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

November 16, 2023

2023 AMERICAN COMMUNITY SURVEY RESEARCH AND EVALUATION REPORT MEMORANDUM SERIES # ACS23-RER-09

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

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

Prepared by: Brian Glassman *BG*
Social, Economic, and Housing Statistics Division

Lauren Contard *LC*
Decennial Statistical Studies Division

Andres Mojica *AM*
American Community Survey Office

Subject: 2022 American Community Survey Content Test Evaluation Report: Supplemental Nutrition Assistance Program (SNAP)

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

If you have any questions about this report, please contact Lauren Contard at 301-763-3580 or Brian Glassman at 301-763-2463.

Attachment
cc:
acso.re.workgroup.list@census.gov
Lindsay Longsine DSSD
Peter Massarone
Broderick Oliver
Michael Risley
Samantha Spiers
Liana Fox SEHSD
Lynda Laughlin
Jonathan Rothbaum
Hyon Shin

This page is intentionally blank.

DRB Clearance Number—CDRB-FY23-ACSO003-B0059

American Community Survey Research and Evaluation Program

November 16, 2023

2022 American Community Survey Content Test Evaluation Report: Supplemental Nutrition Assistance Program (SNAP)

FINAL REPORT

Brian Glassman
Social, Economic and Housing Statistics Division

Lauren Contard
Decennial Statistical Studies Division

Andres Mojica
American Community Survey Office

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 the Supplemental Nutrition Assistance Program (SNAP) question.

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 (CFU) telephone reinterview was conducted to measure response error. However, the SNAP question was not asked in the CFU interview.

The ACS question about recipiency of funding from SNAP was tested to align the reference period of assistance with administrative records data on SNAP. The reference period was modified from “the past 12 months” to the prior calendar year, to be consistent with a similar change tested for ACS questions on Income. The Census Bureau is researching the use of administrative records as a data source in the future.

Households in the Control treatment received the SNAP question as it appears in ACS production, which asks about SNAP receipt “in the past 12 months.” Households in the Test treatment received the new version, which replaced the phrase “in the past 12 months” with “in 2021.” The new question was evaluated on the basis of item missing data rates, response distributions, and similarity to benchmarks.

There were no statistically significant differences in the item missing data rate between the Test and Control treatments in any mode. There were also no statistically significant differences in the estimate of the proportion of households receiving SNAP.

The Content Test estimates were compared to similar benchmark estimates from the Current Population Survey Annual Social and Economic Supplement (CPS ASEC) and from official counts from the Food and Nutrition Service (FNS) at the U.S. Department of Agriculture (USDA). We also compared the Content Test estimates to ACS data from 2017 to 2019 and 2021. As the ACS has traditionally undercounted SNAP receipt, we compared external benchmarks to the Content Test estimates to see if the change in reference period affected our estimation relative to those counts. The comparisons to benchmarks were only nominal comparisons; no statistical tests were performed. Both the Control and Test estimates differed from all most of the benchmarks. However, because the Test estimate was very similar to the Control estimate, they differed from the benchmarks in similar ways.

We also examined measures of respondent burden. Among internet responses, the median completion time for the SNAP question was one second longer for the Test treatment than Control (7 seconds vs. 6 seconds). In addition, the help screen access rate was higher for Test (0.3 percent (standard error < 0.1)) than Control (0.1 percent (standard error < 0.1)). This difference was found to be significant (*p*-value = 0.05). This may indicate that the new question was more difficult to answer; however, because there was no difference in item nonresponse, it seems that ultimately respondents did not have much more difficulty completing the question.

There were no significant differences between the Control and Test treatments in item nonresponse and in the estimated proportion of households receiving SNAP, despite a significantly higher Test treatment help screen access rate and completion time (which were both low overall). Therefore, we recommend implementing the new version of the SNAP question. However, this recommendation is contingent on the recommendation for the Income topic, because the reference periods for SNAP and Income need to align. The reference period for Income will not be changed until further research is conducted on the use of administrative records.

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 are new:

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

This report presents the results of the field test for SNAP.

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.[1] Cognitive interviews were conducted virtually, in English and Spanish.[2] 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 SNAP in the 2022 ACS Content Test

1.3.1 Justification for Inclusion of SNAP in the Content Test

The ACS question about recipiency of funding from SNAP was tested to align the reference period of assistance with administrative records data on SNAP. The reference period was modified from “the past 12 months” to the prior calendar year. The Census Bureau is researching the use of administrative records as a data source in the future.[3] The same change in reference period was also tested for the Labor Force and Income questions, so the test was also intended to keep the reference period consistent for all three topics.

[1] 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.

[2] 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.

[3] For an example of this research, refer to Fox et al., 2022.

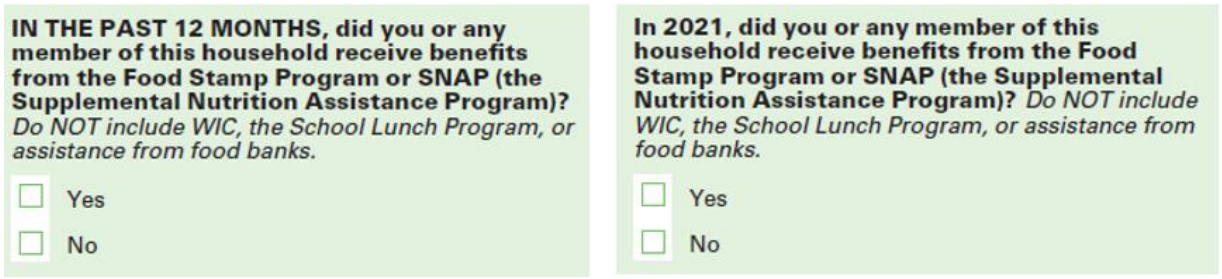
1.3.2 Cognitive Testing Development for SNAP

The only change made to the question was the wording “In the past 12 months” to “In 2021.” In addition to the wording change, we also tested a new placement of the question at the end of the housing section. While testing did not reveal a preference for placement, we ultimately decided to keep the original placement of the question to avoid any confusion or confounding factors in our analysis for the field test.

1.3.3 Question Content

Control and Test versions of each question are shown as they appeared on the paper questionnaire. Automated versions of the questionnaire had the same content formatted accordingly for each mode. The only change made to the SNAP question was the year of reference, from “THE PAST 12 MONTHS” to the prior calendar year. For the 2022 ACS Content Test, the prior calendar year was 2021.

Figure 1. Control and Test Version of SNAP



1.3.4 Research Questions

The questions examined for this research are presented below.

RQ1. How do the estimated proportions of households receiving SNAP in the Test treatments compare with estimates from the 2022 Current Population Survey Annual Social and Economic Supplement (CPS ASEC), 2022 Food and Nutrition Service (FNS) counts, and ACS data from 2017 to 2019 and 2021?

RQ2. How do the estimated proportions of household receiving SNAP in the Control treatment compare with estimates from the 2022 CPS ASEC, 2022 FNS counts, and ACS data from 2017 to 2019 and 2021?

RQ3. Is there a difference in item missing data rates for the Test and Control versions of the SNAP question?

RQ4. Is the proportion of households reporting that they receive SNAP benefits different for the Test versions than for the Control version?

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 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 was 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. However, SNAP was not asked in the CFU reinterview. The primary goal for changing the SNAP question was to align the reference period of the questions more closely with administrative records, which was examined by benchmark comparisons of Content Test data to administrative records and other known data sources.

[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.

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 SNAP 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 SNAP 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. [8]

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.[9]

Comparisons between the Control and test versions of SNAP 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. 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).[10]

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:

[8] 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.

[9] 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).

[10] Using the MULTTEST Procedure in SAS®.

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).[11]

2.4.2 Topic-Level Analysis

To evaluate the changes to SNAP, we calculated a variety of metrics, presented in Sections 2.4.2.1 through 2.4.2.3.

2.4.2.1 Benchmarks

To roughly gauge the accuracy of the responses to SNAP, we compared select estimates derived from these data to similar estimates from the CPS ASEC and from official counts from the FNS at the U.S. Department of Agriculture (USDA). We also compared the test estimates to ACS data from 2017 to 2019 and 2021.[12] We compared the estimates to benchmark estimates, nominally.

2.4.2.2 Item Missing Data Rates

To measure nonresponse to SNAP, 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. We compared item missing data rates via two-tailed t-tests.

2.4.2.3 Response Distributions

To assess how changes to the SNAP question affected the resulting estimates, we calculated the proportion of households reporting that they received SNAP benefits out of all valid responses. We compared these proportions for the Control and Test versions of the SNAP question, using two-tailed t-tests.

2.4.2.4 Response Reliability and Response Bias

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.

Response reliability was not assessed for SNAP.

[11] As part of the 2022 ACS Content Test, we also analyzed respondent burden. The results of this analysis are contained in Virgile et al. (2023).
[12] Since the 2020 ACS estimates were experimental, they were not used as benchmarks.

3 DECISION CRITERIA

Before field testing SNAP, 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 SNAP would be recommended for inclusion in the ACS. The decision criteria for SNAP are presented in Table 2.

Table 2. Decision Criteria for SNAP

Priority	Research Questions	Decision Criteria
1	4	Response Distributions: We hope to see either no difference in the proportion or a greater proportion of people receiving SNAP.
2	3	Item Missing Data Rates: We hope to see no difference or lower item nonresponse.
3	1, 2	Benchmarks: We expect to see comparable estimates.

Regardless of the decision criteria outlined above, the reference period used for SNAP will align with the reference period used for the ACS Income questions.

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. The unit analysis results indicate that this assumption was met for overall response. The only indication of a lack of homogeneity was a difference in the distribution of race for CAPI respondents. Details can be found in Section 5.
- 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).

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 respondent burden analysis used data from the internet and CAPI instruments, and was only conducted for those modes. We did not assess respondent burden in the mail mode, which made up 13.1 percent (SE = 0.3).of Control cases and 12.3 percent (SE = 0.3) of Test cases.
- 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 SNAP. 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 full results are presented in Spiers et al. (2023); the salient results for SNAP are summarized here.

When comparing the Control and Test treatments in the original interview, there were no statistically significant differences in unit response rates, either for all modes combined or each mode separately. We are confident that there are no underlying response rate concerns that would impact topic-specific original interview comparisons between treatments.

We also compared the distributions of demographic and socioeconomic characteristics (age, sex, Hispanic origin, race, building tenure, and language of response) between treatments. There were no statistically significant differences between treatments for any of the examined variables, both for all modes combined and self-response. When looking at CAPI responses, the distribution of race was significantly different between the Test and Control treatments (*p*-value = 0.02). When t-tests were performed on the individual response categories, the only significant difference was in the "Some Other Race" category. The proportion of respondents in this category was 3.6 percentage points lower (standard error (SE) = 1.6, *p*-value = 0.09) in Test than Control.

Since there is a relationship between race and SNAP receipt (Cronquist, 2021), the difference in the race of respondents in the Test and Control treatments in the CAPI mode should be kept in mind when interpreting the CAPI results. However, because the only significant difference was in the other race category, this is not a major cause for concern. Aside from race in the CAPI mode, the results of the unit analysis do not present any concerns for analysis of the SNAP data.

5.1 Benchmark Results for SNAP

To assess whether the change in reference period affected reported SNAP recipiency, we compared Content Test estimates (Control and Test) to estimates from the CPS ASEC SNAP questions. We also compared estimates to official counts from the FNS at USDA. The ACS has traditionally undercounted SNAP receipt relative to FNS official counts (Harris, 2014); we wanted to see if the change in reference period affected our estimation relative to those counts. We also compared Content Test estimates to prior years of ACS data (2017-2019 and 2021).[13] Because all of these benchmark rates come from different sources with different methodologies than the ACS Content Test, statistical tests could not be performed. All comparisons with the Content Test results are only nominal, to assess whether our results were grossly different from other sources of data on SNAP recipiency.

<u>Research Question 1</u>

How do the estimated proportions of households receiving SNAP in the Test treatments compare with estimates from the 2022 CPS ASEC, 2022 FNS counts, and ACS data from 2017 to 2019 and 2021?

[13] Since the ACS 2020 estimates were experimental, they were not used as benchmarks.

Research Question 2

How do the estimated proportions of household receiving SNAP in the Control treatment compare with estimates from the 2022 CPS ASEC, 2022 FNS counts, and ACS data from 2017 to 2019 and 2021?

Content Test SNAP estimates and the estimates from the 2022 CPS ASEC, 2022 FNS counts, and previous ACS data are presented in Tables 3, 4, and 5, respectively.

For both the Test and Control treatments, the Content Test estimates of the proportion of households receiving SNAP were comparable to the estimate from the 2022 CPS ASEC (10.9 (SE = 0.2) and 10.7 (SE = 0.3) percent, respectively, vs. 10.9 percent (SE = 0.1)).[14] The Content Test estimates were lower than the 2022 FNS count of 16.9 percent. Finally, they were lower than the 2017, 2018, and 2021 ACS estimates. The Control estimate was higher than the 2019 ACS estimate, and the Test estimate matched it (see Table 5).

The CPS ASEC and FNS counts are different sources with different data collection methodology than the ACS Content Test, so we would expect there to be some difference in estimated SNAP recipiency rates. In addition, the Content Test methodology differed in several ways from ACS Production (see Section 2.2 and Section 4.2), so the Content Test estimates are not expected to exactly match ACS estimates from prior years.

Both the Control and Test estimates differed from most of our benchmarks (with the exception of the Control estimate and the 2019 ACS, and the Test estimate and the CPS ASEC). However, because the Test and Control estimates were similar (with no statistically significant difference between them (see Section 5.3)), they differed from the benchmarks in similar ways.

Table 3. Percentage of Households Receiving SNAP: 2022 ACS Content Test vs. 2022 CPS ASEC

	Content Test – Test	Content Test – Control	CPS ASEC
SNAP Rate	10.9 (0.2)	10.7 (0.3)	10.9 (0.1)

Source: U.S. Census Bureau, 2022 American Community Survey Content Test | DRB No. CBDRB-FY23-ACSO003-B0059
Note: Minor additive discrepancies are due to rounding. Standard errors are in parentheses.

Table 4. Percentage of Households Receiving SNAP: 2022 ACS Content Test vs. 2022 FNS Counts

	Content Test – Test	Content Test – Control	FNS Count
SNAP Rate	10.9 (0.2)	10.7 (0.3)	16.9

Source: U.S. Census Bureau, 2022 American Community Survey Content Test | DRB No. CBDRB-FY23-ACSO003-B0059
Note: Minor additive discrepancies are due to rounding. Standard errors are in parentheses. Standard errors are not available for FNS count. FNS count is based on the number of households receiving SNAP from FNS (U.S. Department of Agriculture Food and Nutrition Service, 2023) divided by the 2021 ACS estimate of the number of households in the U.S.

[14] Standard errors are not available for FNS counts and were not produced for CPS ASEC SNAP estimates.

Table 5. Percentage of Households Receiving SNAP: 2022 ACS Content Test vs. Prior ACS Years

	Content Test – Test	Content Test – Control	2017 ACS	2018 ACS	2019 ACS	2021 ACS
SNAP Rate	10.9 (0.2)	10.7 (0.3)	11.7 (0.1)	11.3 (0.1)	10.7 (0.1)	12.4 (0.1)

Source: U.S. Census Bureau, 2022 American Community Survey Content Test | DRB No. CBDRB-FY23-ACSO003-B0059
Note: Minor additive discrepancies are due to rounding. Standard errors are in parentheses.

5.2 Item Missing Data Rate Results for SNAP

Research Question 3

Is there a difference in item missing data rates for the Test and Control versions of the SNAP question?

To evaluate item nonresponse, we calculated item missing data rates for the SNAP question and compared the Test and Control treatments, for each response mode separately and all response modes combined.

All households in sample are eligible to answer the SNAP question, since it is in the household section of the survey. It is a Yes/No question, so when no box was checked, the item response was considered missing. For CAPI, a "Don't Know" or "Refusal" was considered a missing response.

The comparisons of item missing data rates can be seen in Table 6. There were no statistically significant differences between the Test and Control item missing data rates for any mode. There is no evidence that the change in reference period in the SNAP question had an effect on item nonresponse.

Table 6. SNAP Item Missing Data Rates by Mode

Mode	Test Percent	Control Percent	Difference	Adjusted P-Value
Overall	3.5 (0.1)	3.6 (0.2)	<0.1 (0.2)	0.86
Internet	3.5 (0.2)	3.3 (0.2)	0.2 (0.3)	0.86
Mail	6.4 (0.3)	7.9 (0.6)	-1.5 (0.7)	0.13
CAPI	1.7 (0.2)	1.1 (0.2)	0.6 (0.3)	0.26

Source: U.S. Census Bureau, 2022 American Community Survey Content Test | DRB No. CBDRB-FY23-ACSO003-B0059
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.

5.3 Response Distribution Results for SNAP

Research Question 4

Is the proportion of households reporting that they receive SNAP benefits different for the Test versions than for the Control version?

We calculated the percentage of households reporting that they received SNAP benefits, out of all households that answered the question. The percentages were compared between Control and Test treatments for each response mode separately and all combined. The results can be seen in Table 7.

There were no statistically significant differences in the proportion of households receiving SNAP between the treatments. There is no evidence that the change in reference period had a significant effect on the measurement of SNAP rates.

Table 7. Percentage of Households Receiving SNAP by Mode

Mode	Test Percent	Control Percent	Difference	Adjusted P-Value
Overall	10.9 (0.2)	10.7 (0.3)	0.2 (0.4)	0.76
Internet	8.0 (0.2)	7.4 (0.3)	0.6 (0.3)	0.29
Mail	11.4 (0.6)	12.6 (0.8)	-1.2 (1.0)	0.68
CAPI	18.5 (0.7)	18.6 (1.1)	-0.1 (1.2)	0.93

Source: U.S. Census Bureau, 2022 American Community Survey Content Test | DRB No. CBDRB-FY23-ACSO003-B0059
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.

5.4 Other Metrics: Respondent Burden Results for SNAP

As part of the Content Test analysis, we also examined several measures of respondent burden to check if there was an impact from the new and revised questions. The measures we analyzed were median completion times, help screen access rates, and breakoff rates. The full results can be found in Virgile et al. (2023). This section discusses the relevant results for SNAP.[15]

Median completion time is the median length of time that respondents spent on the SNAP question screen, recorded in the internet and CAPI instruments. For internet responses, the median completion time was 6 seconds for the Control treatment and 7 seconds for the Test treatment.[16] For CAPI responses, the median completion time was 6 seconds for both treatments. While there was a small increase in internet completion time for the Test question, it does not seem large enough to be a great cause for concern.

The help screen access rate measures the percentage of internet respondents who accessed the help screen for the SNAP question.[17] This rate was 0.1 percent (SE < 0.1) for the Control treatment and 0.3 percent (SE < 0.1) for the Test treatment. The difference between Test and Control was significant (*p*-value = 0.05). It appears that the Test question did prompt more use of the help screen, which may indicate a question that was more difficult to understand.

[15] The respondent burden analysis was only conducted for the internet and CAPI modes.
[16] Because these are median times, there are no standard errors or significance tests for these results.
[17] Help screen access rates and breakoff rates were not weighted. See Virgile et al. (2023) for details.

However, because there was no significant difference in item missing data rates (Section 5.2), it does not appear that this had a major impact on respondents' ability to answer the question.

The breakoff rate measures how many respondents stopped filling out the questionnaire during a particular topic. An increase in the breakoff rate could indicate that more respondents found the new question confusing or overly sensitive to the extent that they stopped answering the survey entirely. The breakoff rate for SNAP was defined as the proportion of respondents who accessed the internet instrument but did not reach the "Submit" action, and whose last action was on the SNAP question screen. This rate was < 0.1% (both SEs < 0.1) for both treatments and there was no significant difference.

Overall, in our view none of the respondent burden results (as measured in the internet and CAPI mode) are a serious impediment to implementing the new version of the SNAP question.

6 CONCLUSIONS AND RECOMMENDATIONS

The purpose of the revised SNAP question in the Content Test was to test whether changing the reference period from the last 12 months to the previous calendar year would affect response rates to the SNAP question or affect SNAP receipt rates. The change in reference period will make it easier for SNAP and other income variables to match up with administrative records.

The overall takeaway from the findings is that, with the exception of help screen access rates, there are no significant differences when using the Test question versus the Control question. More specifically, the notable findings can be grouped into three categories.

First, both the Control and Test estimates nominally differed from most of our benchmarks. However, because the Test estimate was very similar to the Control estimate, they differed from the benchmarks in similar ways. In other words, the difference from the benchmarks was not due to the difference between the Test and Control questions.

Second, there were no statistically significant differences between the Test and Control item missing data rates overall or for any mode. There is no evidence that the revision to the SNAP question had an effect on item nonresponse.

Third, there were no statistically significant differences in the proportion of households receiving SNAP overall or by the different collection modes between the treatments. There is no evidence that the change in reference period had a significant effect on the measurement of SNAP rates.

We also examined measures of respondent burden in the internet and CAPI modes. There were two differences found between the treatments on these measures. First, the median completion time for the SNAP question in the internet mode was one second longer for the Test treatment than Control. Second, there was a statistically significant increase in the help screen access rate. However, because there was not a statistically significant increase in item

nonresponse, we do not consider these results to be a cause for concern if the new question were implemented.

There were no significant differences between the Control and Test treatments in item nonresponse and in the estimated proportion of households receiving SNAP, despite a significantly higher Test treatment help screen access rate and completion time (which were both low overall). Based on these findings, we recommend the adoption of the Test question. However, this recommendation is contingent on the recommendation for the Income topic, because the reference periods for SNAP and Income need to align. The reference period for Income will not be changed until further research is conducted on the use of administrative records.

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

Cronquist, K. (2021). *Characteristics of Supplemental Nutrition Assistance Program households: Fiscal year 2019.* U.S. Department of Agriculture, Food and Nutrition Service, Office of Policy Support. Retrieved August 29, 2023, from https://fns-prod.azureedge.us/sites/default/files/resource-files/Characteristics2019.pdf

Fox, L. Rothbaum, J., & Shantz, K. (2022). Fixing Errors in a SNAP: Addressing SNAP Underreporting to Evaluate Poverty. *American Economic Association Papers and Proceedings*, 112: 330-334.

Harris, B. C. (2014).*Within and across county variation in SNAP misreporting: Evidence from linked ACS and administrative records.* U.S. Census Bureau. Retrieved June 12, 2023, from https://www.census.gov/content/dam/Census/library/working-papers/2014/adrm/carra-wp-2014-05.pdf

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

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., Baumgardner, S., & Mojica, A. (2023). *2022 American Community Survey Content Test evaluation report: Unit-level response*. U.S. Census Bureau. American Community Survey Memorandum Series #ACS23-RER-03. https://www.census.gov/library/working-papers/2023/acs/2023_Spiers_01.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. Department of Agriculture Food and Nutrition Service. (2023). *SNAP: Monthly participation, households, benefits: FY19 through FY23: National view summary.* Retrieved September 7, 2023, from https://fns-prod.azureedge.us/sites/default/files/resource-files/snap-4fymonthly-8.pdf

Virgile, M., Weng, C. F., Mills, G., & Spiers, S. (2023). *2022 American Community Survey Content Test evaluation report: Respondent burden*. U.S. Census Bureau. American Community Survey Memorandum Series #ACS23-RER-04. https://www.census.gov/library/working-papers/2023/acs/2023_Virgile_01.html