis. The results of this report will not be used in the recommendation of changing the reference period for Labor Force or Income, but it serves to inform data users about its potential impact. If implemented, results suggest the change in reference period to a calendar year is not expected to lead to a significant difference in the distribution of full-time year-round workers among the NAICS industry sectors or the SOC major occupations based on the Content Test results for the Test Version 1 or Test Version 2 treatments.

7   ACKNOWLEDGMENTS

The 2022 ACS Content Test would not have been possible without the participation and assistance of many individuals from the Census Bureau and other agencies. We sincerely appreciate their contributions.

- Census Bureau staff in the American Community Survey Office, Application Development and Services Division, Center for Behavioral Science Methods, Decennial Information Technology Division, Decennial Statistical Studies Division, Field Division, National Processing Center, Population Division, and Social, Economic, and Housing Statistics Division.
- Representatives from other agencies in the Federal statistical system serving on OMB's Interagency Working Group for the ACS and the topical subcommittees formed by the Working Group for each topic tested in the 2022 ACS Content Test.
- Staff in OMB's Statistical and Science Policy Office.

8   REFERENCES

Heimel, S. (2016). Postal tracking research on the May 2015 ACS panel. U.S. Census Bureau. 2016 American Community Survey Research and Evaluation Report Memorandum Series #ACS16-RER-01.

Hochberg, Y. (1988). A Sharper Bonferroni Procedure for Multiple Tests of Significance. *Biometrika*, 75 (4), 800-802. Retrieved January 17, 2017, from https://doi.org/10.2307/2336325

Keathley, D. (2022). *Specifications for selecting and weighting the 2022 American Community Survey Content Test sample.* U.S. Census Bureau. DSSD 2022 American Community Survey Memorandum Series #ACS22-S-21.

Mendez-Smith, R., Gumber, C., Longsine, L., & Risley, M. (2023). *2022 American Community Survey Content Test Evaluation Report: Labor Force*. U.S. Census Bureau. Retrieved December 6, 2023, from https://www.census.gov/library/working-papers/2023/acs/2023_Gumber_01.html

Office of Management and Budget (2006). *Standards and guidelines for statistical surveys*. Retrieved February 24, 2022, from https://www.whitehouse.gov/wpcontent/uploads/2021/04/standards_stat_surveys.pdf

Posey, K. G., Unrath, M., Risley, M. & Massarone, P. (2023). *2022 American Community Survey Content Test Evaluation Report: Income*. U.S. Census Bureau. Retrieved December 6, 2023, from https://www.census.gov/library/working-papers/2023/acs/2023_Posey_01.html

Rao, J., & Scott A. (1987). On simple adjustments to chi-square tests with sample survey data. *The Annals of Statistics*, 15 (1), 1987, 385–397. https://doi.org /10.1214/aos/1176350273

RTI International (2022a). *2022 ACS Content Test: Round 1 and Round 2 Cognitive Testing.* Retrieved July 19, 2022, from https://www.census.gov/content/dam/Census/library/working-papers/2022/acs/2022_Wilson_01.pdf.

RTI International (2022b). *Cognitive Testing for the 2022 ACS Content Test: Round 3 Briefing* Retrieved August 29, 2022 from https://www.census.gov/library/working-papers/2022/acs/2022_Wilson_02.html.

Risley, M., & Oliver, B. (2022). *High-level sampling and weighting requirements for the 2022 American Community Survey Content Test.* U.S. Census Bureau. DSSD 2022 American Community Survey Memorandum Series #ACS22-MP-01.

Spiers, S. (2021a). *Requirements for the Content Follow-Up Reinterview Survey in the 2022 American Community Survey Content Test*. U.S. Census Bureau. American Community Survey Memorandum Series #ACS22-MP-02.

Spiers, S. (2021b). Coverage and nonresponse bias in the 2016 American Community Survey Content Follow-Up reinterview. *Proceedings of the 2021 Joint Statistical Meetings, Survey Research Methods Section* (pp. 1664-1677). American Statistical Association.

Spiers, S., Baumgardner, S., & Mojica, A. (2023). *2022 American Community Survey Content Test evaluation report: Unit-level response*. Forthcoming

U.S. Census Bureau. (2022a). *U.S. Census Bureau statistical quality standards*. Retrieved September 16, 2022, from https://www2.census.gov/about/policies/quality/quality-standards.pdf

U.S. Census Bureau. (2022b). *American Community Survey and Puerto Rico Community Survey Design and Methodology, Version 3.0.* Retrieved December 15, 2022, from https://www.census.gov/programs-surveys/acs/methodology/design-and-methodology.html.pdf

Virgile, M., Weng, C. F., Mills, G., & Spiers, S. (2023). *2022 American Community Survey Content Test evaluation report: Respondent burden*. Forthcoming