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

November 27, 2023

2023 AMERICAN COMMUNITY SURVEY RESEARCH AND EVALUATION REPORT MEMORANDUM SERIES # ACS23-RER-07

MEMORANDUM FOR Donna Daily
Chief, American Community Survey Office

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

Prepared by: Molly Cromwell *MC*
Social, Economic and Housing Statistics Division

Lindsay Longsine *LL*
Decennial Statistical Studies Division

Matthew Virgile *MV*
Center for Behavioral Science Methods

Subject: 2022 American Community Survey Content Test Evaluation Report: Electric Vehicles

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

If you have any questions about this report, please contact Molly Cromwell at 301-763-3379 or Lindsay Longsine at 301-763-0364.

cc:

Evan Brassell	SEHSD	Lauren Contard	DSSD	Michael Risley
Kevin Bragdon		Peter Massarone		Samantha Spiers
Hyon Shin		Broderick Oliver		acso.re.workgroup.list

This page is intentionally blank

November 27, 2023

2022 American Community Survey Content Test Evaluation Report: Electric Vehicles

FINAL REPORT

Molly Cromwell
Social, Economic and Housing Statistics Division
Lindsay Longsine
Decennial Statistical Studies Division
Matthew Virgile
Center for Behavioral Science Methods

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 Electric Vehicles 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 telephone reinterview was conducted to measure response error.

The Electric Vehicles question is new to the ACS and had not been tested using the ACS framework. With a growing demand for electric vehicles, it is more important than ever to determine if the national energy supply meets the needs of the country. The ability to assess energy needs at levels of geography available in the ACS also allows for adjustments to municipal infrastructure systems and predictions about energy consumption in the future. Determining the prevalence of electric vehicles at the housing-unit level will help multiple agencies understand the energy supply, demand, and current technology of energy resources when making projections for future energy needs. Currently, federal agencies are able to obtain data on energy consumption and make predictive decisions about energy infrastructure from state and local governments, but there are limitations on their ability to collect data at smaller geographic levels. By testing this question, we hoped to create a lasting source of reliable data from which decisions about energy needs can be made.

During the Content Test, two versions of the electric vehicle question were tested. The Control treatment version gave respondents the opportunity to identify electric vehicle ownership by type of vehicle, either plug-in electric or hybrid. The Test treatment version asked respondents about their electric vehicle ownership but did not allow for a breakdown based on the type of electric vehicle in their household. However, the Test treatment version was able to estimate the rate of plug-in electric vehicles since the question asked respondents to include both all-electric and plug-in hybrid vehicles when marking 'yes.' This affirmative was comparable to marking 'yes' to the first part of the Control treatment question, which explicitly identified plug-in electric vehicles.

The most notable results from the Content Test pertaining to electric vehicles were:

- The Test treatment had higher rates of plug-in electric vehicle ownership overall, as well as in the internet and CAPI modes.
- The Control treatment had higher item missing data rates overall, which were driven by the self-response modes.

Based on the results of the Content Test, we recommend adopting the Electric Vehicles question from the Test treatment into the ACS.

1 BACKGROUND

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

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

This report presents the results of the field test for the Electric Vehicles question.

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 included 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.

1.3 Field Testing Electric Vehicles in the 2022 ACS Content Test

1.3.1 Justification for Inclusion of Electric Vehicles in the Content Test

The proposal for adding a question about the use of electric vehicles to the ACS was submitted by the Energy Information Administration (EIA). The EIA is tasked with determining the current national energy supply and whether it meets the demands of the country. These demands are projected to increase greatly in the next 10-20 years. The EIA is able to collect data from state and local governments but lacks the resources to collect information at smaller levels (e.g., housing units and public buildings).

[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.

Determining the prevalence of electric vehicles at the housing-unit level will help multiple agencies understand the energy supply, demand, and current technology of energy resources when making projections for future energy needs. Understanding energy consumption at these lower levels also helps in making the technological and capital changes necessary within the energy infrastructure. Additionally, the prevalence of electric vehicle ownership may help to evaluate the effectiveness of energy and tax policies.

The current administration has made it a goal that half of all vehicle sales in 2030 will be zero-emissions vehicles, which includes plug-in hybrid electric and all-electric vehicles. In addition to heavily promoting zero-emissions vehicle sales, the current administration has prioritized electric vehicle manufacturing and the establishment of electric vehicle infrastructure in the country, which includes installing the first-ever national network of electric vehicle charging stations. By adding the Electric Vehicles question to the ACS, we will be able to provide data showing which communities have the highest needs for these charging stations.

1.3.2 Cognitive Testing Development for Electric Vehicles

Two versions of this question were presented during the cognitive testing process. The first version asked respondents, “Are any of the following types of electric vehicles kept at home for use by members of this household?” and offered response options of “a plug-in electric vehicle” or “another type of electric vehicle”. The second version of the question presented during cognitive testing asked respondents, “At this house, apartment, or mobile home, do you or any member of this household own or lease any of the following types of electric vehicles?” and offered response options of “a plug-in electric vehicle” or “a hybrid-electric vehicle”.

After the reviewing the cognitive testing findings and discussing optimal phrasing with experts across agencies who have a vested interest in the question, the final versions were created for the field test. The first version asks respondents, “Are any of the following types of electric vehicles kept at home for use by members of this household?” with “yes” or “no” response options for “A plug-in electric vehicle” and “A hybrid electric vehicle”. This version (Version 1) was written and cognitively tested by the Census Bureau. The second version of the Electric Vehicles question (that was field tested but not cognitively tested) asks, “Do you or any member of this household own or lease an electric vehicle? Include both all-electric and plug-in hybrid electric vehicles” with a response option of “yes” or “no”. The second version (Version 2) is borrowed from the Residential Energy Consumption Survey (RECS).

The phrasing was altered for the versions of the questions being tested in the field in order to best capture respondents who are depending upon the current energy supply to charge their vehicles. By splitting the response options of the first version of the question into plug-in electric versus hybrid electric, we were able to separate self-charging vehicles from vehicles that need to be plugged into an outlet for charging. Additionally, the second version of the question being tested in the field encouraged respondents to answer in the affirmative if they

own or lease any vehicles that require being plugged into an outlet for charging. This aspect was not present in previous versions of the questions.

1.3.3 Question Content

The Electric Vehicles question is new to the ACS and had not been tested using the ACS framework. There were two versions field tested in the 2022 ACS Content Test. The first version of the question separated electric vehicles into plug-in electric and hybrid-electric categories. This version was cognitively tested. The second version asked respondents whether they own either an all-electric vehicle or a plug-in electric vehicle. This version of the question used the same wording as the electric vehicle question on the RECS.

Figure 1. Control and Test Version of Electric Vehicles

Control Version of the Electric Vehicles Question (Paper)

Are any of the following types of electric vehicles kept at home for use by members of this household?

a. A plug-in electric vehicle?

- ☐ Yes
- ☐ No

b. A hybrid electric vehicle?

- ☐ Yes
- ☐ No

Test Version of the Electric Vehicles Question (Paper)

Do you or any member of this household own or lease an electric vehicle? *Include both all-electric and plug-in hybrid electric vehicles.*

- ☐ Yes
- ☐ No

1.3.4 Research Questions

The research questions (RQ) examined for this research are presented below.

RQ1. How does the percentage households with plug-in electric vehicles compare to the proportions found in the most recent RECS?

RQ2. Is the overall item missing data rate different between Version 1 and Version 2?

RQ3. Are the item missing data rates different between Version 1 and Version 2 when dividing responses into self-administered (paper and internet) versus interviewer administered (CAPI)?

RQ4. Are the percentages of households with a plug-in electric vehicle different between Version 1 and Version 2?

RQ5. Are the measures of response reliability different between Version 1 and Version 2?

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. [3]

Table 1. Questions by Treatment

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

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

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.

[3] 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.

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.[4]
- 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.[5] 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.[6]

[4] 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.

[5] 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.

[6] 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.

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

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

2.3 Content Follow-Up Operation

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

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.[7]

2.3.2 Content Test Follow-Up for Electric Vehicles

The initial question for Electric Vehicles was asked again during the Content Test Follow-Up in the Control and Test treatments. A survey question has good response reliability if respondents tend to answer the question consistently.

Reliability was assessed by the proportion of households with a different answer between the original response and the CFU response.

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 Electric Vehicles 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 Electric Vehicles 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. The weights used in the CFU analysis also included an adjustment for CFU non-response.[9]

Comparisons between the Control and test versions of Electric Vehicles 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

[7] 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.

[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).

(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.[10] 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).[11]

We estimated the variances of the estimates using the Successive Differences Replication (SDR) method with replicate weights, the standard method used in the ACS (see U.S. Census Bureau, 2022b, Chapter 12). We calculated the variance for each rate and difference using the formula below. The standard error of an estimate (X_0) is the square root of the variance:

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

where:

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

2.4.1 Unit-Level Analysis

The unit response rate is important, as it provided 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).[12]

2.4.2 Topic-Level Analysis

To evaluate the addition of the Electric Vehicles question, we calculated a variety of metrics, presented in Sections 2.4.2.1 through 2.4.2.4. While the Electric Vehicle questions were in the Control, Test, and Roster treatments, all topic-specific analysis was a comparison between the Control and Test treatments. The Roster Test treatment was not used in the analysis.

2.4.2.1 Benchmarks

To roughly gauge the accuracy of the responses to the Electric Vehicles question, we nominally compared the rates of electric vehicle ownership (for both Control and Test) to similar estimates from the RECS, an external reliable source (i.e., a benchmark).

RQ1. How does the percentage households with plug-in electric vehicles compare to the proportions found in the most recent RECS?

[10] See Section 2.4.2.4 for the definition of Gross Difference Rate.
[11] Use the MULTTEST Procedure in SAS®.
[12] 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 wording for Version 2 of the Electric Vehicles question was borrowed from the RECS. The RECS survey is conducted by the Energy Information Administration through in-person interviews, as well as internet and mail survey forms. Every five years, a nationally representative sample of households is surveyed regarding their energy use patterns and household demographics. Household energy suppliers are also surveyed to estimate energy cost and usage patterns. This sample is smaller than the annual ACS sample (roughly 18,500 housing units for RECS vs roughly 3.5 million housing units for ACS). There are also differences in how the RECS and the ACS edit and impute data.

Because of these design differences, we could not make direct comparisons between the electric vehicles rates. Thus, rates of plug-in electric vehicle usage found in the most recent RECS survey and those observed during field testing were compared nominally.

2.4.2.2 Item Missing Data Rates

To measure nonresponse to the Electric Vehicles question, we calculated its item missing data rate, the proportion of eligible housing units for which a required response was missing. A housing unit was considered eligible if the respondent indicated they had at least one vehicle. A high item missing data rate can be indicative of a question that lacked clarity, was sensitive, or was simply too difficult to answer.

RQ2. Is the overall item missing data rate different between Version 1 and Version 2?

For the mail and internet formats of the questions, a missing item response was indicated when there were no response boxes checked. For CAPI, a response of either “Don’t Know” or “Refused” was considered a missing item response. To test for differences in the rate of missing item responses between Version 1 and Version 2 of the question, our analysis involved conducting two-tailed t-tests at the 0.1 level of significance.

RQ3. Are the item missing data rates different between Version 1 and Version 2 when dividing responses into self-administered (paper and internet) versus interviewer-administered (CAPI)?

To test for differences between the item missing data rates of Version 1 and Version 2, we conducted two-tailed t-tests at the 0.1 level of significance.

2.4.2.3 Response Distributions

To assess how adding the Electric Vehicles question affected the resulting estimates, we compared the response distributions of the Control and the Test version of electric vehicles. We calculated the response distributions as the proportion of valid responses in a category to all valid responses.

RQ4. Are the percentages of households with a plug-in electric vehicle different between Version 1 and Version 2?

For this question, we compared the percentage of housing units that marked 'yes' to Version 2 to the percentage of households that marked 'yes' to Part A of Version 1. To test for differences between the proportions of households with a plug-in electric vehicle in Version 1 and Version 2, we conducted two-tailed t-tests at the 0.1 level of significance. We also compared cross-sectionally by household demographics between Version 1 and Version 2. The household demographics that were included in the analysis are building type, tenure, year built, persons per household, and total household income. Table 2 provides the description and valid values of the household demographics used in the analysis.

Table 2. Household Demographics

Crosstab	Group
Building type	House, townhome, or mobile home Small apartment building Large apartment building Unknown
Tenure	Owned Rented Unknown
Year built	2000 or later 1980 to 1999 1950 to 1979 1949 or earlier Unknown
Persons in household	1 person 2 people 3 to 5 people 6 or more people
Total household income	Less than $25,000 a year $25,000 to $49,999 a year $50,000 to $74,999 a year $75,000 to $99,999 a year $100,000 to $149,999 a year $150,000 to $199,999 a year $200,000 or more a year Unknown

*Total income was summed up to the household-level to create total household income.

Comparisons were made using a Rao-Scott chi-square test that checks for a significant difference between two sample distributions (Rao & Scott, 1987). If the chi-square test indicated a significant difference between the Control and Test distributions, we tested for significant differences in the individual category proportions using two-tailed t-tests.

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 electric vehicles, 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 3, "Yes" indicates that the unit was in the category of interest, according to the response from either the original interview or the CFU reinterview. "No" indicates that the unit was not reported to be in the category.

Table 3. 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 were 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 4 shows a cross-tabulation of the original interview and CFU results for a generic analysis topic. Note that if L = 2, then Table 4 is equivalent to Table 3.

Table 4. Cross-Tab of Original Interview and CFU Results: Questions with Response Categories

		Original Interview categories						
		1	2	…	j	…	L	CFU totals
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

$$\mathrm{IOI_L} = \frac{1 - \sum_{i=1}^{L} p_{ii}}{1 - \sum_{i=1}^{L} (p_{i+} p_{+i})} \times 100$$

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.

IOI_L was not calculated for the Electric Vehicles question as it only had two response categories ("yes" or "no").

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

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

3 DECISION CRITERIA

Before field testing the Electric Vehicles question, a team of subject matter experts identified and prioritized which of the research questions presented in Section 1.3.4, would determine which version of the Electric Vehicles question would be recommended for inclusion in the ACS. The decision criteria for Electric Vehicles are presented in Table 5.

Table 5. Decision Criteria for Electric Vehicles

Priority	Research Question	Decision Criteria
1	4, 5	The version which had the highest rate of plug-in electric vehicles and lowest response variance is preferrable. We expected underreporting would be more likely than overreporting, so a higher estimate is more likely to be accurate, but only if the response variance was not statistically higher.
2	2, 3	The version that had lower item missing data rates is preferable.

Research question 1, which compared the results of the field test to the RECS benchmark, was not considered in the decision criteria.

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 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 nonrespondent 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 Electric Vehicles question. 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.

Original Interview

The overall unit response rates were not significantly different between the Control and Test treatments, nor were the response rate portions by mode. When looking at response rates within high and low response areas, there was a difference in the low response area category of the mail mode only, but these results did not appear in the overall comparisons.

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

We are confident there were no differences that would impact original interview comparisons between treatments for the Electric Vehicles question.

CFU Reinterview

The overall unit response rates were not significantly different between the Control and Test treatments, nor were the response rate portions by mode.

When examining demographic and socioeconomic distributions, none of the overall response distributions were significantly different between the Control and Test treatments. When looking at distributions among self-responses and CAPI responses, only the distribution for

tenure among self-responses for the Control and Test treatments was significantly different. This distribution difference showed up in the Owned Free and Clear category.

We are confident there were no differences that would impact response error analyses between treatments for the Electric Vehicles question.

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

5.1 Benchmark Results for Electric Vehicles

RQ1. How does the percentage households with plug-in electric vehicles compare to the proportions found in the most recent RECS?

The most recent RECS was in 2020. Table 6 provides the rates of plug-in electric vehicle ownership for both the Control and Test treatments of the Content Test and the estimate from the 2020 RECS.

Table 6. Electric Vehicle Ownership - 2022 ACS Content Test vs 2020 RECS

Rate	ACS - Test Percent	ACS - Control Percent	RECS Percent
Electric Vehicle Ownership	4.1 (0.2)	2.5 (0.2)	1.5 (<0.1)

Source: U.S. Census Bureau, 2022 American Community Survey (ACS) Content Test and 2020 Residential Energy Consumption Survey (RECS) | DRB No. CBDRB-FY23-ACSO003-B0064
Note: Standard errors are in parentheses.

Due to methodological and design differences between the ACS and RECS, we could not perform statistical testing and can only compare these rates nominally.

5.2 Item Missing Data Rate Results for Electric Vehicles

RQ2. Is the overall item missing data rate different between Version 1 and Version 2?

We compared the item missing data rates for the Control and Test treatments using a two-sided t-test. Table 7 provides the overall item missing data rates.

Table 7. Item Missing Data Rates - Overall

Mode	Test Percent	Control Percent	Difference	P-value
Overall	0.2 (0.1)	0.5 (0.1)	-0.4 (0.1)	<0.01*

Source: U.S. Census Bureau, 2022 American Community Survey Content Test | DRB No. CBDRB-FY23-ACSO003-B0064
Note: Minor additive discrepancies are due to rounding. Significance was tested using a two-tailed t-test at the α=0.1 level. An asterisk (*) indicates a statistically significant result.

The Control treatment had higher item missing data rates than the Test treatment.

RQ3. Are the item missing data rates different between Version 1 and Version 2 when dividing responses into self-administered (paper and internet) versus interviewer-administered (CAPI)?

We also compared the item missing data rates when dividing responses into self-administered versus interviewer-administered.

Table 8. Item Missing Data Rates by Mode

Mode	Test Percent	Control Percent	Difference	Adjusted P-value
Self-response	0.2 (<0.1)	0.7 (0.1)	-0.5 (0.1)	<0.01*
CAPI	0.3 (0.2)	<0.1 (0.1)	0.3 (0.2)	0.12

Source: U.S. Census Bureau, 2022 American Community Survey Content Test | DRB No. CBDRB-FY23-ACSO003-B0064
Note: Minor additive discrepancies are due to rounding. Standard errors are in parentheses. Significance was tested using a two-tailed t-test at the α=0.1 level. An asterisk (*) indicates a statistically significant result. P-values have been adjusted for multiple comparisons using the Hochberg method.

The Control treatment had higher item missing data rates for the self-response modes, but we found no statistically significant difference in the item missing data rates for CAPI.

5.3 Response Distribution Results for Electric Vehicles

In section 5.3, any reference to electric vehicle ownership indicates ownership of a plug-in electric vehicle only (no hybrids were included).

RQ4. Are the percentages of households with a plug-in electric vehicle different between Version 1 and Version 2?

Table 9 provides the rates of plug-in electric vehicle ownership overall and by mode.

Table 9. Electric Vehicle Ownership by Mode

Mode	Test Percent	Control Percent	Difference	Adjusted P-value
Overall	4.1 (0.2)	2.5 (0.2)	1.6 (0.3)	<0.01*
Internet	4.6 (0.3)	2.9 (0.2)	1.7 (0.3)	<0.01*
Mail	1.7 (0.3)	1.2 (0.2)	0.5 (0.4)	0.20
CAPI	4.4 (0.7)	2.3 (0.5)	2.1 (0.8)	0.03*

Source: U.S. Census Bureau, 2022 American Community Survey Content Test | DRB No. CBDRB-FY23-ACSO003-B0064
Note: Minor additive discrepancies are due to rounding. Standard errors are in parentheses. Significance was tested using a two-tailed t-test at the α=0.1 level. An asterisk (*) indicates a statistically significant result. P-values have been adjusted for multiple comparisons using the Hochberg method.

The Test treatment had higher rates of plug-in electric vehicle ownership for internet, CAPI, and overall. There was no statistically significant difference in the rates for the mail mode.

We also compared the distributions of building type, tenure, year built (house), household size, and household income for electric vehicle owners using Rao-Scott chi-square tests (first-order adjustment).

Table 10 shows the distribution of building type for electric vehicle owners.

Table 10. Distribution of Building Type for Electric Vehicle Owners

Building Type	Test Percent	Control Percent	Chi-square	P-value
House	80.6 (2.3)	81.4 (3.0)	3.4	0.33
Small apartment	9.9 (1.5)	9.3 (2.2)		
Large apartment	9.5 (1.8)	8.4 (2.3)		
Unknown	<0.1 (<0.1)	0.9 (0.5)		

Source: U.S. Census Bureau, 2022 American Community Survey Content Test | DRB No. CBDRB-FY23-ACSO003-B0064
Note: Minor additive discrepancies are due to rounding. Standard errors are in parentheses. Significance was tested using a chi-square test at the α=0.1 level.

We found no statistically significant difference in the distributions of building type between Test and Control.

Table 11 shows the distribution of tenure for electric vehicle owners.

Table 11. Distribution of Tenure for Electric Vehicle Owners

Tenure	Test Percent	Control Percent	Chi-square	P-value
Owned	71.2 (2.6)	78.0 (3.3)	2.0	0.36
Rented	27.5 (2.5)	20.9 (3.2)		
Unknown	1.3 (0.7)	1.1 (1.1)		

Source: U.S. Census Bureau, 2022 American Community Survey Content Test | DRB No. CBDRB-FY23-ACSO003-B0064
Note: Minor additive discrepancies are due to rounding. Standard errors are in parentheses. Significance was tested using a chi-square test at the α=0.1 level.

We found no statistically significant difference in the distributions of tenure between Test and Control.

Table 12 shows the distributions of year built for electric vehicle owners.

Table 12. Distribution of Year Built for Electric Vehicle Owners

Year built	Test Percent	Control Percent	Chi-square	P-value
2000 or later	33.8 (2.2)	34.9 (2.9)	3.1	0.54
1980 to 1999	22.9 (2.4)	18.1 (3.1)		
1950 to 1979	26.4 (2.5)	26.5 (3.4)		
1949 or earlier	12.2 (2.2)	12.5 (2.0)		
Unknown	4.7 (1.4)	8.0 (2.1)		

Source: U.S. Census Bureau, 2022 American Community Survey Content Test | DRB No. CBDRB-FY23-ACSO003-B0064
Note: Minor additive discrepancies are due to rounding. Standard errors are in parentheses. Significance was tested using a chi-square test at the α=0.1 level.

We found no statistically significant difference in the distributions of year built between Test and Control.

Table 13 shows the distributions of household size for electric vehicle owners.

Table 13. Distribution of Household Size for Electric Vehicle Owners

Household size	Test Percent	Control Percent	Chi-square	P-value
1 person	16.6 (2.3)	10.9 (2.5)	5.5	0.14
2 persons	35.9 (2.6)	36.6 (3.7)		
3 to 5 persons	45.2 (2.9)	47.3 (3.6)		
6 or more persons	2.3 (0.7)	5.1 (1.8)		

Source: U.S. Census Bureau, 2022 American Community Survey Content Test | DRB No. CBDRB-FY23-ACSO003-B0064Note: Minor additive discrepancies are due to rounding. Standard errors are in parentheses. Significance was tested using a chi-square test at the α=0.1 level.

We found no statistically significant difference in the distributions of household size between Test and Control.

Because we do not know which person or persons in the household uses the electric vehicle, we decided to sum all incomes within the household into a total household income. Table 14 shows the distribution of total household income for electric vehicle owners.

Table 14. Distribution of Total Household Income for Electric Vehicle Owners

Total income	Test Percent	Control Percent	Chi-square	P-value
Less than $25,000	11.0 (1.8)	9.2 (1.6)	17.1	0.02*
$25,000 to $49,999	6.8 (1.4)	9.6 (2.2)		
$50,000 to $74,999	7.8 (1.8)	5.8 (1.4)		
$75,000 to $99,999	11.5 (1.9)	5.0 (1.3)		
$100,000 to $149,999	10.3 (1.7)	14.7 (2.8)		
$150,000 to $199,999	8.6 (1.7)	9.9 (2.2)		
$200,000 or more	25.6 (2.4)	33.2 (2.9)		
Unknown	18.4 (2.3)	12.5 (2.6)		

Source: U.S. Census Bureau, 2022 American Community Survey Content Test | DRB No. CBDRB-FY23-ACSO003-B0064
Note: Minor additive discrepancies are due to rounding. Standard errors are in parentheses. Significance was tested using a chi-square test at the α=0.1 level. An asterisk (*) indicates a statistically significant result.

The chi-square test showed that there was a difference in the distribution of total household income between Test and Control, so we performed a two-sided t-test on these distributions (Table 15).

Table 15. T-tests for Total Household Income for Electric Vehicle Owners

Total income	Test Percent	Control Percent	Difference	Adjusted P-value
Less than $25,000	11.0 (1.8)	9.2 (1.6)	1.9 (2.6)	0.64
$25,000 to $49,999	6.8 (1.4)	9.6 (2.2)	-2.8 (2.5)	0.64
$50,000 to $74,999	7.8 (1.8)	5.8 (1.4)	1.9 (2.1)	0.64
$75,000 to $99,999	11.5 (1.9)	5.0 (1.3)	6.5 (2.2)	0.03*
$100,000 to $149,999	10.3 (1.7)	14.7 (2.8)	-4.5 (3.2)	0.64
$150,000 to $199,999	8.6 (1.7)	9.9 (2.2)	-1.3 (2.7)	0.64
$200,000 or more	25.6 (2.4)	33.2 (2.9)	-7.7 (3.8)	0.28
Unknown	18.4 (2.3)	12.5 (2.6)	5.9 (3.0)	0.28

Source: U.S. Census Bureau, 2022 American Community Survey Content Test | DRB No. CBDRB-FY23-ACSO003-B0064
Note: Minor additive discrepancies are due to rounding. Standard errors are in parentheses. Significance was tested using a two-tailed t-test at the α=0.1 level. An asterisk (*) indicates a statistically significant result. P-values have been adjusted for multiple comparisons using the Hochberg method.

The percent of electric vehicle owners who had a total income of $75,000 to $99,999 was 6.5 percentage points higher for the Test treatment. There were no other statistically significant differences in the distribution of total household income between Test and Control.

5.4 Response Reliability Results for Electric Vehicles

RQ5. Are the measures of response reliability different between Version 1 and Version 2?

To measure response reliability, we compared GDRs and IOIs between the Control and Test questions.

Table 16 provides the GDRs overall and mode.

Table 16. Gross Difference Rates for Electric Vehicles Question

Mode	Test GDR	Control GDR	Difference	Adjusted P-value
Overall	3.5 (0.4)	1.2 (0.3)	2.2 (0.4)	<0.01*
Internet	3.8 (0.5)	1.4 (0.4)	2.4 (0.6)	<0.01*
Mail	2.4 (0.6)	1.0 (0.4)	1.4 (0.7)	0.06*
CAPI	3.2 (0.8)	0.9 (0.3)	2.3 (0.9)	0.02*

Source: U.S. Census Bureau, 2022 American Community Survey Content Test | DRB No. CBDRB-FY23-ACSO003-B0064
Note: Minor additive discrepancies are due to rounding. Standard errors are in parentheses. Significance was tested using a two-tailed t-test at the α=0.1 level. An asterisk (*) indicates a statistically significant result. P-values have been adjusted for multiple comparisons using the Hochberg method.

The Test treatment had higher GDRs for all modes and overall. This would suggest that the Test treatment question (Version 2) is less reliable than the Control question. However, GDRs may not be the best measure of response reliability when it comes to rare occurrences (such as electric vehicle ownership). A better measure of response reliability in these instances would be the index of inconsistency (IOI). IOI metrics take into account the relative size of categories (in this case Yes versus No).

We wanted to measure response reliability by comparing IOIs, but first we needed to test the assumption of parallel measures. If the errors in the original interview and CFU reinterview are positively or negatively correlated, then our IOI rates may be biased and we need to adjust the rates.

We checked this assumption by testing if the net difference rate (NDR) was significantly different from zero separately for each test.

Table 17 provides the results of testing the assumption of parallel measures.

Table 17. NDRs per treatment

Treatment	NDR	P-value
Control	0.2 (0.2)	0.50
Test	-1.5 (0.4)	<0.01*

Source: U.S. Census Bureau, 2022 American Community Survey Content Test | DRB No. CBDRB-FY23-ACSO003-B0064
Note: Standard errors are in parentheses. Significance was tested using a two-tailed t-test at the α=0.1 level with a null hypothesis of NDR=0. An asterisk (*) indicates a statistically significant result.

The Test treatment did not meet the assumption of parallel measures, thus we calculated adjusted IOI rates instead.

Table 18 provides the adjusted IOI rates overall and by mode.

Table 18. Adjusted Index of Inconsistency Rates for Electric Vehicles Question

Mode	Test IOI-Adjusted	Control IOI-Adjusted	Difference	Adjusted P-value
Overall	34.9 (3.5)	30.1 (5.1)	4.8 (5.6)	0.95
Internet	32.9 (4.0)	28.7 (5.7)	4.2 (6.6)	0.95
Mail	49.8 (11.4)	51.2 (17.5)	-1.4 (21.7)	0.95
CAPI	36.4 (9.9)	25.6 (16.8)	10.8 (19.5)	0.95

Source: U.S. Census Bureau, 2022 American Community Survey Content Test | DRB No. CBDRB-FY23-ACSO003-B0064
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. P-values have been adjusted for multiple comparisons using the Hochberg method.

When looking at the adjusted IOI rates, we found no statistically significant difference in response reliability between the Test and Control questions.

5.5 Other Metrics Results for Electric Vehicles

Measures of respondent burden were also analyzed for the 2022 Content Test. Metrics of interest were completion times, help screen access rates, and breakoff rates. These metrics were looked at overall and by each Content Test topic. There were no statistically significant differences found for the Electric Vehicles question between the Control and Test treatments. For more information see Virgile et al (2023).

6 CONCLUSIONS AND RECOMMENDATIONS

Determining the prevalence of electric vehicles in homes allows for the planning of future energy needs. Energy infrastructure can be updated and expanded at the local level when there is a better understanding of the increased energy demand for electric vehicle charging stations. This can not only help to provide the appropriate infrastructure where it is needed, but it can also help avoid blackouts and other forms of energy crises.

The analysis for the research questions on this topic yielded several noteworthy results.

When nominally comparing the electrical vehicle rates from the Content Test to the RECS, we found the Content Test rates to be reasonable. Also, when comparing the Test and Control treatments independent of the RECS question, the Test treatment had higher rates of plug-in electric vehicle ownership in the internet and CAPI modes and overall. There was no statistically significant difference in the rates for the mail mode.

A comparison between the Test and Control treatments also showed the Control treatment had higher item missing data rates overall. When these results were further examined, we found the higher item missing data rates for the Control treatment stemmed from the self-response modes. There was no statistically significant difference between the Test and Control in the item missing data rates for CAPI.

When the distributions of electric vehicle ownership by building type, tenure, year built (house), household size, and household income were compared between the Control and Test treatments, the only difference occurred among electric vehicle owners who had a total income of $75,000 to $99,999. The Test treatment had a higher percentage of these electric vehicle owners.

Finally, when response reliability was measured using adjusted IOI rates, there was no statistically significant difference found between the Test and Control questions. This method is preferable to using GDRs to measure response reliability for this type of question, which measures rare occurrences (such as electric vehicle ownership).

Based on the results of the research questions, we recommend the adoption of the Test question from the Electric Vehicle Content Test to the American Community Survey. In addition to the findings mentioned in the previous section, this version of the question is currently used by the U.S. Energy Information Administration on the Residential Energy Consumption Survey, a federal survey responsible for collecting, analyzing, and disseminating data on energy usage.

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. (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/library/working-papers/2023/acs/2023_Virgile_01.html.