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

November 20, 2023

2023 AMERICAN COMMUNITY SURVEY RESEARCH AND EVALUATION REPORT MEMORANDUM
SERIES # ACS23-RER-08

MEMORANDUM FOR Donna M. Daily
Chief, American Community Survey Office

From: David Waddington *David Waddington*
Chief, Social, Economic, and Housing Statistics Division

Prepared by: William Chapin *WC*
Social, Economic, and Housing Statistics Division

Michael Risley *MR*
Lauren Contard *LC*
Decennial Statistical Studies Division

Subject: 2022 American Community Survey Content Test Evaluation Report:
Solar Panels

Attached is the 2022 American Community Survey (ACS) Content Test report for solar panels.
This report presents the methods and results of the test for a new solar panels question.

If you have any questions about this report, please contact William Chapin at 301-763-5735.

cc:
acso.re.workgroup.list@census.gov
Lindsay Longsine   (DSSD)
Peter Massarone
Broderick Oliver
Samantha Spiers

Evan Brassell   (SEHSD)
Molly Cromwell
Jean Mullin

This page is intentionally blank

American Community Survey Research and Evaluation Program

2022 American Community Survey Content Test Evaluation Report: Solar Panels

FINAL REPORT

United States Census Bureau

William Chapin
Social, Economic, and Housing Statistics Division

Michael Risley and Lauren Contard
Decennial Statistical Studies Division

This page is intentionally blank

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

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. This report presents the results of this field test for solar panels.

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.

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.

The question on solar panels is a new question proposed for inclusion in the ACS and has never been asked in the Decennial Census. The solar panels question is currently asked only in the American Housing Survey (AHS) and the Residential Energy Consumption Survey (RECS), which have a very limited geographic detail. Federal agencies have interest in accessing data on solar panels at smaller geographies. We are testing this question to determine how viable this data is when asked and collected for small area level geographies in the ACS.

The solar panels question was asked similarly across modes; the paper version of the question asks respondents, "Does this house, apartment, or mobile home use solar panels that generate electricity?"

The results of the 2022 ACS Content Test show that the proportion of households with solar panels is nominally higher than previous benchmarks from the 2021 AHS and 2020 RECS. This includes all owner-occupied and renter-occupied units. Furthermore, the item missing data

rates in the Contest Test are nominally higher overall than the missing data rates for AHS. However, the prevalence rates and item missing data rates for the Content Test follow the same trend as the AHS, especially the difference between owner-occupied units and renter-occupied units.

When comparing the field test results to the existing benchmarks, it is important to keep in mind that not only are there methodological differences in how the data for the benchmarks are collected, but the benchmark data is also edited and imputed while the Content Test data is not.

The results also showed that one-family houses[1] had a significantly higher rate of having solar panels than apartment buildings. This was expected as owners of one-family houses are more likely to use this as an investment than other types of units or living situations. Also, reporting error may be a large factor since those living in apartment buildings may not be aware of the presence of solar panels.

Taking into account all of these results, the recommendation is to adopt this question. The ACS can provide data at a much more detailed geographic level; the existing benchmarks can provide only estimates for the national level and for select metropolitan areas.

---

[1] One-family houses refer to the actual structure and not the inhabitants of the housing unit.

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 are new:

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

This report presents the results of the field test for solar panels.

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).

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 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 Field Testing Solar Panels in the 2022 ACS Content Test

1.3.1 Justification for Inclusion of Solar Panels in the Content Test

The question on solar panels is a new question proposed for inclusion in the ACS and has never been asked in the Decennial Census. A similar question is asked on the American Housing Survey (AHS) as well as the Residential Energy Consumption Survey (RECS), but they only produce estimates on the national level and for select metropolitan areas. By asking this question on the ACS, we will be able to obtain data for operational solar panels on a housing unit level at smaller geography levels. This information will help the Energy Information Administration (EIA) match energy consumption to energy production across the United States.

---

[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.

The differential adaptation of energy producing technologies such as rooftop solar (i.e., photovoltaic generating capacity) creates the potential for energy infrastructure for and in U.S. households.

There is only one version of this question being tested due to the succinct nature of both the question and the response categories. We will be comparing results from the Content Test with data obtained in the AHS as data is provided every two years. The question was first asked in the AHS in 2017, where 2.6 percent of occupied units reported having solar panels. That number increased to 3.8 percent in 2021 (U.S. Census Bureau, 2021).

We will also compare data with the RECS where possible, as data is not provided in any report, but data may be obtained from the microdata.

1.3.2 Cognitive Testing Development for Solar Panels

During the first round of the cognitive testing process, two versions of the solar panel question were presented. In the first version, respondents were asked, "Does this house, apartment, or mobile home use solar panels that generate electricity?" In the second version, respondents were asked, "Does this house, apartment, or mobile home use solar power or photovoltaic panels that generate electricity?"

After the first round of cognitive testing, it was determined that the first version was the preferred version due to being less technical and easier understand. For all additional rounds of cognitive testing and for the field test, the only version of the question tested asked respondents, "Does this house, apartment, or mobile home use solar panels that generate electricity?"

1.3.3 Question Content

The solar panels question is shown as it appears on the paper questionnaire. Automated versions of the questionnaire have the same content formatted accordingly for each mode. For the modes where it is possible, help text is available.

Figure 1. Solar Panels Question (Paper)

Figure 2. Solar Panels Question (Internet/CAPI)

23. Does this <mobile home/house/apartment/unit> use solar panels that generate electricity?

- o   Yes
- o   No

1.3.4   Research Questions

The research questions (RQ) examined are presented below.

RQ1. *How does the percentage of housing units with solar panels compare to the proportions found in the most recent American Housing Survey (AHS) overall and for owners and renters?*

RQ2. *How does the item missing data rate for the new solar panels question compare to the item missing data rate for the solar panels question on the most recent American Housing Survey (AHS)?*

RQ3. *How does the percentage of housing units with solar panels compare to the percentages found in the most recent Residential Energy Consumption Survey (RECS) overall and for owners and renters?*

RQ4. *Is there a difference in the item missing data rate between modes of response?*

RQ5. *Is there a difference in item missing data rates between selected housing demographics?*

RQ6. *Are the percentages of housing units with solar panels different between selected housing demographics?*

RQ 7. *Is there a difference in response reliability between self-administered and interviewer administered responses?*

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 treatment and Roster Test treatments and the other in the Test treatment.

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. [4]

---

[4] 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.

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 will be in both test treatments to align with Labor Force and Income that also have a reference period change to the previous calendar year.

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.[5]
- 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.[6] 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 incomplete information on the form or reported more than five people on the roster of a paper questionnaire.[7]
- 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,

---

[5] 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.

[6] 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.

[7] 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.

and the household must have a valid telephone number. See the CFU requirements document for the complete list of eligibility requirements (Spiers, 2021a).

2.3.1   Content Test Follow-Up Protocol

As in previous ACS Content Tests, a case was sent to the CFU operation no sooner than two weeks (14 calendar days) after the original interview and had to be completed within three weeks after being sent to the CFU. This timing attempted to balance two competing needs: (1) to minimize the possibility of real changes in answers due to a change in life circumstances between the two interviews; (2) to minimize the possibility of the respondent repeating their previous answer based on their recollection of the original interview response, rather than considering the most appropriate answer.

All CFU reinterviews were conducted by telephone. At the first contact with a household, interviewers asked to speak with the original respondent. If that person was not available, interviewers scheduled a callback at a time when the original respondent was expected to be available. If this respondent could not be reached at the time of the second contact, the interviewer requested to speak with any other eligible household member (a household member who is 15 years or older). CFU reinterviews for the Content Test were conducted in either English or Spanish.

The CFU data collection instrument included the questions being tested for the 2022 ACS Content Test and some production ACS questions for context. It also included questions on public assistance from the 2022 Current Population Survey Annual Social and Economic Supplement (CPS ASEC) to measure response bias in the income from the public assistance question.

The CFU collected an independent household roster by re-asking the Household Roster questions along with Relationship, Sex, Age, and Date of Birth. The remaining CFU questions were only asked of the original household roster members. Only the Control and Roster Test panels collected an independent household roster. The Test panel used the original household roster to ask housing and detailed person questions.[8]

2.3.2   Content Test Follow-Up for Solar Panels

For the CFU reinterview for solar panels, eligible respondents from the original interview were asked the solar panels question again. The independent answers these respondents provided in the original interview and the CFU reinterview were measured for consistency.

---

[8] The Test panel did not need to collect an independent household roster. The independent roster was needed to calculate the response reliability metrics for the Household Roster topic, which only used data from the Control and Roster Test treatments.

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 solar panels is attributable to differences in the wording and format. 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 solar panels 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 data, such as calculating a person's age based on his or her date of birth, but such edits were minimal. [9]

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.[10]

Comparisons between the Control and test versions of solar panels were conducted using a two-tailed t-test at the $\alpha=0.1$ 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.[11] 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).[12]

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:

---

[9] 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.

[10] 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).

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

[12] Use the MULTTEST Procedure in SAS®.

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

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 provided in a separate report (Spiers et al., 2023).[13]

2.4.2 Topic-Level Analysis

To evaluate the potential inclusion of the solar panels question, we calculated a variety of metrics, presented in Sections 2.4.2.1 through 2.4.2.6.

2.4.2.1 Benchmarks

To roughly gauge the accuracy of the responses to solar panels, we compared select estimates derived from these data to similar estimates from the AHS and the RECS, external reliable sources (i.e., a benchmarks).

The AHS asks a similar question to obtain solar panels data for housing units. However, the methodology of the AHS is different from the ACS in ways that prevent direct comparisons of the estimates.

- The AHS is a longitudinal survey that follows the same housing units over time and data is only collected every other year. Notably, the AHS last collected data in 2021 and the data will be an entire year older than the data collected in the Content Test. This is especially an issue for a rapidly changing topic like solar panels.

- The AHS has no internet or paper questionnaire response option, respondents are only able to respond by interview, either in-person or over the phone.

- AHS has different editing procedures that may lead to different estimates.

Because of the differences in design and methodology, we compare the AHS data to the data collected in the Content Test nominally. Included in this review are comparisons by tenure, which AHS includes in their Table Creator tool, looking at owned units versus rented units.

---

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

The RECS collects data on units with solar panels, however the methodology of RECS data collection is different in ways that prevents direct comparisons of the estimates (U.S. Energy Information Administration, 2022).

- The RECS household survey occurs roughly every five years. Notably, the RECS last collected household data in 2020 and the data will be two years older than the data collected in the Content Test. This is especially an issue for a rapidly changing topic like solar panels.

- In comparison to the ACS, the RECS household survey has a much smaller sample at 48,649 households in 2020, with 18,496 responses. For solar panels, which has a relatively low prevalence, this would result in a relatively high variance for the estimate.

- RECS has different editing and imputation procedures, especially due to the comparatively small sample size, which may lead to different estimates.

Included in this review will be comparisons by type of building and tenure, which are both available on the microdata. For tenure, we look at owned units versus rented units. While data on "Mobile homes and Other" units were collected, it is not included in the data. Rates will be compared nominally.

2.4.2.2 Item Missing Data Rates

To measure nonresponse to the solar panels question, we calculated its item missing data rate, the proportion of eligible housing units 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 item missing data rate was compared between modes of response and between selected housing demographics. Additionally, item missing data is one of the benchmarks being compared to AHS in 2.4.2.1.

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

2.4.2.3 Response Distributions

To assess the inclusion of the solar panels question, we compared the proportion reporting solar panels between different housing demographics. Included in this review will be comparisons by income, type of building, and tenure. For type of building there will be two categories compared: one-family houses (attached and detached combined) and buildings with 2 or more apartments. For tenure, we will look at owned units versus rented units. We calculated the response distributions as the proportion of valid responses in a category to all valid responses.

2.4.2.4 Response Reliability

Survey responses are subject to error. Response error occurs for a variety of reasons, such as flaws in the survey design, misunderstanding of the questions, misreporting by respondents, and interviewer effects. For the 2022 ACS Content Test, response error was measured through response reliability or response bias, not both. This was done to reduce respondent burden and breakoffs during the CFU operation. For solar panels, we measured response error using response reliability.

A survey question has good response reliability if respondents tend to answer the question consistently. For the 2022 ACS Content Test, we measured response reliability for a given question by comparing the responses to this question in the original interview to the responses to this same question in the CFU reinterview.

Re-asking the same question of the same respondent allows us to measure simple response variance, using the following measures:

- 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.

In Table 2, "Yes" indicates that the unit is in the category of interest, according to the response from either the original interview or the CFU reinterview. "No" indicates that the unit is not reported to be in the category.

Table 2. Original Interview and CFU Reinterview Counts for Calculating GDR, IOI, and NDR

		Content Test original interview		reinterview totals
		Yes	No	
CFU reinterview	Yes	a	b	a + b
	No	c	d	c + d
original interview totals		a + c	b + d	n

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

a = units in category for both interview and reinterview
b = units not in category for original interview, but in category for reinterview
c = units in category for original interview, but not in category for reinterview

d = units in category for neither interview nor reinterview

n = total units in the universe = a + b + c + d

These counts were weighted to make them more representative of the population.

We calculated the GDR for this response category as:

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

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.

The SV is the part of total variance resulting from the differences between 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 total variance is due to inconsistency in responses, as measured by SRV. A preliminary definition of the IOI is:

$$IOI = \left(\frac{SRV}{SRV + SV}\right) \times 100$$

We can estimate SRV using the GDR, but also need to estimate the denominator (i.e., total variance) in this expression. Based on previous studies, the estimate we use for total variance is:

$$SRV + SV = \frac{p_1 q_2 + p_2 q_1}{2}$$

where:

$$p_1 = \frac{a + c}{n} = \text{original interview proportion in category}$$

$$q_1 = 1 - p_1 = \frac{b + d}{n} = \text{original interview proportion not in category}$$

$$p_2 = \frac{a + b}{n} = \text{CFU proportion in category}$$

$$q_2 = 1 - p_2 = \frac{c + d}{n} = \text{CFU proportion not in category}$$

In comparing relative reliability (or response error) between treatments, if the response categories are essentially the same, then we looked at the differences in the GDR and IOI for each response category. We tested the significance of these differences, using two-tailed t-tests.

If the response categories did not match up exactly between the compared treatments, we either collapsed response categories to form equivalent categories for comparison, or we conducted comparisons for the response categories where it made sense.

So far, we have only discussed response reliability with respect to single response categories. If a question has three or more response categories (or "comparison categories" in cases where it is necessary to collapse some response categories for comparison), we also measured the overall response reliability of a question using the L-fold index of inconsistency, IOI_L. We looked at the difference in IOI_L between treatments and tested for significance as with the single category measures.

Suppose a question has L response categories. Let X_{ij} be the weighted count of sample units (households or persons) for which we have CFU responses in category *i* and original interview responses in category *j*. Here, both *i* and *j* range from 1 to L. Table 3 shows a cross-tabulation of the original interview and CFU results for a generic analysis topic. Note that if L = 2, then Table 3 is equivalent to Table 2.

Table 3. **Cross-Tab of Original Interview and CFU Results: Questions with Response Categories**

		Original Interview categories						CFU totals
		1	2	...	j	...	L	
CFU categories	1	X_{11}	X_{12}	...	X_{1j}	...	X_{1L}	X_{1+}
	2	X_{21}	X_{22}	...	X_{2j}	...	X_{2L}	X_{2+}
	...	...	...	...	...	...	...	...
	i	X_{i1}	X_{i2}	...	X_{ij}	...	...	X_{i+}
	...	...	...	...	...	...	...	...
	L	X_{L1}	X_{L2}	...	X_{Lj}	...	X_{LL}	X_{L+}
Original interview totals		X_{+1}	X_{+2}	...	X_{+j}	...	X_{+L}	$T = \sum_{i=1}^{L} \sum_{j=1}^{L} X_{ij}$

Now define the following proportions:

$$p_{ij} = \frac{X_{ij}}{T}$$

$$p_{+j} = \frac{X_{+j}}{T}$$

$$p_{i+} = \frac{X_{i+}}{T}$$

The IOI$_L$ is calculated as

$$\text{IOI}_L = \frac{1 - \sum_{i=1}^{L} p_{ii}}{1 - \sum_{i=1}^{L} (p_{i+} p_{+i})} \times 100$$

It can be shown that the IOI$_L$ is a weighted sum of the L category IOI values (Biemer, 2011), but this formula is easier for calculation.

The IOI metrics can be biased if the parallel measures assumption is violated, i.e., if the errors in the original interview and CFU reinterview are positively or negatively correlated (Biemer, 2011). We checked this assumption by testing if the net difference rate (NDR) is significantly different from zero. The NDR is the difference between the original interview proportion of positive responses ("Yes" or in the category of interest) and the CFU proportion of positive responses. The NDR is calculated as follows:

$$NDR = (p_1 - p_2) \times 100 = \left(\frac{c - b}{n}\right) \times 100$$

If the NDR is significantly positive or negative, the assumption of "parallel measures" necessary for the SRV and IOI to be valid is not satisfied (Biemer, 2011). In these situations, we use the following adjustment of the IOI, developed by Flanagan (2001):

$$\text{IOI}_{\text{adjusted}} = \frac{\dfrac{n^2(b+c) - n(c-b)^2}{n-1}}{(a+c)(c+d) + (a+b)(b+d)} \times 100$$

3   DECISION CRITERIA

Before field testing solar panels, a team of subject matter experts identified and prioritized which of the research questions presented in Section 1.3.4, would determine whether the solar panels question would be recommended for inclusion in the ACS. The decision criteria for solar panels is presented in Table 4.

Table 4. Decision Criteria for Solar Panels

Priority	Research Questions	Decision Criteria
1	1, 3	The solar panel prevalence rate should be similar to the AHS and RECS benchmarks overall and for selected housing demographics. A particular concern would be if the difference between the prevalence and one of the benchmarks was larger than the difference between the two benchmarks. Due to a possible correlation between solar panel usage and unit nonresponse, a prevalence rate that is lower than the benchmarks would be a larger concern than one that is higher.
2	2, 4, 5	The item missing data rate should be low and comparable to the item missing data rate in the AHS which was 0.23% in 2021.
3	7	The question should have high response reliability. For IOI this would generally be a value less than 20.
4	6	Rates should be comparable when crossed by type of building within tenure.

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 had the same sample size and used 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).

- We assume that the frequency of real changes in answers due to a change in life circumstances between the original interview and CFU reinterview were similar between treatments.

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.

- 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 CFU reinterviews were conducted by phone only, whereas the original interviews were completed online, by mail, by phone in CAPI, and in person in CAPI. Hence, some of the differences observed between the original interviews and the CFU interviews may be the result of mode effect.

- Not all households who provided a response in the original interview were eligible for the CFU reinterview (see Section 2.3 for more information). As a result, 2.5 percent (standard error 0.2) of households from the original Control interviews, 2.5 percent (standard error 0.2) of households from the original Test interviews, and 3.0 percent (standard error 0.2) of households from the original Roster Test interviews were not eligible for the CFU reinterview. These rates were not significantly different between treatments (chi-square p-value 0.11).

- We reinterviewed the same person who responded in the original interview when possible, but accepted interviewing a different person from the same household after two unsuccessful attempts at reaching the original person. Therefore, differences in results between the original interview and CFU reinterview for these cases could partly be from different people answering the questions. We interviewed a different household member in CFU for 7.3 percent (standard error 0.4) of CFU Control cases, 9.4 percent (standard error 0.5) of CFU Test cases, and 8.5 percent (standard error 0.5) of CFU Roster Test cases. These rates were significantly different between treatments (chi-square p-value 0.01) with the rate of CFU Test cases (t-test p-value <0.01) and CFU Roster Test cases (t-test p-value 0.04) being significantly higher than the rate of CFU Control cases.

- We examined potential differences between CFU respondents and nonrespondents within some socioeconomic and demographic characteristics because there were differences in the 2016 CFU reinterview (Spiers, 2021b). For all treatments combined, there were significant differences between CFU respondents and nonrespondents for *household size, tenure, age, race, Hispanic origin, language of original interview response,* and *high and low response areas*. These differences are similar to the ones found in the 2016 CFU (Spiers, 2021b).

- 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 solar panels question. Additional unit-level (i.e., address level) analyses were performed to access the Content Test treatments as a whole. The results are presented in Spiers et al. (2023). The solar panels metrics were calculated for for the Control and Test Treatments were combined.

In general, the overall unit response rates were not significantly different between treatments, nor were the response rate portions by mode. 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. However, this is unlikely to have an effect on the solar panels analysis due to the metrics being for the two treatments combined.

There is no evidence of underlying CFU response rate issues that would negatively affect topic-level response error analyses. However, when examining demographic and socioeconomic distributions, the distributions for tenure were found to be different between the Control and Test treatments for those that responded during CFU and responded by self-response in the original interview. This difference was due to the Test treatment having a higher percentage of households reporting they own their household free and clear than the Control treatment. Despite tenure being used in the solar panels analysis, it was not used for any of the reliability analysis, so this difference was unlikely to effect on the solar panels analysis.

5.1   Benchmark Results for Solar Panels

Research Question 1

How does the percentage of housing units with solar panels compare to the proportions found in the most recent American Housing Survey (AHS) overall and for owners and renters?

Table 5 shows that the estimated percentage of housing units with solar panels for the 2022 ACS Content Test was nominally higher than the 2021 American Housing Survey but the estimates were within one percentage point.

Table 5. Solar Panel Prevalence – 2022 ACS Content Test vs 2021 AHS

	Content Test	AHS
Solar Panel Prevalence	4.8 (0.2)	3.8 (0.1)

Source: U.S. Census Bureau, 2022 American Community Survey Content Test and 2021 AHS | DRB No. CBDRB-FY23-ACSO003-B0061

As seen in Table 6 and Table 7, the Content Test estimate for the percentage of households with solar panels was also close but nominally higher than the AHS estimate for both owner-occupied units and renter-occupied units.

Table 6. Solar Panel Prevalence for Owners – 2022 ACS Content Test vs 2021 AHS

	Content Test	AHS
Solar Panel Prevalence for Owners	5.6 (0.2)	5.1 (0.1)

Source: U.S. Census Bureau, 2022 American Community Survey Content Test and 2021 AHS| DRB No. CBDRB-FY23-ACSO003-B0061

Table 7. Solar Panel Prevalence for Renters – 2022 ACS Content Test vs 2021 AHS

	Content Test	AHS
Solar Panel Prevalence for Renters	3.2 (0.2)	1.5 (0.1)

Source: U.S. Census Bureau, 2022 American Community Survey Content Test and 2021 AHS| DRB No. CBDRB-FY23-ACSO003-B0061

Research Question 2

How does the item missing data rate for the new solar panels question compare to the item missing data rate for the solar panels question on the most recent American Housing Survey (AHS)?

Shown in Table 8, the solar panel missing data rate for the Content Test was relatively small but was nominally higher than the solar panel missing data rate for the AHS.

Table 8. Solar Panel Missing Data Rate – 2022 ACS Content Test vs 2021 AHS

	Content Test	AHS
Solar Panel Prevalence	2.8 (0.1)	0.2 (<0.1)

Source: U.S. Census Bureau, 2022 American Community Survey Content Test and 2021 AHS| DRB No. CBDRB-FY23-ACSO003-B0061

Research Question 3

How does the percentage of housing units with solar panels compare to the percentages found in the most recent Residential Energy Consumption Survey (RECS) overall and for owners and renters?

Table 9 and Table 10 show that the estimate for the percentage of housing units with solar panels from the Content Test was comparable but nominally higher than the estimate from RECS overall and for owner-occupied units.

Table 9. Solar Panel Prevalence – 2022 ACS Content Test vs 2020 RECS

	Content Test	RECS
Solar Panel Prevalence	4.8 (0.2)	3.0 (0.1)

Source: U.S. Census Bureau, 2022 American Community Survey Content Test and 2020 RECS| DRB No. CBDRB-FY23-ACSO003-B0061

Table 10. Solar Panel Prevalence for Owners – 2022 ACS Content Test vs 2020 RECS

	Content Test	RECS
Solar Panel Prevalence for Owners	5.6 (0.2)	4.3 (0.2)

Source: U.S. Census Bureau, 2022 American Community Survey Content Test and 2020 RECS| DRB No. CBDRB-FY23-ACSO003-B0061

Table 11 shows that, while the estimates from the Content Test and from the 2022 RECS are both relatively small for renter-occupied units, the estimate from the RECS is noticeably smaller than the Content Test estimate. This magnitude of this difference is likely due to RECS giving a value of "Not Applicable" for apartment buildings with two or more units which would include a large portion of the renter-occupied units.

Table 11. Solar Panel Prevalence for Renters – 2022 ACS Content Test vs 2020 RECS

	Content Test	RECS
Solar Panel Prevalence for Renters	3.2 (0.2)	0.3 (0.1)

Source: U.S. Census Bureau, 2022 American Community Survey Content Test and 2020 RECS| DRB No. CBDRB-FY23-ACSO003-B0061

5.2 Item Missing Data Rate Results for Solar Panels

Research Question 4

Is there a difference in the item missing data rate between modes of response?

Shown in Table 12, the solar panel missing data rate was statistically significantly higher for the self-administered modes than for the interview-administered mode.

Table 12. Solar Panel Item Missing Data Rate – Self-Administered vs Interview-Administered

Mode	Self-Administered	Interview-Administered	Difference	P-value
Missing Data Rate	3.1 (0.1)	1.6 (0.2)	1.5 (0.3)	*<0.01

Source: U.S. Census Bureau, 2022 American Community Survey Content Test | DRB No. CBDRB-FY23-ACSO003-B0061
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.

Research Question 5

Is there a difference in item missing data rates between selected housing demographics?

Shown in Table 13, the solar panel missing data rate was statistically significantly lower for owner-occupied units than for renter-occupied units. This was expected due to owners being more likely to be familiar with all of the systems of their home.

Table 13. Solar Panel Item Missing Data Rate – Owners vs Renters

Mode	Owners	Renters	Difference	P-value
Missing Data Rate	0.5 (0.1)	1.5 (0.2)	-1.0 (0.2)	*<0.01

Source: U.S. Census Bureau, 2022 American Community Survey Content Test | DRB No. CBDRB-FY23-ACSO003-B0061
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.

Shown in Table 14, the item missing data rate was statistically lower for one-family houses than for apartments. This is likely due to most apartments being renter-occupied units.

Table 14. Solar Panel Item Missing Data Rate – Houses vs Apartments

Mode	Houses	Apartments	Difference	P-value
Missing Data Rate	1.4 (0.1)	2.4 (0.2)	-1.1 (0.2)	*<0.01

Source: U.S. Census Bureau, 2022 American Community Survey Content Test | DRB No. CBDRB-FY23-ACSO003-B0061
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.

5.3 Response Distribution Results for Solar Panels

Research Question 6

Are the percentages of housing units with solar panels different between selected housing demographics?

Shown in Table 15, the percentage of households with solar panels was found to be statistically significantly higher for owners than for renters, which was expected as solar panels are a major investment and is less likely to be installed for a rental unit.

Table 15. Solar Panel Prevalence – Owners vs Renters

Mode	Owners	Renters	Difference	P-value
Prevalence	5.6 (0.2)	3.2 (0.2)	2.5 (0.3)	*<0.01

Source: U.S. Census Bureau, 2022 American Community Survey Content Test | DRB No. CBDRB-FY23-ACSO003-B0061
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.

Shown in Table 16, the percentage of housing units with solar panels was found to be statistically significantly higher for one-family houses than for apartments. This remained true in Table 17 and Table 18, where the percentage of housing units with solar panels was statistically significantly higher for houses among both owner-occupied units and among renter-occupied units.

Table 16. Solar Panel Prevalence – Houses vs Apartments

Mode	Houses	Apartments	Difference	P-value
Prevalence	5.6 (0.2)	2.9 (0.3)	2.7 (0.4)	*<0.01

Source: U.S. Census Bureau, 2022 American Community Survey Content Test | DRB No. CBDRB-FY23-ACSO003-B0061
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.

Table 17. Solar Panel Prevalence for Owners – Houses vs Apartments

Mode	Houses	Apartments	Difference	P-value
Prevalence	5.9 (0.3)	3.4 (0.9)	2.5 (1.0)	*<0.01

Source: U.S. Census Bureau, 2022 American Community Survey Content Test | DRB No. CBDRB-FY23-ACSO003-B0061
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.

Table 18. Solar Panel Prevalence for Renters – Houses vs Apartments

Mode	Houses	Apartments	Difference	P-value
Prevalence	4.1 (0.5)	2.8 (0.3)	1.3 (0.6)	*0.02

Source: U.S. Census Bureau, 2022 American Community Survey Content Test | DRB No. CBDRB-FY23-ACSO003-B0061
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.

5.4   Response Reliability Results for Solar Panels

Research Question 7

Is there a difference in response reliability between self-administered and interviewer administered responses?

Shown in Table 19 and Table 20, the self-administered modes were found to have higher response reliability due to the statistically significantly lower GDR and IOI. With an IOI value of less than 20, the self-administered modes are considered to have good reliability while the interview-administered mode, with a value above 50, are considered to have poor reliability.

Due to the wording of the question being the same across modes, it is more likely that the results reflect a self-selection difference in the responding populations for CAPI versus self-response. The adjusted IOI was used due to the assumption of "parallel measures" not being met for at least one of the categories.

Table 19. Solar Panel Gross Difference Rate – Self-Administered vs Interview-Administered

Mode	Self-Administered	Interview-Administered	Difference	P-value
GDR	1.2 (0.1)	6.5 (0.9)	-5.3 (0.9)	*<0.01

Source: U.S. Census Bureau, 2022 American Community Survey Content Test | DRB No. CBDRB-FY23-ACSO003-B0061
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.

Table 20. Solar Panel Adjusted Index of Inconsistency – Self-Administered vs Interview-Administered

Mode	Self-Administered	Interview-Administered	Difference	P-value
Adjusted IOI	13.3 (1.7)	54.5 (5.0)	-41.1 (5.4)	*<0.01

Source: U.S. Census Bureau, 2022 American Community Survey Content Test | DRB No. CBDRB-FY23-ACSO003-B0061
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.

6   CONCLUSIONS AND RECOMMENDATIONS

The solar panels question is currently only asked in the AHS and the RECS, which have a very limited geographic detail. Several federal agencies have interest in accessing data on solar panels at smaller geographies. We are testing this question to determine how viable this data is when asked and collected for small area level geographies like in the American Community Survey.

The results of the 2022 ACS Content Test show that the proportion of households with solar panels is nominally higher than previous benchmarks from the 2021 AHS and 2020 RECS across all the areas that were compared. This includes all households, as well as owner and renter-occupied units. Furthermore, the item missing data rates from the Contest Test are nominally higher overall than the missing data rates for AHS. However, the prevalence rates and item missing data rates for the Content Test follow the same trend as the AHS, especially the difference between owner-occupied units and renter-occupied units.

When comparing the field test results to the existing benchmarks, it is important to keep in mind that not only are there methodological differences in how the data for the benchmarks are collected, but the benchmark data is also edited and imputed while the Content Test data is not.

The results also show that houses have a higher prevalence of solar panels than apartment buildings. This was expected as owners of one-family houses are more likely to use this as an investment than other types of units or living situations.

Taking into account all of these results, the recommendation is to adopt this question. The ACS can provide data at a much more detailed geographic level; the existing benchmarks can provide only estimates for the national level and for select metropolitan areas.

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

Biemer, P. (2011). *Latent class analysis of survey error.* John Wiley & Sons, Inc.

Flanagan, P.E. (2001). Measurement errors in survey response. (Unpublished doctoral dissertation). University of Maryland, Baltimore County.

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.

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

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

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.

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

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*. Retrieved November 27, 2023, from https://www.census.gov/library/working-papers/2023/acs/2023_Spiers_01.html

U.S. Census Bureau (2021). *American Housing Survey Table Creator*. Retrieved October 2, 2023, from https://www.census.gov/programs-surveys/ahs/data/interactive/ahstablecreator.html

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

U.S. Energy Information Administration. (2022). *Residential Energy Consumption Survey: About the RECS*. Retrieved June 8, 2022 from https://www.eia.gov/consumption/residential/about.php

Virgile, M., Weng, C. F., Mills, G., & Spiers, S. (2023). *2022 American Community Survey Content Test evaluation report: Respondent burden.* Retrieved November 27, 2023, from https://www.census.gov/content/dam/Census/library/working-papers/2023/acs/2023_Virgile_01.pdf