



November 17, 2023

2023 AMERICAN COMMUNITY SURVEY RESEARCH AND EVALUATION REPORT MEMORANDUM  
SERIES # ACS23-RER-06

MEMORANDUM FOR Donna M. 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

Peter Massarone *PM*  
Decennial Statistical Studies Division

Matthew Virgile *MV*  
Center for Behavioral Science Methods

Subject: 2022 American Community Survey Content Test Evaluation Report:  
Sewer

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

If you have any questions about this report, please contact Molly Cromwell at 301-763-3379 or Peter Massarone at [peter.j.massarone.jr@census.gov](mailto:peter.j.massarone.jr@census.gov).

Attachment

cc:

[acso.re.workgroup.list@census.gov](mailto:acso.re.workgroup.list@census.gov)

Lauren Contard

DSSD

Lindsay Longsine

Broderick Oliver

Michael Risley

Samantha Spiers

Kevin Bragdon

Patrick Madamba

Hyon Shin

SEHSD

*This page is intentionally blank*

---

# 2022 American Community Survey Content Test Evaluation Report: Sewer

FINAL REPORT



**Molly Cromwell**  
Social, Economic, and Housing Statistics Division

**Peter Massarone**  
Decennial Statistical Studies Division

**Matthew Virgile**  
Center for Behavioral Science Methods

*This page is intentionally blank*

---

TABLE OF CONTENTS

**EXECUTIVE SUMMARY ..... iii**

**1 BACKGROUND ..... 1**

1.1 Proposals for New and Revised ACS Questions ..... 1

1.2 Cognitive Testing..... 2

1.3 Field Testing Sewer in the 2022 ACS Content Test ..... 2

1.3.1 Justification for Inclusion of Sewer in the Content Test ..... 2

1.3.2 Cognitive Testing Development for Sewer ..... 3

1.3.3 Question Content ..... 4

1.3.4 Research Questions..... 5

**2 METHODOLOGY ..... 5**

2.1 Sample Design..... 5

2.2 Data Collection..... 6

2.3 Content Follow-Up Operation ..... 8

2.3.1 Content Test Follow-Up Protocol ..... 8

2.3.2 Content Test Follow-Up for Sewer ..... 9

2.4 Analysis Metrics ..... 9

2.4.1 Unit-Level Analysis..... 10

2.4.2 Topic-Level Analysis..... 10

**3 DECISION CRITERIA ..... 15**

**4 ASSUMPTIONS AND LIMITATIONS ..... 15**

4.1 Assumptions ..... 15

4.2 Limitations ..... 16

**5 RESULTS ..... 18**

5.1 Benchmark Results for Sewer ..... 18

5.2 Item Missing Data Rate Results for Sewer..... 20

5.3 Response Distribution Results for Sewer ..... 21

5.4 Response Reliability Results for Sewer..... 21

5.5 Other Metric Results for Sewer ..... 22

**6 CONCLUSIONS AND RECOMMENDATIONS..... 23**

**7 ACKNOWLEDGMENTS ..... 24**

**8 REFERENCES ..... 25**

**LIST OF TABLES**

Table 1. Questions by Treatment..... 6  
Table 2. Original Interview and CFU Reinterview Counts for Calculating GDR, IOI, and NDR ..... 12  
Table 3. Cross-Tab of Original Interview and CFU Results: Questions with Response Categories ..... 14  
Table 4. Decision Criteria for Sewer ..... 15  
Table 5. Type of Sewage System – 2022 ACS Content Test vs. 2021 American Housing Survey..... 19  
Table 6. Item Missing Data Rates by Tenure – Item vs. Telephone Availability ..... 21  
Table 7. Sewer Gross Difference Rates and Adjusted Indexes of Inconsistency..... 22

**LIST OF FIGURES**

Figure 1. Control and Test: Paper Version Sewer..... 4  
Figure 2. Control and Test: Internet Version Sewer ..... 4  
Figure 3. Control and Test: CAPI Version Sewer ..... 4

## **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 Sewer.

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 follow-up 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 Sewer question is a new question proposed for inclusion in the ACS. Knowledge about the status of amenities with a direct link to public health is vital to preserving the well-being of our nation. Including the Sewer question on the ACS provides data on the wastewater infrastructure status in rural and small communities, which allows for the monitoring of water quality. This question was requested by both the Environmental Protection Agency (EPA) and Food and Drug Administration (FDA) due to its role in assessing overall public health quality.

The paper version of the Sewer question that was tested asked respondents, “Is this house, apartment, or mobile home connected to a public sewer?” with response options of “Yes, connected to a public sewer”, “No, connected to a septic system”, or “No, use other type of system”. The internet version was the same except for the auto-filled housing unit type in the question stem. Households that responded via CAPI were asked a yes-or-no question about whether the housing unit was connected to a public sewer and were then conditionally asked a

## **DRB Clearance Number – CBDRB-FY23-ACSO003-B0060**

yes-or-no question about whether the housing unit was connected to a septic tank if they responded with a “No”, “Don’t Know”, or “Refused” to the first question.

The analysis of sewer data collected in the Content Test was guided by research questions focusing on results for response distributions compared to benchmarks, item missing data rates, and response reliability. Data from the 2021 American Housing Survey (AHS) was also used in this analysis for the benchmark comparisons.

- The proportion of households reporting a public sewer as their sewage system was nominally higher in the AHS than in the Content Test, while the Content Test shows nominally higher proportions of households reporting a septic tank or some other type of sewage system.
- Although item missing data rates for the Sewer question are significantly higher overall than the item missing data rates for the telephone availability question, which follows the Sewer question and has similar response categories, the overall difference was less than a percentage point.
- In terms of response reliability, the public sewer and septic tank items were considered to have a low level of inconsistency in responses to the original interview and reinterview overall, while other sewage systems were overall considered to have a high level of inconsistency.

Based on the results as summarized above, the recommendation is to adopt the Sewer question. Numerous federal agencies and user groups have requested these data for their own requirements, and the ACS can provide estimates at the level of geographic specificity that is required.



## **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 Sewer.

### **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.<sup>1</sup> Cognitive interviews were conducted virtually, in English and Spanish.<sup>2</sup> 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 Sewer in the 2022 ACS Content Test**

### **1.3.1 Justification for Inclusion of Sewer in the Content Test**

The Sewer question is a new question proposed for inclusion in the ACS. The Sewer question appeared most recently on the 1990 Census long-form questionnaire. Between 1940 and 2000, the decennial censuses utilized two forms for data collection: the short-form questionnaire and the long-form questionnaire. Most households received the short-form questionnaire while one in every six households received the long-form questionnaire, which allowed for the collection of more detailed demographic and household characteristics. Beginning in 2005, the ACS began

---

<sup>1</sup> 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.

<sup>2</sup> 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.

collecting these long-form data. As a result, beginning with the 2010 Census, the decennial census collects only short-form questionnaire data.

While the 1990 Census was the last time sewer data had been collected on the Census, the American Housing Survey (AHS) has been consistently asking respondents about their public sewer and septic tank status since the 1970s.<sup>3</sup> However, the lowest level of geography available from the AHS is Metropolitan Statistical Areas. Consistent data on the decentralized wastewater infrastructure status in rural and other small communities is needed to protect public health and water quality. As such, the sewer question was requested by both the Environmental Protection Agency (EPA) and Food and Drug Administration (FDA) due to its role in assessing overall public health quality.

Due to the level of geography at which the ACS data can be collected, adding this question to the ACS may help to address changes in housing development and to support regular planning and funding cycles at the local, state, and national level. Additionally, determining the prevalence of existing septic systems and obtaining periodic updates on new septic system construction at the local level would assist in meeting the country's growing infrastructure needs.

### **1.3.2 Cognitive Testing Development for Sewer**

During the first round of the cognitive testing process, two versions of the sewer question were presented. In the first version, respondents were asked, “What is the MAIN type of sewage disposal for this house, apartment, or mobile home?” with response options of “Public sewer”, “Septic system or cesspool”, or “Other type of sewage disposal”. In the second version, respondents were asked, “Is this house, apartment, or mobile home connected to a public sewer?” with response options of “Yes, connected to a public sewer”, “No, connected to septic system or cesspool”, or “No, use other type of system”.

After the first round of cognitive testing, the word “cesspool” was dropped from response options, and the wording from the original Version 2 became the preferred wording. For all additional rounds of cognitive testing and for the field test, the only version of the question tested asked respondents, “Is this house, apartment, or mobile home connected to a public sewer?” with response options of “Yes, connected to a public sewer”, “No, connected to a septic system”, or “No, use other type of system”.

---

<sup>3</sup> For more general information about the American Housing Survey, see U.S. Census Bureau (2021).

### 1.3.3 Question Content

This is a potential new question on the ACS. Only one version of this question will be tested. The paper version of this question is shown in Figure 1. The internet version of this question is shown in Figure 2. The Computer-Assisted Personal Interviewing (CAPI) version of this question is shown in Figure 3. In the CAPI version, respondents are only asked Part B of the question if the response to Part A is “No”. In the instance of a respondent answering “No” for both Part A and Part B, it can be considered equivalent to the “No, use other type of system” response in the paper and internet versions.

**Figure 1. Control and Test: Paper Version Sewer**

**Is this house, apartment, or mobile home connected to a public sewer?**

- Yes, connected to public sewer
- No, connected to septic tank
- No, use other type of system

**Figure 2. Control and Test: Internet Version Sewer**

**Is this <mobile home/house/apartment/unit> connected to a public sewer?**

- Yes, connected to public sewer
- No, connected to septic tank
- No, use other type of system

**Figure 3. Control and Test: CAPI Version Sewer**

**14a. Is this <mobile home/house/apartment/unit> connected to a public sewer?**

- Yes
- No

**14b. Is it connected to a septic tank?**

- Yes
- No

### **1.3.4 Research Questions**

The questions examined for this research are presented below.

**RQ 1.** Is the proportion of households connected to a public sewer, septic tank, or other type of sewage system in the 2022 ACS Content Test different from the respective proportions in the American Housing Survey (AHS)?

**RQ 2.** What are the missing item data rates for the Sewer question, overall and by these subgroups: renters and homeowners? How do these rates compare to the item missing data rate for the telephone availability question that follows the Sewer question on the questionnaire and has similar response categories?

**RQ 3.** Are the answers provided by respondents to the Sewer question in the original interview consistent with the answers they provided in the Content Follow-Up (CFU)?

## **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

## DRB Clearance Number – CBDRB-FY23-ACSO003-B0060

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.<sup>4</sup>

**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 <sup>†</sup>
Labor Force	Production	Test Version 1	Test Version 2
Income	Production	Test Version 1	Test Version 2

<sup>†</sup> The SNAP Test Version will be in both test treatments to align with Labor Force and Income that also have a reference period change to the previous calendar year.

## 2.2 Data Collection

<sup>4</sup> 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.

## DRB Clearance Number – CBDRB-FY23-ACSO003-B0060

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.<sup>5</sup>
- 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.<sup>6</sup> 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

---

<sup>5</sup> 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.

<sup>6</sup> The interviewer did not know which treatment the caller was in and therefore administered the production questionnaire. In 2019, less than one percent (0.6%) of cases responded by TQA and had no other response in a different mode. Based upon this rate, we projected about 744 TQA-only responses would be excluded from the 2022 ACS Content Test analysis.

incomplete information on the form or reported more than five people on the roster of a paper questionnaire.<sup>7</sup>

- 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

---

<sup>7</sup> 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.



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.<sup>8</sup>

### **2.3.2 Content Test Follow-Up for Sewer**

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

## **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 Sewer 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 Sewer 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.<sup>9</sup>

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.<sup>10</sup>

---

<sup>8</sup> 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.

<sup>9</sup> 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.

<sup>10</sup> 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).

Comparisons between the Control and Test versions of Sewer were conducted using a two-tailed t-test at the  $\alpha=0.1$  level of significance. The Content Test sample size was chosen to provide enough statistical power (0.80) to detect a difference in the gross difference rates (measuring differences in adds and deletes from the household roster) of at least two percentage points between the Control and test groups for the Household Roster question.<sup>11</sup> 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).<sup>12</sup>

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 ( $X$ ) is the square root of the variance:

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

where:

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

### **2.4.1 Unit-Level Analysis**

The unit response rate is important, as it provides an indication of the quality of the survey data. As part of our analysis, we examined unit-level (i.e., address-level) responses for the Control and test treatments in the original interviews and CFU reinterviews. These results are provided in a separate report (Spiers et al., 2023).<sup>13</sup>

### **2.4.2 Topic-Level Analysis**

To evaluate the accuracy of the data collected from the Sewer question, we calculated a variety of metrics, presented in Sections 2.4.2.1 through 2.4.2.6.

#### **2.4.2.1 Benchmarks**

To roughly gauge the accuracy of the responses to Sewer, we compared select estimates derived from these data to similar estimates from the AHS, a reliable source (i.e., a benchmark) of household data.

---

<sup>11</sup> See Section 2.4.2.4 for the definition of Gross Difference Rate.

<sup>12</sup> Use the MULTTEST Procedure in SAS®.

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

## **DRB Clearance Number – CBDRB-FY23-ACSO003-B0060**

We compared national level estimates and subgroup estimates like building type and tenure (owner versus renter) for occupied households from both surveys. The comparisons were nominal only due to differences in sampling, data collection, question wording, and other factors. The AHS asks respondents whether their home is connected to a public sewer, and, if the answer to that question is “no”, asks respondents to identify if their home sewage disposal system is a septic tank, a cesspool, a chemical toilet, an outhouse or privy, some other sewage disposal system (with an instruction to specify), or no sewage disposal system. Respondents that identify their sewage disposal system as either a septic tank or cesspool are asked to further specify whether they have a standard septic tank and subsurface leach field (the most common type), a pump to distribute wastewater, a system elevated above the natural soil surface, a system that applies treated wastewater, or some other system that is not listed.

### **2.4.2.2 Item Missing Data Rates**

To measure nonresponse to Sewer, we calculated its item missing data rate, the proportion of eligible housing units for which a required response to the Sewer question 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.

For sewer, the singular Sewer question asked to self-respondents and the two-part question asked to CAPI respondents, Sewer (A) and Septic Tank (B), were of interest. The universe for item missing data analysis for the Sewer question was all occupied households that responded by mail or internet, while the universe for Sewer (A) was similarly all occupied households that responded by CAPI. The universe for Septic Tank (B) was those respondents that were asked the Septic Tank (B) question, which was conditionally asked to CAPI respondents depending on their answer to the Sewer (A) question. No response, responses of “Don’t Know”, and responses of “Refused” were considered missing for the purposes of calculating item missing data rates.

We compared item missing data rates for the Sewer questions to the item missing data rate for the telephone availability question via two-tailed t-tests. The telephone availability question followed the Sewer question on the questionnaire and had similar response categories, and the expectation was the item missing data rates would be comparable between the questions.

### **2.4.2.3 Response Distributions**

No metrics on response distributions were calculated for this research beyond those calculated for benchmark comparisons.

### **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

**DRB Clearance Number – CBDRB-FY23-ACSO003-B0060**

breakoffs during the CFU operation. For Sewer, 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<sub>L</sub>)

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

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

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

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

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

- a = units in category for both interview and reinterview
- b = units not in category for original interview, but in category for reinterview
- c = units in category for original interview, but not in category for reinterview
- d = units in category for neither interview nor reinterview
- n = total units in the universe = a + b + c + d

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

We calculated the GDR for this response category as:

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

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

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

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

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

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

where:

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

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

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

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

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

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

So far, we have only discussed response reliability with respect to single response categories. If a question has three or more response categories (or “comparison categories” in cases where it is necessary to collapse some response categories for comparison), we could also have measured the overall response reliability of a question using the L-fold index of inconsistency,  $IOI_L$ . It is possible to look at the difference in  $IOI_L$  between treatments and test for significance as with the single category measures. We did not examine  $IOI_L$  for this topic.

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

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

		Original Interview categories						
		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

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

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

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

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

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

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

### 3 DECISION CRITERIA

Before field testing Sewer, 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 Sewer would be recommended for inclusion in the ACS. The decision criteria for Sewer are presented in Table 4.

**Table 4. Decision Criteria for Sewer**

Priority	Research Question	Decision Criteria
1	1	The proportion of households connected to a public sewer, septic tank, or other type of sewage is about the same nominally as the respective proportions in the American Housing Survey (AHS).
2	2	The overall item missing data rate is about the same nominally as the item missing data rate for the telephone availability question that follows the Sewer question on the questionnaire.
3	3	The measure of simple response variance, GDR and IOI do not provide an indication of potential problems with questionnaire wording.

## 4 ASSUMPTIONS AND LIMITATIONS

### 4.1 Assumptions

The sample addresses for the Control and test treatments were selected in a manner so that their response propensities and response distributions would be the same. This assumption of homogeneity allows us to conclude that any difference between

treatments is attributable to differences in wording and format. See Section 5 for more details.

- There was no difference between treatments in mail delivery timing or subsequent response time. The treatments had the same sample size and used the same postal sort and mailout procedures. Previous research indicated that postal procedures alone could cause a difference in response rates at a given point in time between experimental treatments of different sizes, with response for the smaller treatments lagging (Heimel, 2016).
- We assume that the frequency of real changes in answers due to a change in life circumstances between the original interview and CFU reinterview were similar between treatments.

#### **4.2 Limitations**

- GQs were not included in the sample for the 2022 ACS Content Test. The results of the Content Test may not extend to GQ populations.
- Housing units from Alaska, Hawaii, and Puerto Rico were not included in the sample for the 2022 ACS Content Test. The results of the Content Test may not extend to the housing unit population in these areas.
- The paper questionnaire was only available in English and was not available in Spanish like in production. The Content Test results related to the English paper questionnaire may not extend to Spanish paper questionnaire.
- We did not have response data for some partial internet responses (179 cases) due to a server issue. These cases were excluded from the analyses.
- TQA responses were excluded from the analysis of the 2022 ACS Content Test response data because survey responses completed via the TQA operation were only conducted using the ACS production data collection instrument.
- CAPI interviewers were assigned 2022 ACS Content Test cases as well as regular production cases. The potential risk of this approach is the introduction of a cross-contamination or carry-over effect among Control and test treatments and production due to the same interviewer administering multiple versions of the same question item (despite their training to read questions verbatim).
- Due to budget constraints, the CAPI workload could not exceed 28,000 housing units. This workload was less than what was subsampled originally because we over-sampled



addresses in low response areas. Limiting the CAPI workload caused an increase in the variances for the analysis metrics used.

- The CFU reinterviews were conducted by phone only, whereas the original interviews were completed online, by mail, by phone in CAPI, and in person in CAPI. Hence, some of the differences observed between the original interviews and the CFU interviews may be the result of mode effect.
- Not all households who provided a response in the original interview were eligible for the CFU reinterview (see Section 2.3 for more information). As a result, 2.5 percent (standard error 0.2) of households from the original Control interviews, 2.5 percent (standard error 0.2) of households from the original Test interviews, and 3.0 percent (standard error 0.2) of households from the original Roster Test interviews were not eligible for the CFU reinterview. These rates were not significantly different between treatments (chi-square p-value 0.11).
- We reinterviewed the same person who responded in the original interview when possible, but accepted interviewing a different person from the same household after two unsuccessful attempts at reaching the original person. Therefore, differences in results between the original interview and CFU reinterview for these cases could partly be from different people answering the questions. We interviewed a different household member in CFU for 7.3 percent (standard error 0.4) of CFU Control cases, 9.4 percent (standard error 0.5) of CFU Test cases, and 8.5 percent (standard error 0.5) of CFU Roster Test cases. These rates were significantly different between treatments (chi-square p-value 0.01) with the rate of CFU Test cases (t-test p-value <0.01) and CFU Roster Test cases (t-test p-value 0.04) being significantly higher than the rate of CFU Control cases.
- We examined potential differences between CFU respondents and 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 Sewer. 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. These results are presented in Spiers et al. (2023), while the salient results for Sewer are summarized below. Analysis on respondent burden was also conducted to ensure that any changes to ACS content will not cause undue burden on future respondents. These results are presented in Virgile et al. (2023), while the salient results for Sewer are summarized below.

Given that the Sewer analysis called for the combination of the Control and Test treatments to create one sample for analysis, and thus no comparisons between treatments were conducted, the only area of potential concern with unit-level analysis results would be results regarding differences in building tenure between the Control and Test treatments. Analysis of the response distributions of demographic and socioeconomic characteristics (age, sex, Hispanic origin, race, building tenure, and language of response) showed no difference in response distributions between the Test and Control treatments by tenure overall, negating any potential concerns regarding the combination of treatments in the Sewer analysis.

The median time of completion for the Sewer question was 6 seconds for both the internet and CAPI response modes for all treatments. For internet respondents, help screen access rates were 1.7 percent, 1.8 percent, and 1.6 percent for the Control, Test, and Roster treatments, respectively. Breakoff rates were 0.1 percent for Control, and less than 0.1 percent for both the Test and Roster treatments.

### **5.1 Benchmark Results for Sewer**

Research Question 1

*Is the proportion of households connected to a public sewer, septic tank, or other type of sewage system in the 2022 ACS Content Test different from the respective proportions in the American Housing Survey (AHS)?*

Table 5 shows the proportion estimates for the type of sewage system for the combined Test and Control treatments in the 2022 ACS Content Test and the 2021 American Housing Survey. Due to many factors including differences in sampling, data collection, and question wording, the results of these surveys can only be compared nominally. Among housing units overall, as well in subgroups based on tenure and building type, the proportion of households with a public sewer is several percentage points higher in the AHS than in the Content Test.

**DRB Clearance Number – CBDRB-FY23-ACSO003-B0060**

Meanwhile, the proportion of households with a septic tank, some other sewage system, or an unknown sewage system is higher in the Content Test than in the AHS.

There are several things to keep in mind when comparing the Content Test and AHS data. First the AHS data are edited and imputed, which would impact several things such as the unknowns from the field test would be allocated into one of the other categories. Additionally, since the AHS is a longitudinal survey, it uses past survey responses to confirm that a respondent is still using a certain system (i.e., conventional septic tank). Confirming past responses can generally be less burdensome on survey respondents than providing a fresh response. Both considerations help explain why the rates of unknown sewage systems are higher in the Content Test than in the AHS.

Other research conducted combing public records with AHS data have also identified that respondents in the AHS tend to overreport being connected to a public sewer system. For further information on these findings, see Madamba (2023).

**Table 5. Type of Sewage System – 2022 ACS Content Test vs. 2021 American Housing Survey**

<b>Response Category</b>	<b>Content Test</b>	<b>2021 AHS</b>
<b>Overall</b>		
Public Sewer	77.0 (0.3)	84.5 (<0.1)
Septic Tank	18.9 (0.3)	15.2 (<0.1)
Other	1.2 (0.1)	0.1 (0.1)
Unknown	2.9 (0.1)	0.2 (0.2)
<b>Tenure – Homeowner</b>		
Public Sewer	73.9 (0.5)	78.6 (<0.1)
Septic Tank	24.4 (0.5)	21.1 (<0.1)
Other	1.0 (0.9)	0.2 (0.2)
Unknown	0.6 (0.1)	0.1 (0.2)
<b>Tenure – Renter</b>		
Public Sewer	88.0 (0.4)	95.0 (<0.1)
Septic Tank	8.0 (0.4)	4.6 (0.1)
Other	1.5 (0.2)	0.1 (0.4)
Unknown	2.6 (0.2)	0.3 (0.2)
<b>Building – One-Family Home</b>		
Public Sewer	75.2 (0.4)	81.1 (<0.1)
Septic Tank	22.7 (0.4)	18.5 (<0.1)
Other	1.0 (0.1)	- ( )
Unknown	1.2 (0.1)	- ( )
<b>Building – Small Apartment Building</b>		
Public Sewer	93.0 (0.4)	99.8 (<0.1)
Septic Tank	2.8 (0.3)	- ( )
Other	0.9 (0.2)	- ( )
Unknown	3.2 (0.4)	- ( )

<b>Building – Large Apartment Building</b>		
Public Sewer	93.2 (0.6)	98.6 (<0.1)
Septic Tank	1.3 (0.3)	- ( )
Other	2.3 (0.3)	- ( )
Unknown	3.2 (0.4)	- ( )
<b>Building – Mobile Home</b>		
Public Sewer	48.0 (1.8)	61.4 (<0.1)
Septic Tank	47.7 (1.9)	38.1 (0.1)
Other	2.6 (0.5)	- ( )
Unknown	1.7 (0.5)	- ( )

**Source:** U.S. Census Bureau, 2022 American Community Survey Content Test, 2021 American Housing Survey | DRB No. CBDRB-FY23-ACSO003-B0060

Note: Minor additive discrepancies are due to rounding. Standard errors are in parentheses. Unknown response category indicates missing responses for the Content Test, and no sewage system or unreported sewage system for AHS. Cells without estimates indicate estimates and standard errors that were withheld due to not meeting publication standards, or to avoid disclosure.

## **5.2 Item Missing Data Rate Results for Sewer**

### Research Question 2

*What are the missing item data rates for the Sewer question, overall and by these subgroups: renters and homeowners? How do these rates compare to the item missing data rate for the telephone availability question that follows the Sewer question on the questionnaire and has similar response categories?*

For the Sewer question, households that responded via paper or internet questionnaire were asked one question about whether the housing unit was connected to a public sewer and, if not, what type of sewage system was the housing unit connected to (see Figures 1 and 2). These households are represented in Table 6 below by the “Sewer” item. Households that responded via CAPI were asked a yes-or-no question about whether the housing unit was connected to a public sewer, and were then conditionally asked a yes-or-no question about whether the housing unit was connected to a septic tank if they responded with a “No”, “Don’t Know”, or “Refused” to the first question (see Figure 3). The universe for the septic tank question was those respondents that were asked this question, which was conditionally asked of CAPI respondents depending on their answer to the sewer question. These two CAPI questions are represented in Table 6 below by the “Sewer (A)” and “Septic Tank (B)” items. The telephone availability question asked respondents “Can you or any member of this household both make and receive phone calls when at this house, apartment, or mobile home?” which, like the Sewer question, could be answered with “Yes”, “No”, “Don’t Know”, or “Refused”.

Item missing data rates for the Sewer and Septic Tank (B) items were significantly higher than the item missing data rates for the telephone availability item both overall and among the subgroups of homeowners and renters; the difference in the overall item missing data rate between the Sewer item and the telephone availability item was less than a percentage point.

**DRB Clearance Number – CBDRB-FY23-ACSO003-B0060**

The differences between the Septic Tank (B) and telephone availability item missing rates both overall and among the subgroups were relatively high, potentially due to a conditional effect of Septic Tank (B). The item missing data rate for the Sewer (A) item was not significantly different from the telephone availability item overall and among homeowners but was significantly higher among renters. Note that the Sewer (A) question is more comparable to the telephone availability question in terms of the associated Yes/No response options and the unconditional asking of all respondents in occupied housing universe.

**Table 6. Item Missing Data Rates by Tenure – Item vs. Telephone Availability**

<b>Item</b>	<b>Item Percent</b>	<b>Item - Telephone Difference</b>	<b>Adjusted P-value</b>
<b>Overall</b>			
Sewer	3.1 (0.1)	0.9 (0.1)	<0.01*
Sewer (A)	2.2 (0.3)	<0.1 (0.3)	0.89
Septic Tank (B)	8.6 (1.2)	6.3 (1.2)	<0.01*
Telephone Availability	2.2 (0.1)	-	-
<b>Tenure – Homeowner</b>			
Sewer	0.6 (0.1)	0.2 (0.1)	0.03*
Sewer (A)	0.5 (0.2)	0.2 (0.2)	0.44
Septic Tank (B)	1.9 (0.8)	1.5 (0.8)	0.09*
Telephone Availability	0.4 (0.1)	-	-
<b>Tenure – Renter</b>			
Sewer	2.2 (0.2)	1.5 (0.2)	<0.01*
Sewer (A)	3.2 (0.5)	2.6 (0.5)	<0.01*
Septic Tank (B)	21.8 (3.3)	21.1 (3.2)	<0.01*
Telephone Availability	0.7 (0.1)	-	-

**Source:** U.S. Census Bureau, 2022 American Community Survey Content Test | DRB No. CBDRB-FY23-ACSO003-B0060

Note: Minor additive discrepancies are due to rounding. Standard errors are in parentheses. Significance was tested using a two-tailed t-test at the  $\alpha=0.1$  level. An asterisk (\*) indicates a statistically significant result. P-values have been adjusted for multiple comparisons using the Hochberg method. Responses that were missing a value for tenure were not included in either the homeowner or renter tables, but were included in the overall table.

**5.3 Response Distribution Results for Sewer**

Research question 1 adequately addresses the response distributions of interest.

**5.4 Response Reliability Results for Sewer**

Research Question 3

*Are the answers provided by respondents to the Sewer question in the original interview consistent with the answers they provided in the CFU?*

The GDR for a specific response category is the percent of inconsistent answers between the original interview and the reinterview. Table 7 shows the GDR and IOI for each response

## DRB Clearance Number – CBDRB-FY23-ACSO003-B0060

category to the sewer question. Overall and among the subgroups of homeowners and renters, the GDR for public sewer was the largest overall at 4.5 percent, among homeowners at 4.1 percent, and among renters at 5.4 percent.

The IOI for a specific response category measures how much of the total estimate’s variance is due to inconsistency in responses. For these data, the assumption of parallel measures as described in Section 2.4.2.4 is not satisfied, and thus the IOI was adjusted for all items. The IOI for public sewers is 13.0 percent overall and 10.4 percent among homeowners, both of which are considered low levels of inconsistency (less than 20 percent). The 28.9 percent IOI for public sewers among renters is considered a moderate level of inconsistency (20 percent to 50 percent). Septic tanks have a low level of inconsistency overall at 8.6 percent, among homeowners at 7.9 percent, and among renters at 14.9 percent. Meanwhile, other sewage systems, despite having generally low GDRs, have high levels of inconsistency (greater than 50 percent) overall at 82.6 percent, among owners at 78.6 percent, and among renters at 89.2 percent.

**Table 7. Sewer Gross Difference Rates and Adjusted Indexes of Inconsistency**

<b>Response Category</b>	<b>GDR</b>	<b>IOI</b>
<b>Overall</b>		
Public Sewer	4.5 (0.3)	13.0 (0.8)
Septic Tank	2.8 (0.2)	8.6 (0.7)
Other	2.9 (0.2)	82.6 (3.8)
<b>Tenure – Homeowner</b>		
Public Sewer	4.1 (0.3)	10.4 (0.7)
Septic Tank	3.0 (0.3)	7.9 (0.7)
Other	2.5 (0.3)	78.6 (5.0)
<b>Tenure – Renter</b>		
Public Sewer	5.4 (0.6)	28.9 (3.2)
Septic Tank	2.3 (0.4)	14.9 (2.7)
Other	3.6 (0.5)	89.2 (6.3)

**Source:** U.S. Census Bureau, 2022 American Community Survey Content Test | DRB No. CBDRB-FY23-ACSO003-B0060

Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. Responses that were missing a value for tenure were not included in either the homeowner or renter tables, but were included in the overall table.

### 5.5 Other Metric Results for Sewer

There are no other metrics that are requested for Sewer. The prior three research questions adequately address the desired metrics.

## 6 CONCLUSIONS AND RECOMMENDATIONS

The results of the 2022 ACS Content Test show that the proportion of households with a public sewer as the sewage system is nominally lower than previously reported benchmarks from the 2021 AHS. While significance testing was not conducted to compare the Content Test and AHS proportions, it can be observed that the proportions from the Content Test fall outside of the confidence intervals from the AHS. Furthermore, the item missing data rates for the Sewer question are significantly higher overall than the item missing data rates for the telephone availability question, which follows the Sewer question on the questionnaire and has similar response categories. However, this difference still amounts to less than a percentage point. Additionally, respondents that identified their households as being connected to either a public sewer or septic tank were reliable and had low to moderate levels of inconsistency in their responses.

When reviewing the field test results and comparing them to results from the American Housing Survey (AHS), there are several things to keep in mind. First the AHS data are edited and imputed, which would impact several things:

- The unknowns from the field test would be allocated into one of the other categories.
- Since the AHS is a longitudinal survey, it uses past survey responses to confirm that a respondent is still using a certain system (i.e., conventional septic tank). Confirming past responses can generally be less burdensome on survey respondents than providing a fresh response, particularly for renters and respondents living in apartment buildings, groups that reported higher rates of “Unknown” sewage systems in the Content Test.

Additionally, research conducted combing public records with AHS data have identified that respondents in the AHS tend to overreport being connected to a public sewer system. For further information, see Madamba (2023).

Lastly, the nonresponse rates for the AHS mimic that which was found during the field test, especially the differences between units by tenure. Renters reported not knowing their sewage system at a higher rate than owners.

With the above in mind, the recommendation is to adopt this question. The response rates and patterns mimic existing benchmarks, and there are numerous federal agencies and user groups who have requested this data to fulfill their public policy requirements. Currently the only source for this data is the AHS, which can only provide estimates at the national and select metro areas, and not the level of geographic specificity needed.

## 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.
- 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.
- 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>
- Madamba, P. (2023). *Comparing Sewage Disposal Method in the American Housing Survey to Administrative Records in Florida*. Retrieved October 27, 2023, from <https://www.census.gov/library/working-papers/2023/demo/SEHSD-WP-2023-30.html>
- Office of Management and Budget (2006). *Standards and guidelines for statistical surveys*. Retrieved February 24, 2022, from [https://georgewbush-whitehouse.archives.gov/omb/inforeg/statpolicy/standards\\_stat\\_surveys.pdf](https://georgewbush-whitehouse.archives.gov/omb/inforeg/statpolicy/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](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](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*. Retrieved November 27, 2023, from [https://www.census.gov/library/working-papers/2023/acs/2023\\_Spiers\\_01.html](https://www.census.gov/library/working-papers/2023/acs/2023_Spiers_01.html)

**DRB Clearance Number – CBDRB-FY23-ACSO003-B0060**

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.

U.S. Census Bureau. (2021). *American Housing Survey: About*. Retrieved June 8, 2022, from <https://www.census.gov/programs-surveys/ahs/about.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>

Virgile, M., Weng, C., Mills, G., & Spiers, S. (2023). *2022 American Community Survey Content Test evaluation report: Respondent burden*. Retrieved November 27, 2023, from [https://www.census.gov/content/dam/Census/library/working-papers/2023/acs/2023\\_Virgile\\_01.pdf](https://www.census.gov/content/dam/Census/library/working-papers/2023/acs/2023_Virgile_01.pdf)