



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

September 27, 2017

2017 AMERICAN COMMUNITY SURVEY RESEARCH AND EVALUATION REPORT
MEMORANDUM SERIES # ACS17-RER-07

MEMORANDUM FOR Victoria Velkoff
Chief, American Community Survey Office

From: David Waddington
Chief, Social, Economic, and Housing Statistics Division (SEHSD)

Prepared by: Edward Berchick
Social, Economic, and Housing Statistics Division (SEHSD)

Subject: 2016 American Community Survey Content Test Evaluation
Report: Health Insurance Coverage, Premiums and Subsidies

Attached is the final report for the 2016 American Community Survey (ACS) Content Test for Health Insurance. This report describes the results of the test for revised versions of the Health Insurance questions.

If you have any questions about this report, please contact Marina Vornovitsky at 301-763-7904 or Edward Berchick at 301-763-6351.

Attachment

cc:
Megan Rabe (ACSO)
Jennifer Ortman (ACSO)
Joanne Pascale (CSM)
Michael Ikeda (CSRM)
Patrick Cantwell (DSSD)
Sarah Heimel (DSSD)
Elizabeth Poehler (DSSD)
Anthony Tersine (DSSD)
Jennifer Day (SEHSD)
Nicole Scanniello (SEHSD)
Marina Vornovitsky (SEHSD)

Intentionally Blank

2016 American Community Survey Content Test Evaluation Report: Health Insurance Coverage, Premiums and Subsidies

FINAL REPORT



Edward Berchick and Brett O'Hara
Social, Economic, and Housing
Statistics Division

Sarah Heimel
Decennial Statistical Studies Division

R. Chase Sawyer
American Community Survey Office

Intentionally Blank

TABLE OF CONTENTS

EXECUTIVE SUMMARY	iv
1. BACKGROUND.....	1
1.1. Justification for Inclusion of Health Insurance Coverage, Premiums, and Subsidies in the Content Test	1
1.2. Question Development	2
1.3. Question Content.....	3
1.4. Research Questions	8
2. METHODOLOGY	10
2.1. Sample Design.....	10
2.2. Data Collection.....	11
2.3. Content Follow-Up.....	13
2.4. Analysis Metrics.....	13
2.4.1. Unit Response Rates and Demographic Profile of Responding Households ..	14
2.4.2. Item Missing Data Rates.....	15
2.4.3. Response Distributions	16
2.4.4. Benchmarks	16
2.4.5. Response Error	17
2.4.6. Standard Error Calculations.....	19
3. DECISION CRITERIA FOR HEALTH INSURANCE COVERAGE, PREMIUM, AND SUBSIDY	20
4. LIMITATIONS	21
5. RESEARCH QUESTIONS AND RESULTS	23
5.1. Unit Response Rates and Demographic Profile of Responding Households	23
5.1.1. Unit Response Rates for the Original Content Test Interview	23
5.1.2. Unit Response Rates for the Content Follow-Up Interview	25
5.1.3. Demographic and Socioeconomic Profile of Responding Households.....	26
5.2. Item Missing Data Rates	27
5.3. Response Distributions.....	32
5.4. Benchmarks	43
5.5. Response Error	46
6. CONCLUSION AND RECOMMENDATIONS	49
7. ACKNOWLEDGEMENTS	50
8. REFERENCES	50
Appendix A: Extra Tables	54
Appendix B. Dual Coverage Edit Checks.....	55

List of Tables

Table 1. Edit Checks Tested, by Mode	5
Table 2: Interview and Reinterview Counts For Each Response Category Used For Calculating the Gross Difference Rate and Index of Inconsistency	18
Table 3. Decision Criteria for Health Insurance Coverage.....	20
Table 4. Decision Criteria for the Health Insurance Premium and Subsidy Question	21
Table 5. Original Interview Unit Response Rates for Control and Test Treatments, Overall and by Mode	24
Table 6. Mail Response Rates by Designated High (HRA) and Low (LRA) Response Areas	25
Table 7. Content Follow-up Interview Unit Response Rates for Control and Test Treatments, Overall and by Mode of Original Interview	25
Table 8. Response Distributions: Test versus Control Treatment	26
Table 9. Comparison of Average Household Size.....	27
Table 10. Comparison of Language of Response	27
Table 11. Health Insurance Coverage Question Item Missing Data Rates for Control and Test Treatments, by Mode.....	28
Table 12. Implied “No” Rates for Test and Control Treatments, by Mode.....	29
Table 13. Medicaid and Direct-Purchase Item Missing Data Rates for Control and Test Treatments	30
Table 14. Health Insurance Premium and Subsidy Item Missing Data Rates for Control and Test Treatments	30
Table 15. Universe for Health Insurance Premium and Subsidy Question by Interview Mode, Control and Test Treatments	31
Table 16. Response Distribution for Any Insurance Coverage by Interview Mode, Control and Test Treatments	32
Table 17. Response Distribution for Specific Health Insurance Coverage Types, Control and Test Treatments	34
Table 18. Response Distribution for Employer-Based Health Insurance by Interview Mode, Control and Test Treatments	34
Table 19. Response Distribution for Direct-Purchase Health Insurance by Interview Mode, Control and Test Treatments	35
Table 20. Response Distribution for Medicare by Interview Mode, Control and Test Treatments	36
Table 21. Response Distribution for Direct-Purchase and Medicaid Coverage by Medicaid Expansion Status, Control and Test Treatments	37
Table 22. Response Distribution for People Who Reported Multiple Types of Health Insurance Coverage by Interview Mode, Control and Test Treatments.....	38
Table 23. Response Distribution for People Who Reported Both Employer-Based and Direct- Purchase Coverage by Interview Mode, Control and Test Treatments.....	38
Table 24. Response Distribution for Health Insurance Coverage Checkboxes (Excluding Write-in Responses), Control and Test Treatments	39
Table 25. Response Distribution for Health Insurance Coverage Type Write-in by Interview Mode, Control and Test Treatments.....	39
Table 26. Response Distribution for Medicare Coverage by Age, Control and Test Treatments	40

Table 27. Response Distribution Direct-Purchase Insurance by Age, Control and Test Treatments	41
Table 28. Response Distribution for Employer-Based Insurance by Age, Control and Test Treatments	41
Table 29. Response Distribution for VA and Military Coverage by Military Status, Control and Test Treatments	42
Table 30. Response Distribution for Subsidized Marketplace Coverage by Interview Mode, Control and Test Treatments	43
Table 31. Comparison of ACS Content Test Treatments to External Benchmarks, Health Insurance Coverage	44
Table 32. Difference in Gross Difference Rates (GDR) between Control and Test Treatments..	46
Table 33. Difference in Index of Inconsistency (IOI) between Control and Test Treatments	47
Table 34. Difference in Gross Difference Rate (GDR) for Multiple Insurance Coverage between Control and Test Treatments.....	47
Table 35. Difference in Index of Inconsistency (IOI) for Multiple Insurance Coverage between Control and Test Treatments.....	48
Table 36. Difference in Gross Difference Rate (GDR) between Control and Test Treatments ...	48
Table 37. Difference in Index of Inconsistency (IOI) between Control and Test Treatments	49
Table A-1. Unit Response Rates by Designated High (HRA) and Low (LRA) Response Areas	54
Table B-1. Counts of Dual-Coverage Edit Triggers and Final Responses, Test Treatment	55

List of Figures

Figure 1: Health Insurance Coverage Question, Control and Test Versions	4
Figure 2: Edit Check for Reports of Multiple Types of Coverage, Employer-based and Direct-purchase	5
Figure 3: “Same as Person 1” Option	6
Figure 4: Health Insurance Premium and Subsidy Question, Control and Test Versions	7

EXECUTIVE SUMMARY

Overview

From February to June of 2016, the U.S. Census Bureau conducted the 2016 American Community Survey (ACS) Content Test, a field test of new and revised content. The primary objective was to test whether changes to question wording, response categories, and definitions of underlying constructs improve the quality of data collected. Both new and revised versions of existing questions were tested to determine if they could provide data of sufficient quality compared to a control version as measured by a series of metrics including item missing data rates, response distributions, comparisons with benchmarks, and response error. The results of this test will be used to help determine the future ACS content and to assess the expected data quality of revised questions and new questions added to the ACS.

The 2016 ACS Content Test consisted of a nationally representative sample of 70,000 residential addresses in the United States, independent of the production ACS sample. The sample universe did not include group quarters, nor did it include housing units in Alaska, Hawaii, or Puerto Rico. The test was a split-panel experiment with one-half of the addresses assigned to the control treatment and the other half assigned to the test treatment. As in production ACS, the data collection consisted of three main data collection operations: 1) a six-week mailout period, during which the majority of self-response via internet and mailback were received; 2) a one-month Computer-Assisted Telephone Interview period for nonresponse follow-up; and 3) a one-month Computer-Assisted Personal Interview period for a sample of the remaining nonresponse. For housing units that completed the original Content Test interview, a Content Follow-Up telephone reinterview was conducted to measure response error.

Health Insurance Coverage, Premiums and Subsidies

This report discusses the Health Insurance Coverage question as well as a proposed new Health Insurance Premium and Subsidy question. A revised version of the Coverage question was tested due to the changes in health insurance coverage with the passage of the Patient Protection and Affordable Care Act in 2010 (ACA) (United States Congress, 2010) and to increase the accuracy of health insurance coverage estimates. The current Coverage question asks respondents to report whether they have coverage through one of seven types of health insurance (or to write in their coverage type if they have an “other” type of insurance). The proposed revision added additional instructions, changed the order of response options, and revised the wording for two coverage types (Medicaid and direct-purchase). The proposed Premium and Subsidy question, which is currently not included on the ACS, asked if a person pays a health insurance premium and, if so, whether he or she received a subsidy to help pay the premium. The proposed question would allow for measurement of subsidized Marketplace coverage, which was introduced through the ACA.

Research Questions and Results

Several research questions concerning item missing data rates, response distributions, and response error guided this evaluation.

Health Insurance Coverage Question

Item missing data rates: About 8.0 percent of the sample in the test treatment did not answer either “Yes” or “No” to at least one item in the Health Insurance Coverage question. This rate did not differ between treatments.

Response distribution: Most of our research questions and primary decision criteria concerned health insurance coverage rates. A higher overall insured rate was considered preferable due to an underreporting of Medicaid and other means-tested programs in the production ACS (Boudreaux, Ziegenfuss, Graven, Davern, & Blewett, 2011; Boudreaux, Call, Turner, & Fried, 2014; Boudreaux, Call, Turner, Fried, & O'Hara, 2015; Lynch, Kenney, Haley, & Resnick, 2011; O'Hara, 2010). The insured rate was significantly different between treatments: the insured rate was 89.1 percent with the test version of the question and 91.4 percent with the control version. Part of the difference between the control and test questions was likely influenced by differences across interview modes. For example, the proportion of Computer-Assisted Personal Interview (CAPI) respondents who reported employer-based health insurance differed, but no difference was observed for any other interview mode.

We also examined response distributions by coverage type. A higher Medicaid coverage rate was considered preferable to a lower one, as earlier research has documented an underreporting of Medicaid and other means-tested programs in the production ACS. However, Medicaid coverage rates were not significantly different between the test and control versions of the Health Insurance Coverage question, regardless of whether a respondent resided in a state that expanded Medicaid eligibility under the ACA.

Rates of direct-purchase and employer-based health insurance were both significantly lower in the test version of the Health Insurance Coverage question than in the control version (11.3 percent versus 13.0 percent for direct-purchase and 51.3 percent versus 55.0 percent for employer- or union-based coverage). However, as discussed in the report, there were substantial differences across interview modes, especially for employer/union coverage. Medicare coverage rates were significantly higher in the test version than in the control version of the question (16.7 percent versus 15.7 percent). Neither TRICARE nor Veterans Administration (VA) coverage rates differed between treatments.

Response reliability: Response reliability generally was not significantly different between treatments.

Health Insurance Premium and Subsidy Question

We also evaluated two versions of a new Health Insurance Premium and Subsidy question. The Premium and Subsidy question measured health insurance coverage via the subsidized Marketplace. We classified persons as having subsidized Marketplace coverage if they reported having direct-purchase insurance, paying a premium, and receiving a subsidy for the premium. Our assessment of the Premium and Subsidy

question was tied to our assessment of the Coverage question, as insurance coverage responses determined who was in universe to report a premium or subsidy.¹

Item missing data rates: For the premium part of the question (part a), the item missing rate was 1.0 percentage point lower in the test version than in the control version. For the subsidy part of the question (part b), the item-missing rate did not differ between the two versions of the question. However, the universe of respondents differed between the versions of the premium part of the question due to differences in the response distribution of the Health Insurance Coverage question. The proportion of individuals in universe was 2.9 percentage points lower in the test treatment than in the control treatment.

Response distribution: About 1.8 percent of individuals in the test version had subsidized Marketplace coverage. This rate did not differ from the rate in the control version.

Response reliability: Response reliability did not significantly differ between treatments for the premium part of the question (part a). However, for the subsidy part of the question (part b), the control version was more reliable than the test version.

Conclusions

We tested a new version of the Health Insurance Coverage question and two new versions of the Health Insurance Premium and Subsidy question. The test version of the Coverage question was introduced to reduce the Medicaid undercount documented in earlier research, to allow respondents to better distinguish direct-purchase insurance from other insurance types, and to improve the accuracy of estimates of health insurance coverage rates in the United States. Results from the Content Test suggest that the proposed revision to the Health Insurance Coverage question did not appreciably and uniformly improve the accuracy of estimates; the test version of the question offered estimates that were largely statistically indistinguishable from the control version of the question.

We also evaluated new questions on premiums and subsidies to determine if individuals purchased their insurance through the subsidized Marketplace. The two versions of the Premium and Subsidy question performed comparably based on our evaluation criteria. However, the Premium and Subsidy question in the control treatment was shorter, resulting in lower respondent burden.

¹ Both versions of the Premium and Subsidy question were new. The “control” version was the version paired with the control version of the Coverage question, and the “test” version was the version paired with the test version of the Coverage question.

1. BACKGROUND

From February to June of 2016, the Census Bureau conducted the 2016 American Community Survey (ACS) Content Test, a field test of new and revised content. The primary objective was to test whether changes to question wording, response categories, and definitions of underlying constructs improve the quality of data collected. Both revised versions of existing questions and new questions were tested to determine if they could provide data of sufficient quality compared to a control version as measured by a series of metrics including item missing data rates, response distributions, comparisons with benchmarks, and response error. The results of this test will be used to help determine the future ACS content and to assess the expected data quality of revised questions and new questions added to the ACS.

The 2016 ACS Content Test included the following topics:

- Relationship
- Race and Hispanic Origin
- Telephone Service
- Computer and Internet Use
- Health Insurance Coverage
- Health Insurance Premium and Subsidy (new questions)
- Journey to Work: Commute Mode
- Journey to Work: Time of Departure for Work
- Number of Weeks Worked
- Class of Worker
- Industry and Occupation
- Retirement, Survivor, and Disability Income

This report discusses the Health Insurance Coverage question, as well as the proposed new Health Insurance Premium and Subsidy question.

1.1. Justification for Inclusion of Health Insurance Coverage, Premiums, and Subsidies in the Content Test

The ACS began collecting information on health insurance coverage in 2008. The question was added “to enable the Department of Health and Human Services (HHS) and other federal agencies to more accurately distribute resources and better understand state and local health insurance needs” (U.S. Census Bureau, 2007).

Research has documented two limitations of the current measure. First, prior research has detailed an underreporting of Medicaid and other means-tested programs (Boudreaux, Ziegenfuss, Graven, Davern, & Blewett, 2011; Boudreaux M. , Call, Turner, Fried, & O'Hara, 2015; O'Hara, 2010) which, all else equal, would result in the underreporting of any health insurance coverage. There is also evidence of overreporting of direct-purchase health insurance coverage. Overreporting likely stems from the inclusion of non-comprehensive health plans (such as vision, dental, or single-service plans), which are not in scope in the ACS, and/or the

report of one plan multiple times (as more than one type of coverage) (Lynch, Kenney, Haley, & Resnick, 2011).

Moreover, with the passage of the Patient Protection and Affordable Care Act (ACA) (Public Law No. 111-148, U.S. Congress, 2010) there is a legislative need to differentiate respondents who have directly purchased coverage through the Health Insurance Marketplace from those who have Medicaid or other means-tested programs (Pascale, Rodean, Leeman, Cosenza, & Schoua-Glusberg, 2013).²

A new Premium and Subsidy question is being considered in response to other changes introduced by the ACA. The ACA introduced coverage purchased through the Health Insurance Marketplace and allowed some income-eligible Marketplace enrollees to receive subsidies for their coverage premiums. The new Health Insurance Premium question and the follow-up Subsidy question, along with the question about Health Insurance Coverage, were designed to allow the Census Bureau to produce estimates of subsidized Marketplace coverage.³

1.2. Question Development

Initial versions of the new and revised questions were proposed by federal agencies participating in the U.S. Office of Management and Budget (OMB) Interagency Committee for the ACS. The initial proposals contained a justification for each change and described previous testing of the question wording, the expected impact of revisions to the time series and the single-year as well as five-year estimates, and the estimated net impact on respondent burden for the proposed revision.⁴ For proposed new questions, the justification also described the need for the new data, whether federal law or regulation required the data for small areas or small population groups, if other data sources were currently available to provide the information (and why any alternate sources were insufficient), how policy needs or emerging data needs would be addressed through the new question, an explanation of why the data were needed with the geographic precision and frequency provided by the ACS, and whether other testing or production surveys had evaluated the use of the proposed questions.

The Census Bureau and the OMB, as well as the Interagency Council on Statistical Policy Subcommittee, reviewed these proposals for the ACS. The OMB determined which proposals moved forward into cognitive testing. After OMB approval of the proposals, topical subcommittees were formed from the OMB Interagency Committee for the ACS, which included all interested federal agencies that use the data from the impacted questions. These subcommittees further refined the specific proposed wording that was cognitively tested.

² Note that as of January 2017, Puerto Rico and the Island Areas do not have access to a Health Insurance Marketplace.

³ Although the questions were also introduced to help distinguish direct-purchase coverage from Medicaid coverage, the current structure of health insurance programs does not make it possible to distinguish between these two coverage types. Some persons pay out-of-pocket premiums for Medicaid and others have fully subsidized direct coverage that they might report as having no (out-of-pocket) premium.

⁴ The ACS produces both single and five-year estimates annually. Single year estimates are produced for geographies with populations of 65,000 or more and five-year estimates are produced for all areas down to the block-group level, with no population restriction.

The Census Bureau contracted with Westat to conduct three rounds of cognitive testing. The results of the first two rounds of cognitive testing informed decisions on specific revisions to the proposed content for the stateside Content Test (Stapleton and Steiger, 2015). In the first round, 208 cognitive interviews were conducted in English and Spanish and in two modes (self-administered on paper and interviewer-administered on paper). In the second round of testing, 120 cognitive interviews were conducted for one version of each of the tested questions, in English and Spanish, using the same modes as in the first round.

A third round of cognitive testing involved only the Puerto Rico Community Survey (PRCS) and Group Quarters (GQ) versions of the questionnaire (Steiger, Anderson, Folz, Leonard, & Stapleton, 2015). Cognitive interviews in Puerto Rico were conducted in Spanish; GQ cognitive interviews were conducted in English. The third round of cognitive testing was carried out to assess the revised versions of the questions in Spanish and identify any issues with questionnaire wording unique to Puerto Rico and GQ populations.⁵ The proposed changes identified through cognitive testing for each question topic were reviewed by the Census Bureau, the corresponding topical subcommittee, and the Interagency Council on Statistical Policy Subcommittee for the ACS. The OMB then provided final overall approval of the proposed wording for field testing.⁶

1.3. Question Content

Health Insurance Coverage

Figure 1 shows the two versions as they appeared on the paper questionnaire mailed to respondents. The current production ACS question is the control version in the Content Test, and the revised version is the test version. All interview modes had the same text content but formatted accordingly for each mode.

⁵ Note that the field testing of the content was not conducted in Puerto Rico or in GQs. See the Methodology section for more information.

⁶ A cohabitation question and domestic partnership question were included in cognitive testing but ultimately we decided not to move forward with field testing these questions.

Figure 1: Health Insurance Coverage Question, Control and Test Versions

Control:	Test:																																																						
<p>16 Is this person CURRENTLY covered by any of the following types of health insurance or health coverage plans? Mark "Yes" or "No" for EACH type of coverage in items a – h.</p> <table border="0" style="width: 100%;"> <tr> <td style="width: 80%;"></td> <td style="text-align: center;">Yes</td> <td style="text-align: center;">No</td> </tr> <tr> <td>a. Insurance through a current or former employer or union (of this person or another family member)</td> <td style="text-align: center;"><input type="checkbox"/></td> <td style="text-align: center;"><input type="checkbox"/></td> </tr> <tr> <td>b. Insurance purchased directly from an insurance company (by this person or another family member)</td> <td style="text-align: center;"><input type="checkbox"/></td> <td style="text-align: center;"><input type="checkbox"/></td> </tr> <tr> <td>c. Medicare, for people 65 and older, or people with certain disabilities</td> <td style="text-align: center;"><input type="checkbox"/></td> <td style="text-align: center;"><input type="checkbox"/></td> </tr> <tr> <td>d. Medicaid, Medical Assistance, or any kind of government-assistance plan for those with low incomes or a disability</td> <td style="text-align: center;"><input type="checkbox"/></td> <td style="text-align: center;"><input type="checkbox"/></td> </tr> <tr> <td>e. TRICARE or other military health care</td> <td style="text-align: center;"><input type="checkbox"/></td> <td style="text-align: center;"><input type="checkbox"/></td> </tr> <tr> <td>f. VA (including those who have ever used or enrolled for VA health care)</td> <td style="text-align: center;"><input type="checkbox"/></td> <td style="text-align: center;"><input type="checkbox"/></td> </tr> <tr> <td>g. Indian Health Service</td> <td style="text-align: center;"><input type="checkbox"/></td> <td style="text-align: center;"><input type="checkbox"/></td> </tr> <tr> <td>h. Any other type of health insurance or health coverage plan – <i>Specify</i> ↘</td> <td style="text-align: center;"><input type="checkbox"/></td> <td style="text-align: center;"><input type="checkbox"/></td> </tr> </table>		Yes	No	a. Insurance through a current or former employer or union (of this person or another family member)	<input type="checkbox"/>	<input type="checkbox"/>	b. Insurance purchased directly from an insurance company (by this person or another family member)	<input type="checkbox"/>	<input type="checkbox"/>	c. Medicare, for people 65 and older, or people with certain disabilities	<input type="checkbox"/>	<input type="checkbox"/>	d. Medicaid, Medical Assistance, or any kind of government-assistance plan for those with low incomes or a disability	<input type="checkbox"/>	<input type="checkbox"/>	e. TRICARE or other military health care	<input type="checkbox"/>	<input type="checkbox"/>	f. VA (including those who have ever used or enrolled for VA health care)	<input type="checkbox"/>	<input type="checkbox"/>	g. Indian Health Service	<input type="checkbox"/>	<input type="checkbox"/>	h. Any other type of health insurance or health coverage plan – <i>Specify</i> ↘	<input type="checkbox"/>	<input type="checkbox"/>	<p>15 Is this person CURRENTLY covered by any of the following types of health insurance or health coverage plans? Mark "Yes" or "No" for EACH type of coverage in items a – h.</p> <p><i>Do NOT include plans that cover only one type of service, such as dental, drug or vision plans.</i></p> <table border="0" style="width: 100%;"> <tr> <td style="width: 80%;"></td> <td style="text-align: center;">Yes</td> <td style="text-align: center;">No</td> </tr> <tr> <td>a. Insurance through a current or former employer or union (of this person or another family member)</td> <td style="text-align: center;"><input type="checkbox"/></td> <td style="text-align: center;"><input type="checkbox"/></td> </tr> <tr> <td>b. Medicare, for people 65 and older, or people with certain disabilities</td> <td style="text-align: center;"><input type="checkbox"/></td> <td style="text-align: center;"><input type="checkbox"/></td> </tr> <tr> <td>c. Medicaid, Medical Assistance, or any kind of state- or government-assistance plan for those with low income</td> <td style="text-align: center;"><input type="checkbox"/></td> <td style="text-align: center;"><input type="checkbox"/></td> </tr> <tr> <td>d. Insurance purchased directly from an insurance company or through a State or Federal Marketplace, HealthCare.gov, or a similar state website (by this person or another family member)</td> <td style="text-align: center;"><input type="checkbox"/></td> <td style="text-align: center;"><input type="checkbox"/></td> </tr> <tr> <td>e. TRICARE or other military health care</td> <td style="text-align: center;"><input type="checkbox"/></td> <td style="text-align: center;"><input type="checkbox"/></td> </tr> <tr> <td>f. VA (including those who have ever used or enrolled for VA health care)</td> <td style="text-align: center;"><input type="checkbox"/></td> <td style="text-align: center;"><input type="checkbox"/></td> </tr> <tr> <td>g. Indian Health Service</td> <td style="text-align: center;"><input type="checkbox"/></td> <td style="text-align: center;"><input type="checkbox"/></td> </tr> <tr> <td>h. Other type of health coverage NOT listed above – <i>Specify</i> ↘</td> <td style="text-align: center;"><input type="checkbox"/></td> <td style="text-align: center;"><input type="checkbox"/></td> </tr> </table>		Yes	No	a. Insurance through a current or former employer or union (of this person or another family member)	<input type="checkbox"/>	<input type="checkbox"/>	b. Medicare, for people 65 and older, or people with certain disabilities	<input type="checkbox"/>	<input type="checkbox"/>	c. Medicaid, Medical Assistance, or any kind of state- or government-assistance plan for those with low income	<input type="checkbox"/>	<input type="checkbox"/>	d. Insurance purchased directly from an insurance company or through a State or Federal Marketplace, HealthCare.gov, or a similar state website (by this person or another family member)	<input type="checkbox"/>	<input type="checkbox"/>	e. TRICARE or other military health care	<input type="checkbox"/>	<input type="checkbox"/>	f. VA (including those who have ever used or enrolled for VA health care)	<input type="checkbox"/>	<input type="checkbox"/>	g. Indian Health Service	<input type="checkbox"/>	<input type="checkbox"/>	h. Other type of health coverage NOT listed above – <i>Specify</i> ↘	<input type="checkbox"/>	<input type="checkbox"/>
	Yes	No																																																					
a. Insurance through a current or former employer or union (of this person or another family member)	<input type="checkbox"/>	<input type="checkbox"/>																																																					
b. Insurance purchased directly from an insurance company (by this person or another family member)	<input type="checkbox"/>	<input type="checkbox"/>																																																					
c. Medicare, for people 65 and older, or people with certain disabilities	<input type="checkbox"/>	<input type="checkbox"/>																																																					
d. Medicaid, Medical Assistance, or any kind of government-assistance plan for those with low incomes or a disability	<input type="checkbox"/>	<input type="checkbox"/>																																																					
e. TRICARE or other military health care	<input type="checkbox"/>	<input type="checkbox"/>																																																					
f. VA (including those who have ever used or enrolled for VA health care)	<input type="checkbox"/>	<input type="checkbox"/>																																																					
g. Indian Health Service	<input type="checkbox"/>	<input type="checkbox"/>																																																					
h. Any other type of health insurance or health coverage plan – <i>Specify</i> ↘	<input type="checkbox"/>	<input type="checkbox"/>																																																					
	Yes	No																																																					
a. Insurance through a current or former employer or union (of this person or another family member)	<input type="checkbox"/>	<input type="checkbox"/>																																																					
b. Medicare, for people 65 and older, or people with certain disabilities	<input type="checkbox"/>	<input type="checkbox"/>																																																					
c. Medicaid, Medical Assistance, or any kind of state- or government-assistance plan for those with low income	<input type="checkbox"/>	<input type="checkbox"/>																																																					
d. Insurance purchased directly from an insurance company or through a State or Federal Marketplace, HealthCare.gov, or a similar state website (by this person or another family member)	<input type="checkbox"/>	<input type="checkbox"/>																																																					
e. TRICARE or other military health care	<input type="checkbox"/>	<input type="checkbox"/>																																																					
f. VA (including those who have ever used or enrolled for VA health care)	<input type="checkbox"/>	<input type="checkbox"/>																																																					
g. Indian Health Service	<input type="checkbox"/>	<input type="checkbox"/>																																																					
h. Other type of health coverage NOT listed above – <i>Specify</i> ↘	<input type="checkbox"/>	<input type="checkbox"/>																																																					

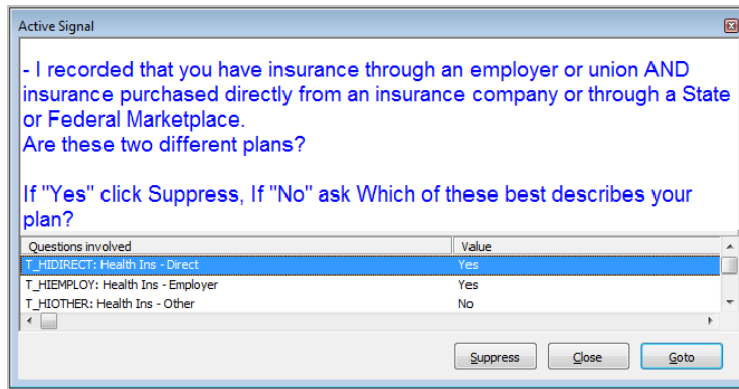
The test version of the Health Insurance Coverage question included the following changes:

- 1) Reordered health insurance types. The order of the health insurance types in the control (production) ACS question was (a) Employer-provided, (b) direct-purchase, (c) Medicare, (d) Medicaid, (e) TRICARE, (f) VA (Veterans Administration), (g) Indian Health Service, (h) other (write-in). The order in the test version of the question was (a) Employer-provided, (b) Medicare, (c) Medicaid, (d) direct-purchase, (e) TRICARE, (f) VA, (g) Indian Health Service, (h) other (write-in). To attempt to reduce over-reporting of direct-purchase insurance (Mach & O’Hara, 2011), the direct-purchase option was moved from (b) to (d), and Medicare and Medicaid were shifted up one position each.
- 2) Added instructions to the Health Insurance question. The test version added an instruction to help focus respondents' attention on comprehensive coverage and, therefore, reduce overreporting of direct-purchase coverage. The instruction stated: “Do NOT include plans that cover only one type of service, such as dental, drug or vision plans.”
- 3) Updated Medicaid question. The test version description of Medicaid was identical to the control version, except that it no longer included references to disability. With the expansion of Medicaid under the ACA, individuals do not need to have a disability to qualify for Medicaid in many states. While the control question asked about: "Medicaid, Medical Assistance, or any kind of state or government-assistance plan for those with low income or a disability," the test question removed the phrase "or a disability."
- 4) Added key terms to the direct-purchase question to measure coverage through the Health Insurance Marketplace. Since 2014, people have been able to buy health insurance through the Health Insurance Marketplace, usually through healthcare.gov or a state-specific website. Although this coverage is direct-purchase, individuals may not know how to report it. Based on recommendations from cognitive testing (Stapleton & Steiger, January 2015), the test

question read (new text italicized): "Insurance purchased directly from an insurance company *or through a State or Federal Marketplace, HealthCare.gov, or a similar state website* (by this person or another family member)." This change was intended to reduce reporting of Marketplace plans as Medicaid or as both Medicaid and direct-purchase.

- 5) Additional edit checks. The current production Computer-Assisted Telephone Interview (CATI) and Computer-Assisted Personal Interview (CAPI) versions of the Health Insurance Coverage question includes an edit check that comes up when someone selects employer-provided insurance (item a) in combination with direct-purchase (item b). With this edit check, a pop-up window appears in the survey instrument to clarify that the respondent is reporting two different plans (Figure 2).

Figure 2: Edit Check for Reports of Multiple Types of Coverage, Employer-based and Direct-purchase



The Content Test introduced two additional edit checks, as well as integration of the edit checks into the internet mode. The new edit checks included: other (item h) in combination with one other type of coverage, and Medicaid (item c) in combination with direct-purchase (item d). The first of these edit checks was in both the control and test versions for CATI/CAPI; the second two were new to CATI/CAPI, and all three edit checks were new to the internet mode. Table 1 below describes the edit checks by mode.

Table 1. Edit Checks Tested, by Mode

Edit Check	Control	Test
Employer-based & Direct-purchase	CATI, CAPI	Internet, CATI, CAPI
Medicaid & Direct-purchase	None	Internet, CATI, CAPI
Checkbox (items a-g) & "Other" (item h)	None	Internet, CATI, CAPI

CATI = Computer-assisted telephone interview; CAPI = Computer-assisted personal interview

The first new edit check (other (h) in combination with one other type of coverage) was hypothesized to reduce reporting the same plan by more than one description. This change could reduce over-reports of direct-purchase coverage, since many write-ins are coded as direct-purchase because they lack any additional information besides the name of an insurance

company (e.g., Blue Cross). The edit check was introduced to help reduce this error by asking the respondent to verify that the two plans were actually separate plans and not just two ways used to describe the same plan.

The second new edit check (Medicaid (c) in combination with direct-purchase (d)) was proposed to classify responses correctly when respondents use both Medicaid and direct-purchase to describe their coverage through a subsidized exchange plan (the healthcare.gov website serves as a portal for both Marketplace and Medicaid registration). Depending on the income reported when applying for direct-purchase Marketplace coverage, lower income individuals may be redirected (on the website) to enroll in Medicaid. Additionally, people who want to sign up for Medicaid might be directed to sign up through the Marketplace in the first place. This new check could help respondents confirm they reported the correct coverage type.

6) Same coverage as Person 1. The test version of the CATI/CAPI instrument allowed interviewers to select an option if respondents volunteered that other persons in the household have the same type of health insurance coverage as the first person. On the test version’s health insurance introduction screen (shown on the top of Figure 3), the interviewer could record that the respondent volunteered that the second (or higher) person had the same type of coverage as the first person in the household. The types of coverage Person 1 did not have were pre-filled with “No”’s, and the interviewer read through the parts of the question corresponding to types that Person 1 did report (as shown on the bottom part of Figure 3). If the respondent changed an answer, all health insurance coverage types were reviewed. Importantly, the interviewer did *not* read anything new to the respondent.

Figure 3: “Same as Person 1” Option

I am now going to ask you some questions about Person Two's health insurance and health coverage. Do NOT include plans that cover only one type of service, such as dental, drug or vision plans.

1. Reset this person
 2. Same type as Person 1

Health Ins - Introduction	2	Health Ins - Premium
Health Ins - Employer		Health Ins - Subsidy
Health Ins - Medicare	2	
Health Ins - Medicaid	2	
Health Ins - Direct	2	
Health Ins - Military	2	
Health Ins - Veteran's Administration	2	
Health Ins - Indian	2	
Health Ins - Other	2	
Other health ins		
Health Ins - Other		

Note: In an interview, “Person One” and “Person Two” was replaced with actual names.

Together, these six modifications to the production ACS question were intended to reduce the Medicaid undercount documented in earlier research, to allow respondents to better distinguish direct-purchase insurance from other insurance types, and to improve the accuracy of estimates of health insurance coverage rates in the United States. All else equal, these changes were not expected to affect coverage rates for employer-based, Medicare, TRICARE, or VA coverage.

Health Insurance Premium and Subsidy

The Content Test also evaluated a new two-part question: one that asked if the health insurance plan had a premium associated with it and, if so, if that premium was subsidized. The new Premium and Subsidy question, in combination with the Health Insurance question, would allow the Census Bureau to produce estimates of coverage obtained through subsidized Marketplace plans.⁷

The question on premiums and subsidies does not currently appear on the production ACS; therefore, we tested two different versions. Initial development and testing of question wording were done earlier this decade with Massachusetts residents for use on both the Current Population Survey Annual Social and Economic Supplement (CPS ASEC) and ACS survey instruments (Pascale, Rodean, Leeman, Cosenza, & Schoua-Glusberg, 2013). Version 1 (labeled the “control” version) is a shorter version than Version 2 (labeled the “test” version).⁸ Figure 4 shows both versions as they appeared on the paper questionnaire mailed to respondents. Automated versions of the questionnaire had the same content formatted accordingly for each mode.

Figure 4: Health Insurance Premium and Subsidy Question, Control and Test Versions

Version 1 (Control):	Version 2 (Test):
<p>17 a. Is there a premium for this plan? <i>A premium is a fixed amount of money paid on a regular basis for health coverage. It does not include copays, deductibles, or other expenses such as prescription costs.</i></p> <p><input type="checkbox"/> Yes</p> <p><input type="checkbox"/> No → SKIP to question 18a</p> <p>b. Does this person or another family member receive a tax credit or subsidy based on family income to help pay the premium?</p> <p><input type="checkbox"/> Yes</p> <p><input type="checkbox"/> No</p>	<p>16 a. Does this person or another family member pay a premium for this health insurance plan? <i>A premium is a fixed amount of money paid on a regular basis for health coverage. It does not include copays, deductibles, or other expenses such as prescription costs.</i></p> <p><input type="checkbox"/> Yes</p> <p><input type="checkbox"/> No → SKIP to question 17a</p> <p>b. Based on family income, does this person or another family member receive financial assistance through a subsidy or tax credit to help pay part or all of the cost of the premium for this plan?</p> <p><input type="checkbox"/> Yes</p> <p><input type="checkbox"/> No</p>

There is an important limitation to this analysis: there were two different versions of the Health Insurance question, each coupled with a different version of the Premium and Subsidy question, posing challenges to the evaluation. The control version of the Health Insurance Premium and Subsidy question was only tested with the control version of the Health Insurance Coverage question, and the test version of the Premium and Subsidy question was only tested with the test

⁷ The questions were also introduced to help distinguish direct-purchase coverage from Medicaid coverage, but the current health insurance landscape does not allow for this distinction. Some persons pay out-of-pocket premiums for Medicaid, and others have fully subsidized direct coverage that they might report as having no premium paid out-of-pocket.

⁸ As noted earlier, the “control” version of the Premium and Subsidy question was the version paired with the control version of the Coverage question, and the “test” version was the version paired with the test version of the Coverage question.

version of the Coverage question. The universe of persons asked the Premium question could have been different between control and test treatments if people answered the Health Insurance Coverage question differently. By design, only people with Medicaid, direct-purchase or an “other” type of health insurance coverage were in universe for the Premium and Subsidy question.⁹ Differences in how people responded to the Medicaid, direct-purchase, and “other” checkboxes, therefore, could have led to differences in the universe for the new question.

Additionally, we do not know the extent to which the test Health Insurance Coverage question could prime respondents to answer the Premium and Subsidy question differently. The test version of the Health Insurance Coverage question includes the term “Health Insurance Marketplace.” This additional text could have primed respondents to think about those plans that were subsidized through the Marketplace in a way that respondents in the control condition did not.

1.4. Research Questions

The following research questions were formulated to guide the analyses of the Health Insurance Coverage and Premium and Subsidy questions.¹⁰ The analyses assess how the test versions of the questions performed compared to the control versions in the following ways: how often the respondents answered the questions, the consistency and accuracy of the responses, and how the responses affect the resulting estimates.

Health Insurance Coverage Question

- 1. Is the overall item missing data rate lower for the test treatment than for the control treatment?*

⁹ To reduce respondent confusion, there was no explicit skip instruction on the mailed questionnaire.

¹⁰ Questions have been reordered and renumbered from the Research & Evaluation Analysis Plan (REAP). A number of research questions have also been omitted. The Health Insurance Coverage and Premium and Subsidy questions can only measure subsidized Marketplace coverage, as they cannot distinguish unsubsidized nongroup plans purchased from the Marketplace from those purchased elsewhere. As a result, the following two research questions were omitted:

Using the Health Insurance Coverage question, Premium question and Subsidy question in combination, what are the rates of Marketplace coverage in each version?

Using the Health Insurance Coverage question, Premium question and Subsidy question in combination, what are the rates of unsubsidized Marketplace coverage in each version?

Since some persons pay out-of-pocket premiums for Medicaid, and others have fully subsidized direct coverage that they might report as having no premium paid out-of-pocket, the following research question was not addressed:

Which version [of the Health Insurance Coverage question] needs to be more heavily recoded using the Premium and Subsidy question?

Finally, based on the recommendation not to adopt the test version of the question, we did not assess the research question examining edit checks in the test version, but included frequencies in this report’s appendix. As noted in the REAP, this question was asked for informational purposes only and was not part of the specified decision criteria.

2. *Does the rate of implied “No” responses differ between the test and control versions?*¹¹
3. *Is the item missing data rate for the Medicaid and direct-purchase boxes, and their combination, lower in the test treatment than for the control treatment?*
4. *Does the test version of the Health Insurance question have a higher proportion of persons with any type of health insurance coverage compared with the control version?*¹²
5. *For each of the above questions, is there a difference when dividing responses into self-administered (mail and internet) versus interviewer-administered (CATI and CAPI)?*¹³
6. *Does the test version of the Health Insurance question have a higher proportion of persons with Medicaid compared with the control version?*
7. *Are rates of coverage by employer-based insurance, direct-purchase insurance, TRICARE, VA Care, and Medicare consistent between the test and control versions?*
8. *Does the test version of the Health Insurance Coverage question have a lower proportion of persons with multiple types of health insurance compared with the control version?*
9. *Does the test version of the Health Insurance Coverage question result in a smaller proportion of persons who reported having both employer-provided insurance and insurance purchased directly compared with the control version?*
10. *Does the test version of the Health Insurance Coverage question decrease the proportion of persons who write-in an “other” type of health insurance coverage compared with the control version?*
11. *Is there a difference in coverage types when divided by age or military status?*
12. *Is there a difference in coverage types, especially means-tested and direct-purchase, between test and control responses, by state Medicaid Expansion status?*
13. *How do the proportions in each treatment compare with proportions found in other surveys?*
14. *Are the measures of response reliability (gross difference rate, index of inconsistency) better for the test treatment than for the control treatment?*

¹¹ Implied “No” responses are those person records with at least one “Yes” or “No” box marked and at least one blank checkbox.

¹² Based on findings, we also evaluated a more general version of this question: Does the proportion of persons with any type of health insurance coverage differ across the test and control versions? This revision implies a two-tailed t-test, whereas the original question wording implies a one-tailed test.

¹³ Based on findings, we also examined whether differences by mode also existed for the overall insured rate, specific coverage type rates, and Premium and Subsidy question response distribution. We also compared the overall insured rate between treatments excluding CAPI respondents.

15. In CATI/CAPI interviews, how often was the option for respondents to volunteer the same type of coverage as Person 1 (for Persons 2+) used in the test version?

Health Insurance Premium and Subsidy Question

16. For each part of the question (Premium and Subsidy), which version has a lower missing data rate?
17. Do the versions have a different proportion of individuals who are in the universe for the Premium question?
18. Do the versions have different proportions of individuals who are in the universe for the Subsidy question? (Coming from part a.)
19. For each part of the question (Premium and Subsidy), is the percent of people who say “Yes” to each question (for those in the universe for each question) different between the two versions?
20. Using the Health Insurance Coverage question and Premium and Subsidy question in combination, what are the rates of subsidized Marketplace coverage in each version?
21. Which version produces estimates of subsidized Marketplace coverage that more closely match benchmarks (including administrative data)?¹⁴
22. Are estimates of the proportions of persons who pay a premium comparable to estimates from National Health Interview Survey (NHIS)?
23. For each question (Premium and Subsidy), which version (control/test) has better response reliability?

2. METHODOLOGY

2.1. Sample Design

The 2016 ACS Content Test consisted of a nationally representative sample of 70,000 residential addresses in the United States, independent of the production ACS sample. The Content Test sample universe did not include GQs, nor did it include housing units in Alaska, Hawaii, or Puerto Rico.¹⁵ The sample design for the Content Test was largely based on the ACS production

¹⁴ This revised question is a combination of two different questions specified in the REAP: *Which version produces estimates of Marketplace coverage (subsidized and unsubsidized) that more closely match benchmarks?* and *Additionally, if timely data are available, we will compare results to benchmarks from administrative data. This could include Medicaid participation as well as participation in the Health Insurance Marketplace (only in states with Federal-based Marketplaces).* Given our ability to only measure subsidized Marketplaces and given the benchmarks available, these two questions became duplicative.

¹⁵ Alaska and Hawaii were excluded for cost reasons. GQs and Puerto Rico were excluded because the sample sizes required to produce reliable estimates would be overly large and burdensome, as well as costly.

sample design with some modifications to better meet the test objectives.¹⁶ The modifications included adding an additional level of stratification by stratifying addresses into high and low self-response areas, oversampling addresses from low self-response areas to ensure equal response from both strata, and sampling units as pairs.¹⁷ The high and low self-response strata were defined based on ACS self-response rates at the tract level. Sampled pairs were formed by first systematically sampling an address within the defined sampling stratum and then pairing that address with the address listed next in the geographically sorted list. Note that the pair was likely not neighboring addresses. One member of the pair was randomly assigned to receive the control version of the question and the other member was assigned to receive the test version of the question, thus resulting in a sample of 35,000 control cases and 35,000 test cases.

As in the production ACS, if efforts to obtain a response by mail or telephone were unsuccessful, attempts were made to interview in person a sample of the remaining nonresponding addresses (see Section 2.2 Data Collection for more details). Addresses were sampled at a rate of 1-in-3, with some exceptions that were sampled at a higher rate.¹⁸ For the Content Test, the development of workload estimates for CATI and CAPI did not take into account the oversampling of low response areas. This oversampling resulted in a higher than expected workload for CATI and CAPI and therefore required more budget than was allocated. To address this issue, the CAPI sampling rate for the Content Test was adjusted to meet the budget constraint.

2.2. Data Collection

The field test occurred in parallel with the data collection activities for the March 2016 ACS production panel, using the same basic data collection protocol as production ACS with a few differences as noted below. The data collection protocol consisted of three main data collection operations: 1) a six-week mailout period, during which the majority of internet and mailback responses were received; 2) a one-month CATI period for nonresponse follow-up; and 3) a one-month CAPI period for a sample of the remaining nonresponse. Internet and mailback responses were accepted until three days after the end of the CAPI month.

As indicated earlier, housing units included in the Content Test sample were randomly assigned to a control or test version of the questions. CATI interviewers were not assigned specific cases; rather, they worked the next available case to be called and therefore conducted interviews for both control and test cases. CAPI interviewers were assigned Content Test cases based on their geographic proximity to the cases and therefore could also conduct both control and test cases.

¹⁶ The ACS production sample design is described in Chapter 4 of the ACS Design and Methodology report (U.S. Census Bureau, 2014).

¹⁷ Tracts with the highest response rate based on data from the 2013 and 2014 ACS were assigned to the high response stratum in such a way that 75 percent of the housing units in the population (based on 2010 Census estimates) were in the high response areas; all other tracts were designated in the low response strata. Self-response rates were used as a proxy for overall cooperation. Oversampling in low response areas helps to mitigate larger variances due to CAPI subsampling. This stratification at the tract level was successfully used in previous ACS Content Tests, as well as the ACS Voluntary Test in 2003.

¹⁸ The ACS production sample design for CAPI follow-up is described in Chapter 4, Section 4.4 of the ACS Design and Methodology report (U.S. Census Bureau, 2014).

The ACS Content Test’s data collection protocol differed from the production ACS in a few significant ways. The Content Test analysis did not include data collected via the Telephone Questionnaire Assistance (TQA) program since those who responded via TQA used the ACS production TQA instrument. The Content Test excluded the telephone Failed Edit Follow-Up (FEFU) operation.¹⁹ Furthermore, the Content Test had an additional telephone reinterview operation used to measure response reliability. We refer to this telephone reinterview component as the Content Follow-Up, or CFU. The CFU is described in more detail in Section 2.3.

ACS production provides Spanish-language versions of the internet, CATI, and CAPI instruments, and callers to the TQA number can request to respond in Spanish, Russian, Vietnamese, Korean, or Chinese. The Content Test had Spanish-language automated instruments; however, there were no paper versions of the Content Test questionnaires in Spanish.²⁰ Any case in the Content Test sample that completed a Spanish-language internet, CATI, or CAPI response was included in analysis. However, if a case sampled for the Content Test called TQA to complete an interview in Spanish or any other language, the production interview was conducted and the response was excluded from the Content Test analysis. This was due to the low volume of non-English language cases and the operational complexity of translating and implementing several language instruments for the Content Test. CFU interviews for the Content Test were conducted in either Spanish or English. The practical need to limit the language response options for Content Test respondents is a limitation to the research, as some respondents self-selected out of the test.

Analyses also used coded data to determine individuals’ health insurance coverage. In both versions of the Coverage question, respondents could write-in an “other” type of coverage into a response field (see Figure 1). Responses were machine-coded: if the write-in text matched the text of previously-used write-in entry, then the text was assigned one of the listed coverage types. If the text did not match a previous-used write-in entry, it was sent to a Census Bureau analyst to manually code to an insurance type. As the Content Test data did not go through full production processing, entries coded to “public or private” were used to calculate rates of insurance but were ignored when calculating rates of specific coverage types.²¹

¹⁹ In ACS production, paper questionnaires with an indication that there are more than five people in the household or questions about the number of people in the household, and self-response returns that are identified as being vacant or a business or lacking minimal data are included in FEFU. FEFU interviewers call these households to obtain any information the respondent did not provide.

²⁰ In the 2014 ACS, respondents requested 1,238 Spanish paper questionnaires, of which 769 were mailed back. From that information, we projected that fewer than 25 Spanish questionnaires would be requested in the Content Test.

²¹ Consistent with production ACS data, write-ins were coded to one of twelve categories: employer-based, direct-purchase, Medicare, Medicaid, TRICARE, VA, Indian Health Service, family/dependent coverage, private or public, out-of-scope, other, or no coverage. In production, the private or public category would be allocated to employer-based, direct-purchase, Medicare, or Medicaid based on edit specifications. However, Content Test data did not undergo full processing.

2.3. Content Follow-Up

For housing units that completed the original interview, a CFU telephone reinterview was also conducted to measure response error.²² A comparison of the original interview responses and the CFU reinterview responses was used to answer research questions about response error and response reliability.

A CFU reinterview was attempted with every household that completed an original interview for which there was a telephone number. A reinterview was conducted no sooner than two weeks (14 calendar days) after the original interview. Once the case was sent to CFU, it was to be completed within three weeks. This timing balanced two competing interests: (1) conducting the reinterview as soon as possible after the original interview to minimize changes in truth between the two interviews, and (2) not making the two interviews so close together that the respondents were simply recalling their previous answers. Interviewers made two call attempts to interview the household member who originally responded, but if that was not possible, the CFU reinterview was conducted with any other eligible household member (15 years or older).

The CFU asked basic demographic questions and a subset of housing and detailed person questions that included all of the topics being tested, with the exception of Telephone Service, and any questions necessary for context and interview flow to set up the questions being tested.²³ All CFU questions were asked in the reinterview, regardless of whether or not a particular question was answered in the original interview. Because the CFU interview was conducted via telephone, the wording of the questions in CFU followed the same format as the CATI nonresponse interviews. Housing units assigned to the control version of the questions in the original interview were asked the control version of the questions in CFU; housing units assigned to the test version of the questions in the original interview were asked the test version of the questions in CFU. The only exception was for retirement, survivor, and disability income, for which a different set of questions was asked in CFU.²⁴

2.4. Analysis Metrics

This section describes the metrics used to assess the revised version of the Health Insurance Coverage question and the proposed Health Insurance Premium and Subsidy question. These metrics include the item missing data rate, response distributions, comparisons to benchmarks, and response error. This section also describes the methodology used to calculate unit response rates and standard errors for the test.

All Content Test data were analyzed without imputation due to our interest in how question changes or differences between versions of new questions affected “raw” responses, not the final edited variables. Some editing of responses was done for analysis purposes, such as collapsing

²² Throughout this report, the “original interview” refers to responses completed via paper questionnaire, internet, CATI, or CAPI.

²³ Because the CFU interview was conducted via telephone, the Telephone Service question was not asked. We assume that CFU respondents have telephone service.

²⁴ Refer to the 2016 ACS Content Test report on Retirement Income for a discussion on CFU questions for survivor, disability, and retirement income.

response categories or modes together or calculating a person's age based on his or her date of birth.

All estimates from the ACS Content Test were weighted. Analysis involving data from the original interviews used the final weights that take into account the initial probability of selection (the base weight) and CAPI subsampling. For analysis involving data from the CFU interviews, the final weights were adjusted for CFU nonresponse to create CFU final weights.

The significance level for all hypothesis tests is $\alpha = 0.1$. Since we are conducting numerous comparisons between the control and test treatments, there is a concern about incorrectly rejecting a hypothesis that is actually true (a "false positive" or Type I error). The overall Type I error rate is called the familywise error rate and is the probability of making one or more Type I errors among all hypotheses tested simultaneously. When adjusting for multiple comparisons, the Holm-Bonferroni method was used (Holm, 1979).

2.4.1. Unit Response Rates and Demographic Profile of Responding Households

The unit response rate is generally defined as the proportion of sample addresses eligible to respond that provided a complete or sufficient partial response.²⁵ Unit response rates from the original interview are an important measure to look at when considering the analyses in this report that compare responses between the control and test versions of the survey questionnaire. High unit response rates are important in mitigating potential nonresponse bias.

For both control and test treatments, we calculated the overall unit response rate (all modes of data collection combined) and unit response rates by mode: internet, mail, CATI, and CAPI. We also calculated the total self-response rate by combining internet and mail modes together. Some Content Test analyses focused on the different data collection modes for topic-specific evaluations, thus we felt it was important to include each mode in the response rates section. In addition to those rates, we calculated the response rates for high and low response areas because analysis for some Content Test topics was done by high and low response areas. Using the Census Bureau's Planning Database (U.S. Census Bureau, 2016), we defined these areas at the tract level based on the low response score.

The universe for the overall unit response rates consists of all addresses in the initial sample (70,000 addresses) that were eligible to respond to the survey. Some examples of addresses ineligible for the survey were a demolished home, a home under construction, a house or trailer that was relocated, or an address determined to be a permanent business or storage facility. The universe for self-response (internet and mail) rates consists of all mailable addresses that were eligible to respond to the survey. The universe for the CATI response rate consists of all nonrespondents at the end of the mailout month from the initial survey sample that were eligible to respond to the survey and for whom we possessed a telephone number. The universe for the CAPI response rates consists of a subsample of all remaining nonrespondents (after CATI) from the initial sample that were eligible to respond to the survey. Any nonresponding addresses that were sampled out of CAPI were not included in any of the response rate calculations.

²⁵ A response is deemed a "sufficient partial" when the respondent gets to the first question in the detailed person questions section for the first person in the household.

We also calculated the CFU interview unit response rate overall and by mode of data collection of the original interview and compared the control and test treatments because response error analysis (discussed in Section 5.5.) relies upon CFU interview data. Statistical differences between CFU response rates for control and test treatments will not be taken as evidence that one version is better than the other. For the CFU response rates, the universe for each mode consists of housing units that responded to the original questionnaire in the given mode (internet, mail, CATI, or CAPI) and were eligible for the CFU interview. We expected the response rates to be similar between treatments; however, we calculated the rates to verify that assumption.

Another important measure to look at in comparing experimental treatments is the demographic profile of the responding households in each treatment. The Content Test sample was designed with the intention of having respondents in both control and test treatments exhibit similar distributions of socioeconomic and demographic characteristics. Similar distributions allow us to compare the treatments and conclude that any differences are due to the experimental treatment instead of underlying demographic differences. Thus, we analyzed distributions for data from the following response categories: *age*, *sex*, *educational attainment*, and *tenure*. The topics of *race*, *Hispanic origin*, and *relationship* are also typically used for demographic analysis; however, those questions were modified as part of the Content Test, so we could not include them in the demographic profile. Additionally, we calculated *average household size* and the *language of response* for the original interview.²⁶

For response distributions we used chi-square tests of independence to determine statistical differences between control and test treatments. If the distributions were significantly different, we performed additional testing on the differences for each response category. To control for the overall Type I error rate for a set of hypotheses tested simultaneously, we performed multiple-comparison procedures with the Holm-Bonferroni method (Holm, 1979). A family for our response distribution analysis was the set of p-values for the overall characteristic categories (*age*, *sex*, *educational attainment*, and *tenure*) and the set of p-values for a characteristic's response categories if the response distributions were found to have statistically significant differences. To determine statistical differences for *average household size* and the *language of response* of the original interview we performed two-tailed hypothesis tests.

For all response-related calculations mentioned in this section, addresses that were either sampled out of the CAPI data collection operation or that were deemed ineligible for the survey were not included in any of the universes for calculations. Unmailable addresses were also excluded from the self-response universe. For all unit response rate estimates, differences, and demographic response analysis, we used replicate base weights adjusted for CAPI sampling (but not adjusted for CFU nonresponse).

2.4.2. Item Missing Data Rates

Respondents leave items blank for a variety of reasons including not understanding the question (clarity), their unwillingness to answer a question as presented (sensitivity), and their lack of

²⁶ Language of response analysis excludes paper questionnaire returns because there was only an English questionnaire.

knowledge of the data needed to answer the question. The item missing data rate (for a given item) is the proportion of eligible units, housing units for household-level items or persons for person-level items, for which a required response (based on skip patterns) is missing.

There are three classes of responses that are of interest when analyzing the item missing data rate; our evaluation primarily focuses on the second type but also considers the third.

1. Person records with no response to any part of the question (completely blank), “Don’t know” and “Refused” were treated as missing
2. Person records with one or more boxes marked “Yes” or “No” and the rest of the boxes left blank (implied “No”)
3. Person records with either “Yes” or “No” marked for each coverage type, as desired (full response)

We expected the item missing data rate to be lower in the test treatment than in control treatment due to the introduction of additional instructions at the beginning of the Health Insurance Coverage question and the addition of key terms to the direct-purchase question.

2.4.3. Response Distributions

Comparing the response distributions between the control version of a question and the test version of a question allowed us to assess whether the question change affected the resulting estimates. Comparisons were made using Rao-Scott chi-squared tests (Rao & Scott, 1987) for distribution and t-tests for single categories when the corresponding distributions were found to be statistically different.

Proportion estimates were calculated as:

$$\text{Category proportion} = \frac{\text{weighted count of valid responses in category}}{\text{weighted count of all valid responses}}$$

2.4.4. Benchmarks

For the topic of health insurance coverage, we compared data from both control and test treatments to information from the 2015 National Health Interview Survey (NHIS) and the 2016 Current Population Survey Annual Social and Economic Supplement (CPS ASEC). The NHIS is an annual cross-sectional survey of American households that collects information about household members’ health, health insurance, and sociodemographic characteristics (Parsons et al., 2014). The CPS ASEC is an annual household survey that obtains detailed information about the social and demographic characteristics of household members, including their education, income, and health status (U.S. Census Bureau, 2006). Comparing ACS results with NHIS and CPS results allowed us to assess whether our results on coverage rates are markedly different from other national-level, governmental sources.

For the Health Insurance Premium and Subsidy Question, we compared data from both control and test treatments to information from the NHIS and from the Centers for Medicare and Medicaid Services (CMS). Data from these sources allowed us to compare the proportion of insured respondents who pay a health insurance premium and the subsidized Marketplace

coverage rates with information from other government sources. We did not compare Content Test and external data with formal statistical tests; rather, we used them to ensure that estimates were reasonable.

Benchmark limitations

Estimates from other surveys may differ for a variety of reasons. First, health insurance questions differ across surveys in ways that could contribute to differences in coverage rates. For example, ACS, NHIS, and some CPS ASEC estimates reflect health insurance coverage at the time of the interview, while other CPS ASEC benchmarks correspond to coverage for the previous calendar year. Second, the mode of interview may lead to different results, and the ACS is the only survey among the three that includes a mail mode and internet mode (NHIS and CPS are both collected via interview). Third, the context of the survey may differ across the ACS, CPS ASEC, and NHIS in ways that could prime respondents when they answer survey questions: the NHIS primarily focuses on health, while the CPS ASEC primarily focuses on labor force and income. Finally, Content Test data have not been through complete data processing, and both Content Test and NHIS data contain missing values.

2.4.5. Response Error

Response error occurs for a variety of reasons, such as flaws in the survey design, misunderstanding of the questions, misreporting by respondents, or interviewer effects. There are two components of response error: response bias and simple response variance. Response bias is the degree to which respondents consistently answer a question incorrectly. Simple response variance is the degree to which respondents answer a question inconsistently. A question has good response reliability if respondents tend to answer the question consistently. Re-asking the same question of the same respondent (or housing unit) allows us to measure response variance.

We measured simple response variance by comparing valid responses to the CFU reinterview with valid responses to the corresponding original interview.²⁷ The Census Bureau has frequently used content reinterview surveys to measure simple response variance for large demographic data collection efforts, including the 2010 ACS Content Test, and the 1990, 2000, and 2010 decennial censuses (Dusch & Meier, 2012).

The following measures were used to evaluate consistency:

- Gross difference rate (GDR)
- Index of inconsistency (IOI)

The two measures – GDR and IOI – were calculated for individual response categories. The GDR, and subsequently the simple response variance, are calculated using the following table and formula.

²⁷ A majority of the CFU interviews were conducted with the same respondent as the original interview (see the Limitations section for more information).

Table 2: Interview and Reinterview Counts for Each Response Category Used for Calculating the Gross Difference Rate and Index of Inconsistency

	Original Interview “Yes”	Original Interview “No”	Reinterview Totals
CFU Reinterview “Yes”	a	b	a + b
CFU Reinterview “No”	c	d	c + d
Original Interview Totals	a + c	b + d	n

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

a = weighted count of units in the category of interest for both the original interview and reinterview

b = weighted count of units NOT in the category of interest for the original interview, but in the category for the reinterview

c = weighted count of units in the category of interest for the original interview, but NOT in the category for the reinterview

d = weighted count of units NOT in the category of interest for either the original interview or the reinterview

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

The GDR for a specific response category is the percent of inconsistent answers between the original interview and the reinterview (CFU). We calculate the GDR for a response category as:

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

Statistical significance between the GDR for a specific response category between the control and test treatments is determined using a two-tailed t-test.

In order 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 (Flanagan, 1996).

SV is the part of total variance resulting from the differences among 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 the total variance is due to inconsistency in responses, as measured by SRV and is calculated as:

$$\text{IOI} = \frac{n(b + c)}{(a + c)(c + d) + (a + b)(b + d)} \times 100$$

Per the Census Bureau’s general rule, index values of less than 20 percent indicate low inconsistency, 20 to 50 percent indicate moderate inconsistency, and over 50 percent indicate high inconsistency.

When the sample size is small, the reliability estimates are unstable. Therefore, we do not report the IOI and GDR values for categories with a small sample size, as determined by the following formulas: $2a + b + c < 40$ or $2d + b + c < 40$, where a, b, c, and d are unweighted counts as shown in Table 1 above (see Flanagan 1996, p. 15).

The measures of response error assume that those characteristics in question did not change between the original interview and the CFU interview. To the extent that this assumption is incorrect, we assume that it is incorrect at similar rates between the control and test treatments. For instance, an individual might have changed coverage types before the CFU interview and then accurately reported a different response than in the original interview.

In calculating the IOI reliability measures, the assumption is that the expected value of the error in the original interview is the same as in the CFU reinterview. This assumption of parallel measures is necessary for the SRV and IOI to be valid. In calculating the IOI measures for this report, we found this assumption was not met for the response categories specified in the limitations section (see Section 4).

Biemer (2011, pp. 56-58) provides an example where the assumption of parallel measures is not met, but does not provide definitive guidelines for addressing it. In Biemer’s concluding remarks, he states, “...both estimates of reliability are biased to some extent because of the failure of the parallel assumptions to hold.” Flanagan (2001) addresses this bias problem and offers the following adjustment to the IOI formula:

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

This formula was tested on selected topics in the 2016 ACS Content Test. The IOI_{estimate} resulted in negligible reduction in the IOI values. For this reason, we did not recalculate the IOI values using IOI_{estimate} . Similar to Biemer (2011, p. 58), we acknowledge that for some cases, the estimate of reliability is biased to some extent.

2.4.6. Standard Error Calculations

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, 2014, Chapter 12). We calculated the variance for each rate and difference using the formula below. The standard error of the estimate (X_0) is the square root of the variance:

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

where:

X_0 = the estimate calculated using the full sample,

X_r = the estimate calculated for replicate r .

3. DECISION CRITERIA FOR HEALTH INSURANCE COVERAGE, PREMIUM, AND SUBSIDY

Before fielding the 2016 ACS Content Test, we identified which of the metrics would be given higher importance in determining which version of the question would be recommended for inclusion in the ACS moving forward. The following tables identify the research questions and associated metrics in priority order.

Table 3. Decision Criteria for Health Insurance Coverage

Research Questions	Decision Criteria, in order of priority
4	The version that had a lower uninsured rate is preferable (because of evidence that people underreport Medicaid, Medical Assistance, and other means-tested state-provided health plans).
6	Given the evidence that people underreport Medicaid, Medical Assistance, and other means-tested state-provided health plans, the version that had a higher Medicaid rate is preferable.
8-9, 15	Given that people over-report direct-purchase plans that are not comprehensive or that write-ins could be coded as direct-purchase incorrectly, a decrease in direct-purchase plans in combination with another type of coverage is preferable.
7, 11-13	We use current healthcare policy to guide our analysis wherever possible. This potentially includes information about eligibility for receiving subsidies or tax credits for premiums paid to buy insurance through an exchange. ²⁸
10	We expect that write-ins are being used for one of two reasons: first, that respondents did not understand the question or their type of insurance, so they write in their coverage type for the Census Bureau to figure out, or second, that respondents are providing additional information about a plan they already reported. Given this information, a decrease in the proportion of people who write in a type of health insurance coverage is preferable.
1-3	A lower item missing data rate for the health insurance coverage variables is preferable.
14	Higher response reliability (based on gross difference rate and index of inconsistency) is preferable.

Note: Research Question 5, which concerns differences by interview mode for Research Questions 1-4, spans multiple criteria and, therefore, is not included in Table 3. We consider mode differences for all criteria that concern these questions.

²⁸ This criterion has been updated to reflect sample size limitations and narrower research questions.

Table 4. Decision Criteria for the Health Insurance Premium and Subsidy Question

Research Questions	Decision Criteria, in order of priority
20-22, 19	Together with and apart from the Health Insurance question, the version that produced an estimate of subsidized Marketplace coverage ²⁹ comparable to benchmarks is preferable. If both versions performed the same, the shorter version (on the control instrument) is preferable.
17-18	We expect the number of persons in universe for each version to be comparable. Likewise, we expect the rates of subsidized coverage to be comparable; a difference would be noteworthy.
16	The version with a lower item missing data rate is preferable.
23	The version with lower measures of response error (gross difference rate, index of inconsistency) is preferable.

4. LIMITATIONS

CATI and CAPI interviewers were assigned control and test treatment cases, as well as production cases. The potential risk of this approach is the introduction of a cross-contamination or carry-over effect due to the same interviewer administering multiple versions of the same question item. Interviewers are trained to read the questions verbatim to minimize this risk, but there still exists the possibility that an interviewer may deviate from the scripted wording of one question version to another. This could potentially mask a treatment effect from the data collected.

Interviews were only conducted in English and Spanish. Respondents who needed language assistance in another language were not able to participate in the test. Additionally, the 2016 ACS Content Test was not conducted in Alaska, Hawaii, or Puerto Rico. Any conclusions drawn from this test may not apply to these areas or populations.

For statistical analysis specific to the mail mode, there may be bias in the results because of unexplained unit response rate differences between the control and test treatments.

We were not able to conduct demographic analysis by relationship status, race, or ethnicity because these topics were tested as part of the Content Test.

The CFU reinterview was not conducted in the same mode of data collection for households that responded by internet, by mail, or by CAPI in the original interview since CFU interviews were only administered using a CATI mode of data collection. As a result, the data quality measures derived from the reinterview may include some bias due to the differences in mode of data collection.

²⁹ In the REAP, we stated that we would make our decision based on subsidized and unsubsidized Marketplace coverage. However, the Premium and Subsidy question does not permit an estimate of unsubsidized Marketplace coverage.

To be eligible for a CFU reinterview, respondents needed to either provide a telephone number in the original interview or have a telephone number available to the Census Bureau through reverse address look up. As a result, 2,284 of the responding households (11.8 percent with a standard error of 0.2) from the original control interviews and 2,402 of the responding households (12.4 percent with a standard error of 0.2) from the original test interviews were not eligible for the CFU reinterview. The difference between the control and test treatments is statistically significant (p-value=0.06).

Although we reinterviewed the same person who responded in the original interview when possible, we interviewed a different member of the household in the CFU for 7.5 percent (standard error of 0.4) of the CFU cases for the control treatment and 8.4 percent (standard error of 0.5) of the CFU cases for the test treatment.³⁰ The difference between the test and control treatments is not statistically significant (p-value=0.26). This means that differences in results between the original interview and the CFU for these cases could be due in part to having different people answering the questions. However, those changes were not statistically significant between the control and test treatments and should not impact the conclusions drawn from the reinterview.

The Content Test does not include the production weighting adjustments for seasonal variations in ACS response patterns, nonresponse bias, and under-coverage bias. As a result, any estimates derived from the Content Test data do not provide the same level of inference as the production ACS and cannot be compared to production estimates.

In developing initial workload estimates for CATI and CAPI, we did not take into account the fact that we oversampled low response areas as part of the Content Test sample design. Therefore, workload and budget estimates were too low. In order to stay within budget, the CAPI workload was subsampled more than originally planned. This caused an increase in the variances for the analysis metrics used.

An error in addressing and assembling the materials for the 2016 ACS Content Test caused some Content Test cases to be mailed production ACS questionnaires instead of Content Test questionnaires. There were 49 of these cases that returned completed questionnaires, and they were all from the test treatment. These cases were excluded from the analysis. Given the small number of cases affected by this error, there is very little effect on the results.

Questionnaire returns were expected to be processed and keyed within two weeks of receipt. Unfortunately, a check-in and keying backlog prevented this requirement from being met, thereby delaying eligible cases from being sent to CFU on a schedule similar to the other modes. Additionally, the control treatment questionnaires were processed more quickly in keying than the test treatment questionnaires resulting in a longer delay for test mail questionnaire cases to be eligible for CFU. On average, it took 18 days for control cases to become eligible for CFU; it took 20 days for test cases. The difference is statistically significant. This has the potential to impact the response reliability results.

³⁰ This is based on comparing the first name of the respondent between the original interview and the CFU interview. Due to a data issue, we were not able to use the full name to compare.

The assumption of parallel measures for GDR and IOI calculations was not met for the following categories: employer-based, Medicare, TRICARE, VA, and “other” health insurance coverage, as well as for the combination of direct-purchase insurance and Medicare, direct-purchase and employer-based, and Medicare and Medicaid. The assumption was also not met for health insurance premiums. For these categories, the GDR and IOI estimates are biased to some extent.

Content Test data did not undergo processing prior to analysis. As a result of item nonresponse and the need to process other data, we were not able to stratify analyses by income or poverty status. Research questions, therefore, were narrowed to consider a smaller subset of social and demographic characteristics. We were also unable to stratify results by relationship status, race, and/or ethnicity due to differences in those questions between the test and control treatments.

Finally, the universe of people who were eligible to answer the Premium and Subsidy question differed between control and test respondents. Only people who reported coverage via Medicaid, direct-purchase, or an “other” type of coverage to the Health Insurance Coverage question are asked if they paid a premium.³¹ Each version of the Premium and Subsidy question followed a different version of the Health Insurance Coverage question. Second, we do not know if the test Health Insurance Coverage question primed respondents to answer the Premium and Subsidy question differently. The test Health Insurance Coverage question includes the term “Health Insurance Marketplace,” which should prime respondents to think about those plans that are subsidized through the Marketplace.

5. RESEARCH QUESTIONS AND RESULTS

This section presents the results from the analyses of the 2016 ACS Content Test data for the Health Insurance Coverage question and the Health Insurance Premium and Subsidy question. An analysis of unit response rates is presented first followed by topic-specific analyses. For the topic-specific analyses, each research question is restated, followed by corresponding data and a brief summary of the results.

5.1. Unit Response Rates and Demographic Profile of Responding Households

This section provides results for unit response rates for both control and test treatments for the original Content Test interview and for the CFU interview. It also provides results of a comparison of socioeconomic and demographic characteristics of respondents in both control and test treatments.

5.1.1. Unit Response Rates for the Original Content Test Interview

The unit response rate is generally defined as the proportion of sample addresses eligible to respond that provided a complete or sufficient partial response. We did not expect the unit response rates to differ between treatments. This is important because the number of unit responses should also affect the number of item responses we receive for analyses done on

³¹ The skip pattern was built into the instrument for the CATI/CAPI and Internet modes. To reduce respondent confusion, there is no explicit skip instruction for mail responses. However, we restricted mail responses to the same universe in our analyses for comparability across modes.

specific questions on the survey. Similar item response universe sizes allow us to compare the treatments and conclude that any differences are due to the experimental treatment instead of differences in the populations sampled for each treatment.

Table 5 shows the unit response rates for the original interview for each mode of data collection (internet, mail, CATI, and CAPI), all modes combined, and both self-response modes (internet and mail combined) for the control and test treatments. When looking at the overall unit response rate (all modes combined), the difference between control (93.5 percent) and test (93.5 percent) is less than 0.1 percentage points and is not statistically significant.

Table 5. Original Interview Unit Response Rates for Control and Test Treatments, Overall and by Mode

Mode	Test Interviews	Test Percent	Control Interviews	Control Percent	Test minus Control	P-Value
All Modes	19,400	93.5 (0.3)	19,455	93.5 (0.3)	<0.1 (0.4)	0.98
Self-Response	13,131	52.9 (0.5)	13,284	53.7 (0.5)	-0.8 (0.6)	0.23
Internet	8,168	34.4 (0.4)	8,112	34.1 (0.4)	0.4 (0.6)	0.49
Mail	4,963	18.4 (0.3)	5,172	19.6 (0.3)	-1.2 (0.5)	0.01*
CATI	872	8.7 (0.4)	880	9.2 (0.4)	-0.4 (0.6)	0.44
CAPI	5,397	83.5 (0.7)	5,291	83.6 (0.6)	<0.1 (0.9)	0.96

Source: U.S. Census Bureau, 2016 American Community Survey Content Test

Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test at the $\alpha=0.1$ level. The weighted response rates account for initial sample design as well as CAPI subsampling.

When analyzing the unit response rates by mode of data collection, the only modal comparison that shows a statistically significant difference is the mail response rate. The control treatment had a higher mail response (19.6 percent) than the test treatment (18.4 percent) by 1.2 percentage points. As a result of this difference, we looked at how mail responses differed in the high and low response areas. Table 6 shows the mail response rates for both treatments in high and low response areas.³² The difference in mail response rates appears to be driven by the difference of rates in the high response areas.

It is possible that the difference in the mail response rates between control and test is related to the content changes made to the test questions. There are some test questions that could be perceived as being too sensitive by some respondents (such as the test question relating to same-sex relationships) and some test questions that could be perceived to be too burdensome by some respondents (such as the new race questions with added race categories). In the automated modes (internet, CATI, and CAPI) there is a higher likelihood of obtaining a sufficient partial response (obtaining enough information to be deemed a response for calculations before the respondent stops answering questions) than in the mail mode. If a respondent is offended by the questionnaire or feels that the questions are too burdensome, they may just throw the questionnaire away and not respond by mail. This could be a possible explanation for the unit response rate being lower for test than control in the mail mode.

³² Table A-1 (including all modes) can be found in Appendix A.

We note that differences between overall and total self-response response rates were not statistically significant. As most analysis was conducted at this level, we are confident the response rates were sufficient to conduct topic-specific comparisons between the control and test treatments and that there are no underlying response rate concerns that would impact those findings.

Table 6. Mail Response Rates by Designated High (HRA) and Low (LRA) Response Areas

	Test Interviews	Test Percent	Control Interviews	Control Percent	Test minus Control	P-Value
HRA	2,082	20.0 (0.4)	2,224	21.5 (0.4)	-1.5 (0.6)	0.02*
LRA	2,881	13.8 (0.3)	2,948	14.1 (0.3)	-0.3 (0.4)	0.43
Difference		6.2 (0.5)		7.4 (0.4)	-1.1 (0.7)	0.11

Source: U.S. Census Bureau, 2016 American Community Survey Content Test

Note: Standard errors are in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test at the $\alpha=0.1$ level. The weighted response rates account for initial sample design as well as CAPI subsampling.

5.1.2. Unit Response Rates for the Content Follow-Up Interview

Table 7 shows the unit response rates for the CFU interview by mode of data collection of the original interview and for all modes combined, for control and test treatments. Overall, the differences in CFU response rates between the treatments are not statistically significant. The rate at which CAPI respondents from the original interview responded to the CFU interview is lower for test (34.8 percent) than for control (37.7 percent) by 2.9 percentage points. While the protocols for conducting CAPI and CFU were the same between the test and control treatments, we could not account for personal interactions that occur in these modes between the respondent and interviewer. This can influence response rates. We do not believe that the difference suggests any underlying CFU response issues that would negatively affect topic-specific response reliability analysis for comparing the two treatments.

Table 7. Content Follow-Up Interview Unit Response Rates for Control and Test Treatments, Overall and by Mode of Original Interview

Original Interview Mode	Test Interviews	Test Percent	Control Interviews	Control Percent	Test minus Control	P-Value
All Modes	7,867	44.8 (0.5)	7,903	45.7 (0.6)	-0.8 (0.8)	0.30
Internet	4,078	51.9 (0.6)	4,045	52.5 (0.7)	-0.6 (0.8)	0.49
Mail	2,202	46.4 (0.9)	2,197	44.2 (0.9)	2.1 (1.3)	0.11
CATI	369	48.9 (1.9)	399	51.5 (2.5)	-2.5 (2.9)	0.39
CAPI	1,218	34.8 (1.2)	1,262	37.7 (1.1)	-2.9 (1.6)	0.07*

Source: U.S. Census Bureau, 2016 American Community Survey Content Test

Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test at the $\alpha=0.1$ level.

5.1.3. Demographic and Socioeconomic Profile of Responding Households

One of the underlying assumptions of our analyses in this report is that the sample for the Content Test was selected in such a way that responses from both treatments would be comparable. We did not expect the demographics of the responding households for control and test treatments to differ. To test this assumption, we calculated distributions for respondent data for the following response categories: *age*, *sex*, *educational attainment*, and *tenure*.³³ The response distribution calculations can be found in Table 8. Items with missing data were not included in the calculations. After adjusting for multiple comparisons, none of the differences in the categorical response distributions shown below are statistically significant.

Table 8. Response Distributions: Test versus Control Treatment

Item	Test Percent	Control Percent	Adjusted P-Value
AGE	(n=43,236)	(n=43,325)	0.34
Under 5 years old	5.7 (0.2)	6.1 (0.2)	
5 to 17 years old	17.8 (0.3)	17.6 (0.3)	
18 to 24 years old	8.6 (0.3)	8.1 (0.3)	
25 to 44 years old	25.1 (0.3)	26.2 (0.3)	
45 to 64 years old	26.8 (0.4)	26.6 (0.4)	
65 years old or older	16.0 (0.3)	15.4 (0.3)	
SEX	(n=43,374)	(n=43,456)	1.00
Male	48.8 (0.3)	49.1 (0.3)	
Female	51.2 (0.3)	50.9 (0.3)	
EDUCATIONAL ATTAINMENT [#]	(n=27,482)	(n=27,801)	1.00
No schooling completed	1.3 (0.1)	1.2 (0.1)	
Nursery to 11 th grade	8.1 (0.3)	8.0 (0.3)	
12 th grade (no diploma)	1.7 (0.1)	1.6 (0.1)	
High school diploma	21.7 (0.4)	22.3 (0.4)	
GED [†] or alternative credential	3.5 (0.2)	3.6 (0.2)	
Some college	21.0 (0.4)	20.2 (0.4)	
Associate's degree	8.8 (0.3)	9.1 (0.3)	
Bachelor's degree	20.9 (0.4)	20.3 (0.4)	
Advanced degree	13.1 (0.3)	13.7 (0.3)	
TENURE	(n=17,190)	(n=17,236)	1.00
Owned with a mortgage	43.1 (0.6)	43.2 (0.5)	
Owned free and clear	21.1 (0.4)	21.2 (0.4)	
Rented	33.8 (0.6)	34.0 (0.5)	
Occupied without payment of rent	1.9 (0.2)	1.7 (0.1)	

Source: U.S. Census Bureau, 2016 American Community Survey Content Test

[#]For ages 25 and older

[†]General Educational Development

Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. Significance testing done at the $\alpha=0.1$ level. P-values have been adjusted for multiple comparisons using the Holm-Bonferroni method.

³³ We were not able to conduct demographic analysis by relationship status, race, or ethnicity because these topics were tested as part of the Content Test.

We also analyzed two other demographic characteristics shown by the responses from the survey: *average household size* and *language of response*. The results for the remaining demographic analyses can be found in Table 9 and Table 10.

Table 9. Comparison of Average Household Size

Topic	Test (n=17,608)	Control (n=17,694)	Test minus Control	P-value
Average Household Size (Number of People)	2.51 (<0.1)	2.52 (<0.1)	>-0.01 (<0.1)	0.76

Source: U.S. Census Bureau, 2016 American Community Survey Content Test

Note: Standard errors are shown in parentheses. Significance was tested based on a two-tailed t-test at the $\alpha=0.1$ level.

Table 10. Comparison of Language of Response

Language of Response	Test Percent (n=17,608)	Control Percent (n=17,694)	Test minus Control	P-value
English	96.1 (0.2)	96.2 (0.2)	<0.1 (0.3)	0.52
Spanish	2.7 (0.2)	2.6 (0.2)	<0.1 (0.2)	0.39
Undetermined	1.2 (0.1)	1.2 (0.1)	<0.1 (0.2)	0.62

Source: U.S. Census Bureau, 2016 American Community Survey Content Test

Note: Standard errors are shown in parentheses. Significance was tested based on a two-tailed t-test at the $\alpha=0.1$ level.

The Content Test was available in two languages, English and Spanish, for all modes except the mail mode. However, the language of response variable was missing for some responses, so we created a category called “undetermined” to account for those cases.

There are no detectable differences between control and test for *average household size* or *language of response*. There are also no detectable differences for any of the response distributions that we calculated. As a result of these analyses, it appears that respondents in both treatments do exhibit comparable demographic characteristics since none of the resulting findings is significant, which verifies our assumption of demographic similarity between treatments.

5.2. Item Missing Data Rates

Is the overall item missing data rate lower for the test treatment than for the control treatment?

A number of research questions concern the missing data rate. We first examined the proportion of eligible people who do not respond “Yes” or “No” to any part of the Health Insurance Coverage question. Table 11 displays item missing data rates for the Health Insurance Coverage question. Mode-specific rates are also presented in this table and in subsequent tables to highlight any interview-mode differences (where present). As noted in the Decision Criteria, a lower missing rate was preferable to a higher one.

Table 11. Health Insurance Coverage Question Item Missing Data Rates for Control and Test Treatments, by Mode

Mode	Test Sample Size	Test Percent	Control Sample Size	Control Percent	Test minus Control	Adjusted P-value
Self-response	31,228	9.4 (0.3)	31,484	8.9 (0.3)	0.6 (0.4)	0.61
Internet	21,102	10.5 (0.4)	20,861	10.3 (0.4)	0.2 (0.6)	1.00
Mail	10,126	6.8 (0.4)	10,623	5.8 (0.4)	1.0 (0.5)	0.36
Computer-assisted	12,365	6.0 (0.6)	12,187	4.8 (0.5)	1.2 (0.7)	0.49
CATI	1,963	5.8 (1.2)	1,894	5.2 (1.3)	0.6 (1.9)	1.00
CAPI	10,402	6.0 (0.6)	10,293	4.7 (0.5)	1.2 (0.8)	0.49
Overall	43,593	8.0 (0.3)	43,671	7.2 (0.3)	0.8 (0.4)	0.36

Source: U.S. Census Bureau, 2016 American Community Survey Content Test

Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. Significance was tested based on a one-tailed t-test (test < control) at the $\alpha=0.1$ level. P-values have been adjusted for multiple comparisons using the Holm-Bonferroni method.

About 7.2 percent of persons in the control treatment were missing data for all items on the Health Insurance Coverage question. Contrary to hypotheses, the item missing data rate did not significantly differ between the test and control versions of the question. Differences were not observed for any of the four modes.

Unless otherwise specified, the remainder of analyses for the Health Insurance Coverage question are limited to person records with one or more boxes marked “Yes” or “No.”

Does the rate of implied “No” responses differ between the test and control versions?

For subsequent analyses (unless stated otherwise), we consider a “Yes” response to at least one coverage type and nonresponse to all other coverage type checkboxes to imply that the person does not have the nonchecked coverage types (implied “No”).³⁴ For example, if a person reported Medicare coverage, but the remaining coverage boxes were blank, we recoded the missing coverage-type responses to “No”s. Table 12 shows implied “No” rates across the test and control treatments.

³⁴ In the case where all other health insurance questions are blank, the write-in field must be codeable in order to treat missing values as implied “No” responses.

Table 12. Implied “No” Rates for Test and Control Treatments, by Mode

Mode	Test Sample Size	Test Percent	Control Sample Size	Control Percent	Test minus Control	Adjusted P-value
Self-response	27,811	51.1 (0.6)	28,130	53.2 (0.6)	-2.0 (0.9)	0.02*
Internet	18,547	48.9 (0.8)	18,314	51.5 (0.8)	-2.5 (1.1)	0.02*
Mail	9,264	56.1 (1.0)	9,816	56.7 (0.9)	-0.6 (1.4)	0.34
Computer-assisted	11,684	1.6 (0.3)	11,593	1.5 (0.2)	0.1 (0.3)	0.60
CATI	1,815	3.9 (0.5)	1,799	3.9 (0.6)	>-0.1 (0.8)	0.96
CAPI	9,869	1.4 (0.3)	9,794	1.3 (0.2)	0.1 (0.3)	0.96
Overall	39,495	30.8 (0.5)	39,723	32.2 (0.5)	-1.4 (0.7)	0.09*

Source: U.S. Census Bureau, 2016 American Community Survey Content Test

Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference based on a one-tailed t-test (test < control) at the $\alpha=0.1$ level. P-values have been adjusted for multiple comparisons using the Holm-Bonferroni method.

Consistent with hypotheses, the implied “No” rate was significantly lower in the test version of the question than in the control version: about 30.8 percent of persons had at least one implied “No” in the test treatment compared with 32.2 percent in the control treatment. These differences, however, were limited to internet responses, where the rate was 2.5 percentage points lower in the test version (48.9 percent in the test treatment versus 51.5 percent in the control version). Rates did not differ between treatments for mailed questionnaires or for the two computer-assisted interview modes.³⁵

Is the item missing data rate for the Medicaid and direct-purchase boxes, and their combination, lower in the test treatment than for the control treatment?

Given the motivation to improve the accuracy of reports of Medicaid and direct coverage, we also examined item missing data rates for Medicaid, direct-purchase, and their combination (Table 13).³⁶ For this research question, we examined whether a respondent marked a checkbox, treating an implied “No” as missing data. As seen in Table 13, the item missing data rates for these types of health insurance coverage were not lower in the test version than in the control version.

³⁵ CATI and CAPI implied “No” rates were expected to be low based on instrument design.

³⁶ Research has found that Medicaid and other means-tested programs are underreported in the ACS (Boudreaux, Ziegenfuss, Graven, Davern, & Blewett, 2011; O'Hara, 2010). All things being equal, a Medicaid undercount would result in the underreporting of health insurance coverage. Research has also found that direct-purchase health insurance coverage is overreported in the ACS (Lynch, Kenney, Haley, & Resnick, 2011).

Table 13. Medicaid and Direct-Purchase Item Missing Data Rates for Control and Test Treatments

Coverage Type	Test Percent (n=39,495)	Control Percent (n=39,723)	Test minus Control	Adjusted P-value
Medicaid	25.4 (0.5)	26.5 (0.4)	-1.2 (0.6)	0.11
Direct-purchase	24.9 (0.5)	25.7 (0.4)	-0.7 (0.7)	0.35
Medicaid & Direct-purchase	21.9 (0.5)	22.4 (0.4)	-0.5 (0.6)	0.35

Source: U.S. Census Bureau, 2016 American Community Survey Content Test

Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding.

Significance was tested based on a one-tailed t-test (test < control) at the $\alpha=0.1$ level. P-values have been adjusted for multiple comparisons using the Holm-Bonferroni method.

Health Insurance Premium and Subsidy Question

For each part of the question (Premium and Subsidy), which version has a lower missing data rate?

We also analyzed item missing data rates for the Health Insurance Premium and Subsidy question (Table 14). For the premium part of the question (part a), there was a 1.0 percentage point difference in item missing data rates between treatments. About 3.6 percent of the sample in the control treatment did not respond to the question compared with 2.6 percent in the test treatment.³⁷ However, mode-specific results indicated that the item missing data rate differed between treatments for computer-assisted interview modes (CATI and CAPI). Item missing data rates were not significantly different between treatments, overall or for any modes, for the subsidy part of the question (part b).

Table 14. Health Insurance Premium and Subsidy Item Missing Data Rates for Control and Test Treatments

Mode	Test Count	Test Percent	Control Count	Control Percent	Test minus Control	Adjusted P-value
Premium	12,137	2.6 (0.2)	13,745	3.6 (0.3)	-1.0 (0.4)	0.04*
Internet	4,304	0.7 (0.3)	5,177	0.8 (0.2)	-0.1 (0.3)	0.94
Mail	3,118	9.0 (0.8)	3,662	9.9 (0.7)	-0.9 (1.2)	0.94
CATI	704	1.3 (0.5)	749	5.6 (1.4)	-4.4 (1.5)	0.01*
CAPI	4,011	1.3 (0.3)	4,157	2.6 (0.4)	-1.3 (0.4)	0.04*
Subsidy	4,441	2.9 (0.4)	5,494	3.2 (0.3)	-0.3 (0.5)	0.77
Internet	2,065	0.6 (0.2)	2,845	1.0 (0.2)	-0.4 (0.3)	0.87
Mail	1,383	6.8 (0.9)	1,680	5.7 (0.8)	1.1 (1.2)	0.63
CATI	297	2.1 (1.1)	308	5.7 (2.1)	-3.6 (2.4)	0.87
CAPI	696	2.7 (1.0)	661	4.3 (1.1)	-1.5 (1.5)	0.87

Source: U.S. Census Bureau, 2016 American Community Survey Content Test

Note: Standard errors are shown in parentheses. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test at the $\alpha=0.1$ level. P-values have been adjusted for multiple comparisons using the Holm-Bonferroni method.

³⁷ The “control” version was the version paired with the control version of the Coverage question, and the “test” version was the version paired with the test version of the Coverage question.

In the paper version of the questionnaire, only persons with reported health insurance coverage were instructed to complete the Premium and Subsidy question. In all other modes, only persons with direct-purchase coverage, Medicaid, or “other” coverage were asked about whether they paid a health insurance premium. To maintain a consistent universe across modes, all subsequent analyses were restricted to respondents with direct-purchase or Medicaid coverage (unless otherwise stated). For all interview modes, only individuals who reported paying a premium in part a were asked whether the premium was subsidized in part b.

Do the versions have a different proportion of individuals who are in the universe for the Premium part of the Premium and Subsidy Question? Do the versions have different proportions of individuals who are in the universe for the Subsidy part of the question? (Coming from the Premium part of the question.)

Table 15. Universe for Health Insurance Premium and Subsidy Question by Interview Mode, Control and Test Treatments

Mode	Test Sample Size	Test Percent	Control Sample Size	Control Percent	Test minus Control	Adjusted P-value
Premium	39,495	29.0 (0.6)	39,723	31.8 (0.5)	-2.9 (0.7)	<0.01*
Internet	18,547	20.6 (0.7)	18,314	25.3 (0.6)	-4.7 (0.9)	<0.01*
Mail	9,264	31.5 (0.9)	9,816	34.2 (1.0)	-2.7 (1.3)	0.10
CATI	1,815	33.7 (1.8)	1,799	38.4 (2.0)	-4.7 (3.0)	0.22
CAPI	9,869	36.4 (1.1)	9,794	37.1 (1.2)	-0.7 (1.5)	0.65
Subsidy	12,137	39.5 (1.0)	13,745	41.6 (1.0)	-2.1 (1.3)	0.31
Internet	4,304	55.2 (1.5)	5,177	62.5 (1.3)	-7.2 (2.1)	<0.01*
Mail	3,118	52.2 (1.5)	3,662	54.9 (1.4)	-2.8 (2.0)	0.35
CATI	704	52.8 (3.3)	749	49.8 (3.8)	3.0 (5.4)	0.58
CAPI	4,011	23.8 (1.5)	4,157	19.5 (1.4)	4.3 (2.0)	0.12

Source: U.S. Census Bureau, 2016 American Community Survey Content Test

Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test at the $\alpha=0.1$ level. P-values have been adjusted for multiple comparisons using the Holm-Bonferroni method.

Due to differences in the response distribution for the lead-in Health Insurance Coverage question (described in Section 5.3), the universe for the premium part of the Health Insurance Premium and Subsidy question was significantly lower in the test treatment than in the control treatment. About 29.0 percent of the test treatment and 31.8 percent of the control treatment was in-universe for the premium part of the question, a 2.9 percentage point difference. This difference, however, was mode-specific: the percentage of individuals in universe only differed for internet responses.

Additionally, while there was no overall difference in the percent of individuals in universe for the Subsidy question, there were mode-specific differences. The percentage of individuals in universe for the Subsidy question was 7.2 percentage points smaller in the test treatment than in the control treatment for internet respondents but did not differ for the other interview modes.

5.3. Response Distributions

Several research questions concerned the response distributions of the Health Insurance Coverage question. Results in this section also take into account any coverage types that respondents may have included in the write-in field.

*Does the test version of the Health Insurance question have a higher proportion of persons with any type of health insurance coverage compared with the control version?*³⁸

One of the primary criteria for evaluating the two versions of the Health Insurance Coverage question was the overall insured rate. Specifically, the version of the question that results in a higher insured rate (a lower uninsured rate) compared with the other treatment was preferable, as previous research has documented an underreporting of Medicaid and other means-tested programs (O'Hara, 2010)(Boudreaux, Ziegenfuss, Graven, Davern, & Blewett, 2011). All things being equal, this undercount would result in a lower health insurance coverage rate.

As a result, the test version of the question was hypothesized to have a higher proportion of persons reporting any type of health insurance coverage than the existing question. However, this was not the case. About 89.1 percent of persons in the test treatment had any type of health insurance coverage compared with 91.4 percent in the control treatment, a 2.3 percentage point difference (Table 16).³⁹

Table 16. Response Distribution for Any Insurance Coverage by Interview Mode, Control and Test Treatments

Mode	Test Sample Size	Test Percent	Control Sample Size	Control Percent	Test minus Control	Adjusted P-Value
Internet	18,547	95.8 (0.2)	18,314	95.5 (0.3)	0.3 (0.3)	0.37
Mail	9,264	94.1 (0.4)	9,816	94.3 (0.4)	-0.2 (0.5)	0.72
CATI	1,815	93.9 (1.0)	1,799	91.5 (1.4)	2.4 (1.7)	0.16
CAPI	9,869	79.3 (0.9)	9,794	85.6 (0.6)	-6.3 (1.1)	<0.01*
Overall	39,495	89.1 (0.4)	39,723	91.4 (0.3)	-2.3 (0.5)	<0.01*

Source: U.S. Census Bureau, 2016 American Community Survey Content Test

Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference at the $\alpha=0.1$ level. P-values have been adjusted for multiple comparisons using the Holm-Bonferroni method.

Comparing results across interview modes revealed substantial heterogeneity. Coverage rates did not differ across treatments for the two self-administered modes (internet and mail). The insured rate was 2.4 percentage points higher in the test treatment than in the control treatment for telephone (CATI) responses (93.9 percent versus 91.5 percent, respectively). For in-person responses, rates differed by 6.3 percentage points: 85.6 percent and 79.3 percent for the control

³⁸ Since Indian Health Service (IHS) coverage is not comprehensive (State Health Access Data Assistance Center, 2005), it is not included in the total insured rate. Unless otherwise specified, differences in IHS coverage rates are not analyzed in this report.

³⁹ The REAP specified a one-tailed t-test because the test version was expected to have a higher insured rate. However, given the observed response distribution, a two-tailed t-test was also estimated.

and test treatments, respectively. This mode-specific difference is discussed in greater detail in the next section. In sum, the revised test Health Insurance Coverage question did not outperform the existing control question according to a key decision criterion and performed inconsistently across survey interview modes.

In CATI/CAPI interviews, how often was the option for respondents to volunteer the same type of coverage as Person 1 (for Persons 2+) used in the test version?

As noted in the Background, the test version of the question differed from the control version in six ways. Among these differences was the introduction of a new option present in computer-assisted interview modes for respondents who reported that they had the same coverage type as the first person in the household. Its introduction was intended to help reduce respondent burden and/or respondent error. In households with at least two persons present, the respondent could report that Persons 2+ had the same health insurance coverage type(s) as Person 1. If the respondent volunteered that the second (or higher) person has the same type of coverage as the first person in the household, the field representative could select this option.

The “same as Person 1” option was used in 35.2 percent of reports for Persons 2+ (n= 2,526). This option was more frequently used in CAPI interviews than in CATI interviews. In CATI, 975 persons were Persons 2+ and, in CAPI, there were 6,197 Persons 2+. About 13.8 percent of CATI persons and 38.6 percent of CAPI persons employed the option (135 persons and 2,391 persons, respectively).

Variation by coverage type

Are rates of coverage by employer-based insurance, direct purchase insurance, TRICARE, VA Care, and Medicare consistent between the test and control versions?

There is also substantial heterogeneity by type of health insurance coverage. The overall insured rate is comprised of six different types of insurance coverage: three types of private coverage (employer-based, direct, and TRICARE) and three types of public coverage (Medicare, Medicaid, and VA). Table 17 reports the overall response distribution for the two versions of the Health Insurance Coverage question across coverage types. The next several sections describe these results in greater detail.

Table 17. Response Distribution for Specific Health Insurance Coverage Types, Control and Test Treatments

Coverage Type	Test Percent (n=39,495)	Control Percent (n=39,723)	Test minus Control	Adjusted P-Value
Private	64.1 (0.5)	68.6 (0.5)	-4.5 (0.7)	<0.01*
Employer-based	51.3 (0.6)	55.0 (0.6)	-3.7 (0.8)	<0.01*
Direct purchase	11.3 (0.4)	13.0 (0.4)	-1.7 (0.5)	<0.01*
TRICARE	2.6 (0.2)	2.8 (0.2)	-0.2 (0.3)	0.49
Public	32.7 (0.5)	32.6 (0.5)	0.1 (0.6)	0.91
Medicare	16.7 (0.4)	15.7 (0.3)	1.0 (0.5)	0.07*
Medicaid	17.2 (0.5)	17.8 (0.5)	-0.7 (0.6)	0.54
VA	1.9 (0.1)	2.0 (0.1)	-0.2 (0.2)	0.54
Any insurance	89.1 (0.4)	91.4 (0.3)	-2.3 (0.5)	<0.01*

Source: U.S. Census Bureau, 2016 American Community Survey Content Test

Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test at the $\alpha=0.1$ level. P-values have been adjusted for multiple comparisons using the Holm-Bonferroni method. Some people may have more than one coverage type.

Employer-Based Coverage

Examining the response distribution for employer-based coverage, which is the most common insurance type (Barnett & Vornovitsky, 2016), revealed a substantial difference in coverage rates between versions of the Health Insurance Coverage question (Table 18). Overall, employer-based coverage rates were 55.0 and 51.3 percent for the control and test versions, respectively, a 3.7 percentage point difference. However, differences between treatments were only observed for CAPI responses (7.8 percentage point difference); significant differences were not present in any other mode.

Table 18. Response Distribution for Employer-Based Health Insurance by Interview Mode, Control and Test Treatments

Mode	Test Sample Size	Test Percent	Control Sample Size	Control Percent	Test minus Control	Adjusted P-value
Internet	18,547	65.6 (0.7)	18,314	67.5 (0.7)	-1.8 (0.9)	0.15
Mail	9,264	49.1 (0.9)	9,816	49.8 (0.9)	-0.8 (1.2)	1.00
CATI	1,815	47.5 (2.2)	1,799	47.9 (2.0)	-0.3 (2.9)	1.00
CAPI	9,869	37.3 (1.2)	9,794	45.1 (1.3)	-7.8 (1.8)	<0.01*
Overall	39,495	51.3 (0.6)	39,723	55.0 (0.6)	-3.7 (0.8)	<0.01*

Source: U.S. Census Bureau, 2016 American Community Survey Content Test

Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test at the $\alpha=0.1$ level. P-values have been adjusted for multiple comparisons using the Holm-Bonferroni method.

Given the large difference between treatments for the most common insurance type in the second most common interview mode (comprising about a quarter of the sample), we also recalculated the overall insured rate excluding information from CAPI responses. Without CAPI responses, the insured rate did not significantly differ between treatments.

Direct-Purchase Coverage

Several of the changes in the test version were designed to reduce the overreporting of direct-purchase coverage. For example, the test version added instructions to “NOT include plans that cover only one type of service, such as dental, drug or vision plans” (which are often directly purchased) and moved the direct purchase checkbox from the second item listed to the fourth item listed. Additionally, the test version of the Coverage question explicitly mentioned “a State or Federal Marketplace, HealthCare.gov, or a similar state website” as a means of obtaining direct-purchase coverage.

Direct-purchase rates differed between the test and control treatments (Table 19). Overall direct-purchase rates were 13.0 percent and 11.3 percent for the control and test versions of the Health Insurance Coverage question, respectively. Prior research suggested that ACS direct-purchase coverage rates are higher than other surveys’ rates (Mach & O’Hara, 2011); therefore, a lower direct-purchase rate was considered preferable.

Table 19. Response Distribution for Direct-Purchase Health Insurance by Interview Mode, Control and Test Treatments

Mode	Test Sample Size	Test Percent	Control Sample Size	Control Percent	Test minus Control	Adjusted P-value
Internet	18,547	10.4 (0.5)	18,314	14.1 (0.5)	-3.7 (0.7)	<0.01*
Mail	9,264	16.3 (0.7)	9,816	18.5 (0.7)	-2.2 (1.0)	0.06*
CATI	1,815	18.9 (1.6)	1,799	21.6 (2.0)	-2.6 (2.5)	0.57
CAPI	9,869	9.4 (0.7)	9,794	8.5 (0.7)	0.9 (0.9)	0.57
Overall	39,495	11.3 (0.4)	39,723	13.0 (0.4)	-1.7 (0.5)	<0.01*

Source: U.S. Census Bureau, 2016 American Community Survey Content Test

Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test at the $\alpha=0.1$ level. P-values have been adjusted for multiple comparisons using the Holm-Bonferroni method.

As indicated in Table 19, lower direct-purchase rates in the test treatment than in the control treatment were limited to the two self-administered interview modes. Compared with the control treatment, coverage rates were 3.7 percentage points lower in the test version among internet responses and 2.2 percentage points lower among mail responses. Coverage rates did not significantly differ between test and control treatments in the two computer-assisted interview modes (CATI and CAPI). However, as described later in the report, direct-purchase rates did differ between treatments for select populations.

Medicare

The final insurance type that differed between the test and control treatments was Medicare. The test version of the question listed Medicare second; the control version listed it third. As shown in Table 20, overall Medicare coverage rates were 15.7 and 16.7 percent for the control and test versions, a 1.0 percentage point difference. The results in Table 20 show that Medicare coverage rates were higher in the test treatment than in the control treatment only for mail and CAPI responses. Rates did not differ between treatments in the other two modes.

Table 20. Response Distribution for Medicare by Interview Mode, Control and Test Treatments

Mode	Test Sample Size	Test Percent	Control Sample Size	Control Percent	Test minus Control	Adjusted P-value
Internet	18,547	12.8 (0.4)	18,314	13.8 (0.4)	-1.0 (0.6)	0.18
Mail	9,264	33.1 (0.8)	9,816	30.3 (0.6)	2.8 (1.0)	0.02*
CATI	1,815	39.1 (2.0)	1,799	39.1 (2.3)	<0.1 (3.0)	0.99
CAPI	9,869	11.3 (0.6)	9,794	8.3 (0.5)	3.0 (0.8)	<0.01*
Overall	39,495	16.7 (0.4)	39,723	15.7 (0.3)	1.0 (0.5)	0.07*

Source: U.S. Census Bureau, 2016 American Community Survey Content Test

Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test at the $\alpha=0.1$ level. P-values have been adjusted for multiple comparisons using the Holm-Bonferroni method.

Medicaid

Does the test version of the Health Insurance question have a higher proportion of persons with Medicaid compared with the control version?

Several of the changes in the test version were designed to increase the reporting of Medicaid. The test version of the Health Insurance Coverage question reordered coverage types, moving Medicaid from the fourth type listed (after employer-based, direct-purchase, and Medicare) to the third type listed (after employer-based and Medicare). The test version also removed the phrase "or a disability" from the existing question. With the expansion of Medicaid as part of the Affordable Care Act, people in many states do not need to have a disability to be eligible for Medicaid.

Contrary to hypotheses, the test version of the Health Insurance Coverage question did not have a higher proportion of persons with Medicaid compared with the control version, overall or for any interview mode. That is, the Content Test results do not suggest that the test version of the Health Insurance Coverage question increased the Medicaid coverage rate (Table 17).

Is there a difference in coverage types, especially means-tested and direct-purchase, between test and control responses, by state Medicaid Expansion status?

We also examined Medicaid and direct-purchase coverage rates by Medicaid expansion status. Provisions of the ACA allowed states to expand Medicaid eligibility (Table 21). By the time the Content Test was fielded in March 2016, 30 states and the District of Columbia had expanded their Medicaid eligibility.

Table 21. Response Distribution for Direct-Purchase and Medicaid Coverage by Medicaid Expansion Status, Control and Test Treatments

Coverage Type	Test Sample Size	Test Percent	Control Sample Size	Control Percent	Test minus Control	Adjusted P-value
Expansion ¹						
Direct- purchase	22,101	9.8 (0.5)	22,356	11.4 (0.4)	-1.6 (0.6)	0.03*
Medicaid	22,101	18.4 (0.7)	22,356	18.9 (0.6)	-0.6 (0.9)	1.00
Non-expansion						
Direct- purchase	17,394	11.3 (0.5)	17,367	12.4 (0.6)	-1.1 (0.9)	0.60
Medicaid	17,394	14.3 (0.7)	17,367	14.3 (0.6)	<-0.1 (0.9)	1.00

Source: U.S. Census Bureau, 2016 American Community Survey Content Test

¹As of March 2016, Medicaid expansion states are AK, AR, AZ, CA, CT, CO, DE, DC, HI, IA, IL, IN, KY, MA, MD, MI, MN, MT, ND, NH, NM, NJ, NY, NV, OH, OR, PA, RI, VT, WA, and WV.

Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test at the $\alpha=0.1$ level. P-values have been adjusted for multiple comparisons using the Holm-Bonferroni method.

Medicaid coverage rates also did not differ between the test and control versions regardless of whether the respondent resided in a Medicaid Expansion state. However, the overall direct coverage rate is 1.6 percentage points lower in the test treatment than the control treatment only for expansion states.

Does the test version of the Health Insurance Coverage question have a lower proportion of persons with multiple types of health insurance compared with the control version?

Does the test version of the Health Insurance Coverage question result in a smaller proportion of persons who reported having both employer-provided insurance and insurance purchased directly compared with the control version?

Revisions to the Health Insurance Coverage question also sought to decrease the number of individuals who reported multiple types of coverage, particularly in unexpected combinations.

The revised Coverage question appeared to help respondents report insurance coverage type more accurately. Compared with respondents in the control treatment, individuals in the test treatment were less likely to have multiple types of health insurance coverage (Table 22). About 13.6 percent of respondents in the control treatment had more than one type of health insurance reported, while only 10.6 percent of respondents in the test treatment did.

Table 22. Response Distribution for People Who Reported Multiple Types of Health Insurance Coverage by Interview Mode, Control and Test Treatments

Mode	Test Percent (n=39,495)	Control Percent (n=39,723)	Test minus Control	Adjusted P-value
Internet	6.0 (0.2)	12.8 (0.4)	-6.8 (0.5)	<0.01*
Mail	22.6 (0.7)	23.6 (0.6)	-1.1 (0.9)	0.48
CATI	29.8 (2.0)	34.4 (2.1)	-4.6 (2.9)	0.34
CAPI	8.4 (0.5)	7.6 (0.6)	0.8 (0.7)	0.48
Overall	10.6 (0.3)	13.6 (0.3)	-3.0 (0.4)	<0.01*

Source: U.S. Census Bureau, 2016 American Community Survey Content Test

Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test at the $\alpha=0.1$ level. P-values have been adjusted for multiple comparisons using the Holm-Bonferroni method.

The test version of the health insurance coverage question also resulted in a smaller proportion of responses with the specific combination of employer-provided and direct-purchase insurance (Table 23). In the test version, 0.5 percent of persons had both of these coverage types marked, compared with 1.4 percent in the control version. Differences were limited to self-reported interview modes (mail and internet).

Table 23. Response Distribution for People Who Reported Both Employer-Based and Direct-Purchase Coverage by Interview Mode, Control and Test Treatments

Mode	Test Percent (n=39,495)	Control Percent (n=39,723)	Test minus Control	Adjusted P-value
Internet	0.4 (0.1)	1.9 (0.2)	-1.5 (0.2)	<0.01*
Mail	0.9 (0.1)	1.7 (0.2)	-0.9 (0.3)	<0.01*
CATI	2.4 (0.8)	2.7 (0.6)	-0.3 (1.0)	0.77
CAPI	0.4 (0.1)	0.6 (0.1)	-0.2 (0.2)	0.68
Overall	0.5 (0.1)	1.4 (0.1)	-0.9 (0.1)	<0.01*

Source: U.S. Census Bureau, 2016 American Community Survey Content Test

Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test at the $\alpha=0.1$ level. P-values have been adjusted for multiple comparisons using the Holm-Bonferroni method. Estimates based on checkboxes only; for this analysis, write-in coverage types were not considered.

Checkboxes and Write-Ins

The Health Insurance Coverage question allows respondents to write in an “other” type of health insurance. As noted in the Methodology section, these entries are assigned (where possible) to one of the seven health insurance coverage types listed on the Coverage question. The results presented thus far have included coverage that was determined through the coding of write-in entries. To examine how write-in entries might have influenced estimates, Table 24 below shows results for the response distribution of health insurance coverage excluding write-in responses. In general, the same patterns as described in Table 17 were observed: the employer-based and direct coverage rates were significantly lower in the test version than in the control version, and the Medicare rate was significantly higher in the test version than in the control version.

Table 24. Response Distribution for Health Insurance Coverage Checkboxes (Excluding Write-in Responses), Control and Test Treatments

Coverage Type	Test Percent (n=39,495)	Control Percent (n=39,723)	Test minus Control	Adjusted P-value
Employer-based	51.2 (0.6)	54.8 (0.6)	-3.7 (0.8)	<0.01*
Direct	10.4 (0.4)	11.8 (0.4)	-1.4 (0.5)	0.03*
Medicare	16.6 (0.4)	15.6 (0.3)	1.0 (0.5)	0.10*
Medicaid	16.8 (0.5)	17.1 (0.5)	-0.3 (0.6)	1.00
TRICARE	2.6 (0.2)	2.7 (0.2)	-0.2 (0.3)	1.00
VA	1.9 (0.1)	2.0 (0.1)	-0.2 (0.2)	0.92

Source: U.S. Census Bureau, 2016 American Community Survey Content Test

Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test at the $\alpha=0.1$ level. P-values have been adjusted for multiple comparisons using the Holm-Bonferroni method.

Does the test version of the Health Insurance Coverage question decrease the proportion of persons who write-in an “other” type of health insurance coverage compared with the control version?

We hypothesize that respondents used write-ins for one of two reasons. First, respondents might not have understood the question or might have been unsure how to classify their insurance coverage. They could have written in information for the Census Bureau to classify their coverage. Second, respondents might have provided additional information about a plan they already reported via checkbox. In both cases, a decrease in the proportion of people who write in a type of health insurance coverage would be preferable.

The percentage of respondents in each treatment that elected to write in a type of health insurance differed between versions (Table 25). The write-in rate was 1.6 percentage points lower for the test version of the health insurance coverage question compared with the control version.⁴⁰ Whereas 3.0 percent of persons in the test treatment had information in the write-in field, 4.5 percent of persons in the control treatment did.

Table 25. Response Distribution for Health Insurance Coverage Type Write-in by Interview Mode, Control and Test Treatments

Mode	Test Percent (n=39,495)	Control Percent (n=39,723)	Test minus Control	Adjusted P-value
Internet	1.8 (0.2)	3.9 (0.2)	-2.1 (0.3)	<0.01*
Mail	5.0 (0.4)	8.1 (0.5)	-3.1 (0.6)	<0.01*
CATI	2.7 (0.7)	5.2 (1.4)	-2.5 (1.6)	0.23
CAPI	3.2 (0.4)	3.3 (0.4)	-0.1 (0.6)	0.85
Overall	3.0 (0.2)	4.5 (0.2)	-1.6 (0.3)	<0.01*

Source: U.S. Census Bureau, 2016 American Community Survey Content Test

Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test at the $\alpha=0.1$ level. P-values have been adjusted for multiple comparisons using the Holm-Bonferroni method.

⁴⁰ These rates are for persons who included any information in the write-in field, regardless of whether it was ultimately coded as a specific coverage type.

Health Insurance Coverage Response Distribution by Subgroup: Is there a difference in coverage types when divided by age or military status?

Comparing the two versions of the Health Insurance Coverage question by age and by military status helps to identify whether respondents might be misclassifying their health insurance coverage. Our analyses were limited to these social and demographic characteristics. For example, due to sample size constraints, the testing of other questions (e.g., about household relationships), and the unavailability of edited data, we were unable to examine differences by health insurance unit (HIU) income-to-poverty ratios.

Table 26 below shows differences in Medicare coverage by age.⁴¹ Consistent with enrollment criteria, the highest Medicare enrollment rates are among adults aged 65 or older. However, results presented in the table indicate that the 1.0 percentage point higher Medicare coverage rate in the test version (compared with the control version) was only present for working-age adults. About 3.2 percent of 19-64 year-old adults in the control treatment reported having Medicare compared with 3.8 percent in the test treatment. In this age bracket, only individuals with kidney disease and certain other specific illnesses/disabilities would be eligible for Medicare coverage.

Table 26. Response Distribution for Medicare Coverage by Age, Control and Test Treatments

Age	Test Percent (n=39,495)	Control Percent (n=39,723)	Test minus Control	Adjusted P-value
Under 19	0.7 (0.2)	0.4 (0.1)	0.3 (0.2)	0.32
19-64	3.8 (0.2)	3.2 (0.1)	0.6 (0.2)	0.05*
65+	86.9 (0.6)	87.3 (0.5)	-0.3 (0.8)	0.68
Overall	16.6 (0.4)	15.6 (0.3)	1.0 (0.5)	0.07*

Source: U.S. Census Bureau, 2016 American Community Survey Content Test

Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test at the $\alpha=0.1$ level. P-values have been adjusted for multiple comparisons using the Holm-Bonferroni method.

In contrast, direct-purchase coverage rates only differed between the two treatments for older adults (individuals 65 or older). As shown in Table 27, the direct-purchase coverage rate was 5.7 percentage points lower in the test treatment than in the control treatment (25.6 percent in the control treatment versus 19.9 percent in the test treatment). The test version included a number of changes that could have affected direct purchase reporting, including instructions that respondents should only report comprehensive plans. The majority of older adults (about 87 percent in both treatments) report Medicare coverage.

⁴¹ Children under the age of 19 are eligible for Medicaid/Children’s Health Insurance Program (CHIP) coverage.

Table 27. Response Distribution Direct-Purchase Insurance by Age, Control and Test Treatments

Age	Test Percent (n=39,495)	Control Percent (n=39,723)	Test minus Control	Adjusted P-value
Under 19	5.4 (0.6)	6.1 (0.5)	-0.6 (0.8)	0.44
19-64	9.9 (0.3)	10.6 (0.4)	-0.7 (0.5)	0.33
65+	19.9 (0.9)	25.6 (0.9)	-5.7 (1.0)	<0.01*
Overall	10.4 (0.4)	11.8 (0.4)	-1.4 (0.5)	0.02*

Source: U.S. Census Bureau, 2016 American Community Survey Content Test

Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test at the $\alpha=0.1$ level. P-values have been adjusted for multiple comparisons using the Holm-Bonferroni method.

For all age groups, the employer-based coverage rates were higher in the control treatment than in the test treatment (Table 28).

Table 28. Response Distribution for Employer-Based Insurance by Age, Control and Test Treatments

Age	Test Percent (n=39,495)	Control Percent (n=39,723)	Test minus Control	Adjusted P-value
Under 19	48.0 (1.1)	52.3 (1.2)	-4.3 (1.6)	0.01*
19-64	59.8 (0.7)	62.4 (0.6)	-2.6 (0.9)	0.01*
65+	24.4 (0.8)	29.5 (0.8)	-5.1 (1.2)	<0.01*
Overall	51.2 (0.6)	54.8 (0.6)	-3.7 (0.8)	<0.01*

Source: U.S. Census Bureau, 2016 American Community Survey Content Test

Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test at the $\alpha=0.1$ level. P-values have been adjusted for multiple comparisons using the Holm-Bonferroni method.

Table 28 presents results stratified by military status. For VA care, we examined veterans and nonveterans, as veterans would be eligible to have VA coverage. For TRICARE, we examined the presence of active military in the household, the absence of active military in the household, and all other households. TRICARE is a type of coverage for active duty military, their families, and certain military retirees. As Content Test data, including military status, did not go through the standard ACS editing process, the “other” category includes those persons with missing data on military status.

Rates of reported VA coverage were similar across military status. About 30 percent of veterans and 0.3 percent of non-veterans in each treatment reported VA coverage. Moreover, TRICARE coverage rates did not differ between the test and control treatments for any of the military-status groups.

Table 29. Response Distribution for VA and Military Coverage by Military Status, Control and Test Treatments

Military Status	Test Sample Size	Test Percent	Control Sample Size	Control Percent	Test minus Control	Adjusted P-value
VA						
Veteran	2,286	29.0 (1.3)	2,364	30.0 (1.4)	-1.0 (1.9)	1.00
Non-veteran	37,209	0.3 (<0.1)	37,359	0.3 (0.1)	0.0 (0.1)	1.00
Overall	39,495	1.9 (0.1)	39,723	2.0 (0.1)	-0.1 (0.2)	0.92
TRICARE						
Active military in household	1,434	24.8 (2.6)	1,340	27.6 (3.9)	-2.8 (4.9)	1.00
No active military in household	4,928	10.5 (1.2)	5,063	10.6 (0.8)	-0.1 (1.4)	1.00
Other	33,133	0.4 (<0.1)	33,320	0.5 (0.1)	-0.2 (0.1)	0.11
Overall	39,495	2.6 (0.2)	39,723	2.7 (0.2)	-0.2 (0.3)	1.00

Source: U.S. Census Bureau, 2016 American Community Survey Content Test

Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test at the $\alpha=0.1$ level. P-values have been adjusted for multiple comparisons using the Holm-Bonferroni method.

Health Insurance Premium and Subsidy Question

The Health Insurance Premium and Subsidy question was introduced to help determine whether respondents with direct-purchase health insurance coverage had subsidized Marketplace coverage. As a result, the response distribution was examined in tandem with the Health Insurance Coverage question. However, because the test and control versions of the Premium and Subsidy Question were directly tied to test and control versions of the Coverage question and because the universes for the two versions of the Premium and Subsidy question differed (see Table 15), it is difficult to isolate the unique contribution of each question.⁴²

Moreover, only persons with direct-purchase, Medicaid, or “other” coverage completed the Premium question in the computer-assisted versions of the questionnaire. All analyses of the response distribution for the Premium and Subsidy question (i.e., subsidized Marketplace coverage) were restricted to those reporting one of those two coverage types (either via checkbox or write-in) regardless of response mode to permit comparisons across modes.

Using the Health Insurance Coverage question and Premium and Subsidy question in combination, what are the rates of subsidized Marketplace coverage in each version?

Table 30 shows the estimated subsidized Marketplace coverage rates across the control and test treatments. We classified persons as having such coverage if they reported direct purchase insurance, paying a premium, and receiving a subsidy. Overall, 1.8 percent of persons in the control treatment had subsidized Marketplace coverage, a rate that was not statistically different

⁴² The “control” version was the version paired with the control version of the Coverage question, and the “test” version was the version paired with the test version of the Coverage question.

from the rate in the test treatment. Indeed, subsidized Marketplace coverage rates did not differ in any mode.⁴³

Table 30. Response Distribution for Subsidized Marketplace Coverage by Interview Mode, Control and Test Treatments

Mode	Test Percent (n=39,495)	Control Percent (n=39,723)	Test minus Control	Adjusted P-value
Internet	2.1 (0.2)	2.3 (0.2)	-0.2 (0.3)	1.00
Mail	2.5 (0.3)	1.9 (0.2)	0.6 (0.4)	0.59
CATI	1.6 (0.4)	1.8 (0.7)	-0.2 (0.9)	1.00
CAPI	1.3 (0.2)	1.0 (0.2)	0.3 (0.3)	1.00
Overall	1.8 (0.1)	1.7 (0.1)	0.1 (0.2)	1.00

Source: U.S. Census Bureau, 2016 American Community Survey Content Test

Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. Significance was tested based on a two-tailed t-test at the $\alpha=0.1$ level. P-values have been adjusted for multiple comparisons using the Holm-Bonferroni method.

Some persons with fully subsidized coverage might not have any reported out-of-pocket premium. As a result, we also examined the percent of individuals with direct coverage who reported no premium. About 1.6 percent of persons with direct coverage in the control treatment explicitly reported that they did not pay premium. This rate was not significantly different from the rate in the test treatment.

5.4. Benchmarks

Health Insurance Coverage Benchmarks

How do the proportions in each treatment compare with proportions found in other surveys?

To assess the accuracy of estimates from each version of the survey, we compared results to external benchmarks. For the Health Insurance Coverage question, estimates from the 2015 and 2016 National Health Interview Survey (NHIS) and the 2016 Current Population Survey Annual Social and Economic Supplement (CPS ASEC) served as benchmarks.

Although these three data sources offer health insurance coverage estimates, they differ in their primary focus, data collection timing, and reference period. ACS Content Test estimates represent current coverage at the time of interview (March 2016). However, NHIS estimates correspond to coverage at the time of interview (in January – March 2016 or October – December 2015), and CPS ASEC estimates include coverage at the time of interview (February – April 2016) as well as during the previous calendar year (2015). Comparisons were not formally tested via any statistical test.

As noted earlier, the insured rate in the 2016 ACS Content Test (fielded in March 2016) was 89.1 percent in the test treatment and 91.4 percent in the control treatment. These estimates are

⁴³ We also calculated the subsidized Marketplace rate for respondents in-universe to answer the two questions (i.e., those with Medicaid or direct-purchase coverage) (not shown). Once again, rates did not differ between treatments.

fairly close to benchmark estimates. According to the 2016 CPS ASEC, 90.4 percent of persons had health insurance at the time of survey (February-April 2016), and 90.9 percent had coverage at any point in 2015 (Barnett & Vornovitsky, 2016). ACS Content Test estimates of the overall insured rate are also similar to 2016 NHIS Early Release estimates (Cohen, Martinez, & Zammiti, 2016), which indicate that 91.4 percent of Americans were insured at the time of survey (January – March 2016).

Table 31. Comparison of ACS Content Test Treatments to External Benchmarks, Health Insurance Coverage

Category	Content Test Treatment	Content Test Control Treatment	2016 CPS ASEC Current Coverage	2016 CPS ASEC Prior Year	2016 NHIS
Reference period	March 2016	March 2016	February - March 2016	January - December 2015	January - March 2016
Any insurance	89.1 (0.3)	91.4 (0.4)	90.4 (0.1)	90.9 (0.2)	91.4 (0.3)
Any private	64.1 (0.5)	68.6 (0.5)	66.3 (0.2)	67.2 (0.4)	63.8 (0.7)
Any public	32.7 (0.5)	32.6 (0.5)	36.6 (0.2)	37.1 (0.3)	36.2 (0.6)
Employer-based	51.3 (0.6)	55.0 (0.6)	--	55.7 (0.4)	--
Direct	11.3 (0.4)	13.0 (0.4)	--	16.3 (0.3)	--
Medicare	16.7 (0.4)	15.7 (0.3)	16.3 (0.1)	16.3 (0.1)	--
Medicaid	17.2 (0.5)	17.8 (0.5)	18.9 (0.2)	19.6 (0.3)	--

Source: U.S. Census Bureau, 2016 American Community Survey Content Test; Barnett & Vornovitsky 2016; Cohen, Martinez, & Zammiti, 2016; National Center for Health Statistics 2015 National Health Interview Survey
 Note: Standard errors are shown in parentheses. Hyphens (--) indicate that data were not available.

Table 31 also compares 2016 ACS Content Test estimates of coverage rates with estimates from the 2016 CPS ASEC and 2016 NHIS. Although the surveys define private and public coverage differently (e.g., TRICARE is considered private coverage in the ACS and public coverage in the CPS ASEC), all three estimates are relatively similar to one another.

Employer-based coverage rates in the control treatment appeared closer to CPS ASEC estimates than did rates in the test treatment: 55.7 percent in the CPS ASEC compared with 55.0 percent and 51.3 percent in the ACS control and test conditions, respectively. Direct-purchase rates appeared higher in the CPS ASEC than in either Content Test treatment, and rates were higher in the ACS control treatment than in the test treatment.

As noted earlier, Medicaid coverage rates did not differ between the two Content Test treatments. Consistent with earlier research, Content Test Medicaid coverage rates appeared lower than benchmark rates. Compared with the control treatment, Medicare rates in the test treatment appeared more similar to CPS ASEC benchmarks. Benchmark estimates might nonetheless underestimate actual Medicare enrollment. Bhaskar and colleagues (2016) used administrative records to show that CPS ASEC estimates underestimate Medicare enrollment for adults aged 65 and older. However, as noted earlier, differences in Medicare coverage rates were limited to working-age adults.

Health Insurance Premium and Subsidized Marketplace Benchmarks

Which version produces estimates of subsidized Marketplace coverage that more closely match benchmarks (including administrative data)?

Are estimates of the proportions of persons who pay a premium comparable to estimates from National Health Interview Survey (NHIS)?

We compared estimates from the Premium Question to the 2016 NHIS. Though NHIS data were from the prior calendar year, they were the most recent data available. To increase comparability, NHIS data were limited to information from respondents interviewed during the final quarter of 2015 (October-December).

The percentages of persons with Medicaid or direct-purchase health insurance who paid a premium were lower in both Content Test treatments than in the NHIS. In 2015, 28.2 percent of NHIS respondents with one of those insurance types reported paying a premium compared with 39.5 percent in the Content Test control treatment.⁴⁴ The control treatment rate was not statistically different from the test treatment rate. The ACS has lower rates of Medicaid and other means-tested coverage than other surveys (see previous table) (Boudreaux, Ziegenfuss, Graven, Davern, & Blewett, 2011) that would result in a smaller denominator. If not reporting Medicaid is associated with not having a premium, then this difference could contribute to lower premium rates in the NHIS than in the ACS insofar as the association is stronger in the ACS than in the NHIS.

We also compared our subsidized Marketplace rates with CMS enrollment statistics. In the Content Test, we classify people as having subsidized Marketplace coverage if they report direct-purchase health insurance and report having a premium and subsidy (i.e., responding “Yes” to both parts of the Premium and Subsidy question). In the CMS statistics, we focused on the number of Marketplace enrollees with Advance Premium Tax Credits (APTC) to capture subsidies. An APTC is a subsidy or tax credit to reduce premiums. In March 2016 quarterly data, CMS reported 9,389,609 persons were APTC enrollees. We used these numbers as a numerator and the Census Bureau’s July 2016 population estimates as the denominator (323,127,513 persons) (U.S. Census Bureau, 2017).

This constructed measure based on CMS estimates suggested that the subsidized Marketplace coverage rate was 2.9 percent. This estimate was larger than Content Test estimates, which were not significantly different between treatments. About 1.8 percent of persons with at least one valid health insurance response in the control treatment had subsidized Marketplace coverage (standard error = 0.1 percent).

⁴⁴ To make the ACS and NHIS estimates more comparable, this external benchmark includes persons who received coverage through Medicaid, State Children Health Insurance Programs (SCHIP), and other government programs. As a result, this benchmarked value may differ from other published NHIS estimates.

5.5. Response Error

Per our research questions, we also assessed whether response reliability differed across treatments.⁴⁵ Respondents were reinterviewed so their responses could be compared across two time points. We focused on two measures of reliability: gross difference rates (GDR) and indices of inconsistency (IOI) (see Section 2.4).⁴⁶

Health Insurance Coverage

Are the measures of response reliability (gross difference rate, index of inconsistency) better for the test treatment than for the control treatment?

As the revised Health Insurance Coverage question helped to clarify coverage types, we hypothesized that responses in the test treatment would be more reliable than responses in the control treatment. Statistical significance between the GDR and IOI was determined using a two-tailed t-test.

Table 32. Difference in Gross Difference Rates (GDR) between Control and Test Treatments

Coverage Type	Test Sample Size	Test GDR Percent	Control Sample Size	Control GDR Percent	Test minus Control	Adjusted P-value
Employer-based	14,032	8.4 (0.6)	14,260	7.7 (0.5)	0.7 (0.8)	1.00
Direct-purchase	11,082	9.7 (0.6)	11,358	10.4 (0.5)	-0.8 (0.8)	1.00
Medicare	12,174	3.3 (0.2)	12,161	2.8 (0.3)	0.5 (0.3)	0.77
Medicaid	11,200	6.6 (0.5)	11,300	7.1 (0.5)	-0.5 (0.7)	1.00
TRICARE	10,748	1.1 (0.3)	10,794	1.0 (0.3)	0.1 (0.3)	1.00
VA	10,743	1.4 (0.2)	10,750	1.4 (0.2)	0.1 (0.3)	1.00
Indian Health	10,563	0.2 (0.1)	10,552	0.3 (0.1)	-0.1 (0.1)	1.00
Other	10,378	5.2 (0.3)	10,580	7.3 (0.4)	-2.1 (0.6)	<0.01*

Source: U.S. Census Bureau, 2016 American Community Survey Content Test.

Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test at the $\alpha=0.1$ level. P-values have been adjusted for multiple comparisons using the Holm-Bonferroni method.

Table 32 presents GDRs for both the test and control treatments. As seen in the table, there were no significant differences for any of the six types of comprehensive health insurance coverage or for Indian Health Service coverage. However, the GDR was 2.1 percentage points lower in the test treatment than in the control treatment for the “other” health insurance checkbox. The percent of inconsistent answers for the “other” health insurance checkbox between the original interview and the reinterview (CFU) was 5.2 percent for the test treatment compared with 7.3 percent for the control treatment.

⁴⁵ As both versions of the Health Insurance Coverage question ask about *current* coverage, we would expect some change in responses between the initial interview and re-interview. However, we did not expect changing health insurance coverage to affect the two treatments differentially.

⁴⁶ For measures of response error, coded information from the write-in field was not used.

Table 33. Difference in Index of Inconsistency (IOI) between Control and Test Treatments

Coverage Type	Test Sample Size	Test IOI Percent	Control Sample Size	Control IOI Percent	Test minus Control	Adjusted P-value
Employer-based	14,032	17.1 (1.2)	14,260	16.5 (1.1)	0.6 (1.7)	1.00
Direct-purchase	11,082	38.6 (2.2)	11,358	37.6 (1.9)	1.0 (3.1)	1.00
Medicare	12,174	8.5 (0.6)	12,161	7.4 (0.6)	1.1 (0.8)	1.00
Medicaid	11,200	20.0 (1.5)	11,300	21.1 (1.3)	-1.1 (2.0)	1.00
TRICARE	10,748	12.1 (3.1)	10,794	10.2 (2.4)	2.0 (3.2)	1.00
VA	10,743	19.1 (2.6)	10,750	20.2 (2.7)	-1.0 (3.6)	1.00
Indian Health	10,563	25.6 (7.3)	10,552	20.1 (9.1)	5.4 (11.0)	1.00
Other	10,378	89.9 (2.6)	10,580	83.6 (2.4)	6.4 (3.6)	0.64

Source: U.S. Census Bureau, 2016 American Community Survey Content Test

Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test at the $\alpha=0.1$ level. P-values have been adjusted for multiple comparisons using the Holm-Bonferroni method.

The second measure of reliability, IOI, did not differ between treatments for any of the coverage type (Table 33). As noted earlier, per the Census Bureau’s general rule, index values of less than 20 percent indicate low inconsistency, 20 to 50 percent indicate moderate inconsistency, and over 50 percent indicate high inconsistency. Only “other” health insurance was in the high inconsistency category. Employer-based coverage, Medicare, and TRICARE in both treatments and VA coverage in the test treatment were in the low inconsistency category. The remainder of coverage types fell in the moderate inconsistency category. Overall, the GDR and IOI results did not provide evidence that the test version was more reliable than the control version.

We also examined the GDR and IOI for persons with multiple types of comprehensive coverage. These combinations are of particular interest because they are generally improbable and/or impossible.

Table 34. Difference in Gross Difference Rate (GDR) for Multiple Insurance Coverage between Control and Test Treatments

Coverage Type	Test Sample Size	Test Statistic Percent	Control Sample Size	Control Statistic Percent	Test minus Control	Adjusted P-value
Direct & Medicare	10,604	5.0 (0.4)	10,769	4.5 (0.4)	0.5 (0.5)	0.99
Direct & Medicaid	10,344	1.2 (0.2)	10,415	1.1 (0.1)	<0.1 (0.2)	1.00
Direct & Employer	10,400	1.7 (0.2)	10,612	3.3 (0.3)	-1.6 (0.4)	<0.01*
Medicare & Medicaid	10,602	2.4 (0.3)	10,589	2.5 (0.3)	-0.1 (0.4)	1.00

Source: U.S. Census Bureau, 2016 American Community Survey Content Test

Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test at the $\alpha=0.1$ level. P-values have been adjusted for multiple comparisons using the Holm-Bonferroni method.

Table 35. Difference in Index of Inconsistency (IOI) for Multiple Insurance Coverage between Control and Test Treatments

Coverage Type	Test Sample Size	Test Statistic Percent	Control Sample Size	Control Statistic Percent	Test minus Control	Adjusted P-value
Direct & Medicare	10,604	45.9 (3.4)	10,769	33.5 (2.3)	12.4 (3.9)	<0.01*
Direct & Medicaid	10,344	87.1 (4.7)	10,415	86.6 (4.5)	0.5 (6.6)	1.00
Direct & Employer	10,400	95.5 (2.4)	10,612	92.0 (3.0)	3.5 (3.9)	1.00
Medicare & Medicaid	10,602	35.2 (3.4)	10,589	38.1 (3.7)	-2.9 (4.4)	1.00

Source: U.S. Census Bureau, 2016 American Community Survey Content Test

Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test at the $\alpha=0.1$ level. P-values have been adjusted for multiple comparisons using the Holm-Bonferroni method.

Table 34 presents differences in response reliability between treatments for four combinations of insurance coverage (direct purchase and Medicare, direct purchase and Medicaid, direct-purchase and employer, and Medicare and Medicaid) as measured by the GDR. The GDR was 1.6 percentage points lower in the test treatment than in the control treatment for the report of both direct-purchase and employer-based health insurance (1.7 percent versus 13.3 percent, respectively). That is, the percent of inconsistent answers between the original interview and the reinterview was lower in the test treatment for this combination.⁴⁷ However, the GDR did not significantly differ between treatments for any other dual-coverage report.

Table 35 shows the IOI for both treatments for the same four combinations of insurance coverage. The IOI only significantly differed for the persons who reported both direct purchase coverage and Medicare. For this combination, the IOI was 45.9 percent in the test treatment and 33.5 percent in the control treatment. Both of these values fell into the “moderate inconsistency” category.

Health Insurance Premium and Subsidy

For each part of the question (Premium and Subsidy), which version (control/test) has better response reliability?

Table 36. Difference in Gross Difference Rate (GDR) between Control and Test Treatments

Question Part	Test Sample Size	Test Statistic Percent	Control Sample Size	Control Statistic Percent	Test minus Control	Adjusted P-value
Premium	3,822	7.5 (0.9)	4,228	8.7 (0.9)	-1.2 (1.1)	0.27
Subsidy	1,345	11.9 (1.8)	1,657	6.4 (0.9)	5.5 (1.9)	<0.01*

Source: U.S. Census Bureau, 2016 American Community Survey Content Test

Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. Significance was tested based on a two-tailed t-test at the $\alpha=0.1$ level. P-values have been adjusted for multiple comparisons using the Holm-Bonferroni method.

⁴⁷ As noted in the Response Distribution section (Section 5.3), both employer-based and direct-purchase rates differed between treatments, and these differences were not present for all interview modes.

Table 37. Difference in Index of Inconsistency (IOI) between Control and Test Treatments

Question Part	Test Sample Size	Test Statistic Percent	Control Sample Size	Control Statistic Percent	Test minus Control	Adjusted P-value
Premium	3,822	15.6 (1.9)	4,228	17.5 (1.8)	-1.9 (2.2)	0.39
Subsidy	1,345	32.7 (4.5)	1,657	19.8 (2.9)	12.9 (5.1)	0.02*

Source: U.S. Census Bureau, 2016 American Community Survey Content Test

Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. Significance was tested based on a two-tailed t-test at the $\alpha=0.1$ level. P-values have been adjusted for multiple comparisons using the Holm-Bonferroni method.

Table 36 and 37 report response reliability metrics for the Health Insurance Premium and Subsidy question. The GDR and IOI did not significantly differ between the test and control treatments for the premium part of the question (part a). However, response reliability significantly differed between treatments for the subsidy part of the question (part b). Table 36 shows that 11.9 percent of answers were inconsistent between the original interview and the reinterview for the test version of the question, but 6.4 percent of answers were inconsistent for the control version. Table 37 shows that the IOI was 12.9 percentage points higher in the test version than in control version. The test version fell into the “moderate inconsistency” category, whereas the control version fell into the “low inconsistency” category.

6. CONCLUSION AND RECOMMENDATIONS

The American Community Survey is the principal source of detailed health insurance coverage information for state and sub-state geographies due to its large sample size. The results of the 2016 ACS Content Test show that the proposed revision to the Health Insurance Coverage question did not appreciably and uniformly improve the accuracy of estimates. The *a priori* primary decision criterion concerned health insurance coverage rates; a higher insured rate was considered preferable. Prior research has detailed an underreporting of Medicaid and other means-tested programs (O'Hara, 2010), which, all else equal, would result in the underreporting of any health insurance coverage. Contrary to hypotheses, the revised version of the question did not produce a higher insured rate than the control (current production) version. The insured rate was 89.1 percent in the test version of the question and 91.4 percent in the control version.

Given the evidence that people underreport Medicaid, Medical Assistance, and other means-tested state-provided health plans, the second primary evaluation criterion stated that the version of the question that produced a higher Medicaid coverage rate was preferable. However, Medicaid coverage rates did not differ between versions.

The test version of the Coverage question did perform better than the control version based on some of the other evaluation criteria, particularly criteria related to survey production and processing (e.g., lower implied “No” rates, fewer reports of multiple coverage types, a lower direct purchase rate, and fewer coverage type write-ins). However, these differences were fairly limited in scope and are not consistent across all interview modes.⁴⁸

⁴⁸ Given indirect evidence that some of the six changes might have improved the accuracy of estimates, some of the changes could be tested in the future.

We also evaluated the two versions of the Health Insurance Premium and Subsidy question, which was introduced to measure subsidized Marketplace coverage. Statistical evidence did not suggest that one version of the question performed substantially better than the other. However, responses to the Health Insurance Coverage question determine who should respond to the Premium and Subsidy question. As a result, the contribution of each question to estimates of subsidized Marketplace coverage rates could not be fully isolated. For example, the direct purchase rate was lower in the test version of the Coverage question than in the control version, perhaps, because the test version explicitly reminded respondents to only report comprehensive health plans. Persons in the control treatment, therefore, might have reported information about supplemental plans when responding to the Premium and Subsidy question. Nonetheless, the “control” version of the Premium and Subsidy question has been tested with the preferred version of the Health Insurance Coverage question and is appreciably shorter, resulting in lower respondent burden.

7. ACKNOWLEDGEMENTS

The 2016 ACS Content Test would not have been possible without the participation and assistance of many individuals from the Census Bureau and other agencies. Their contributions are sincerely appreciated and gratefully acknowledged.

- Census Bureau staff in the American Community Survey Office, Application Development and Services Division, 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 the Office of Management and Budget’s Interagency Working Group for the ACS and the Topical Subcommittees formed by the Working Group for each topic tested on the 2016 ACS Content Test.
- Staff in the Office of Management and Budget’s Statistical and Science Policy Office.

The authors would like to thank the following individuals for their contributions to the analysis and review of the Health Insurance Coverage, Premiums, and Subsidies report: Jennifer Cheeseman Day, Michael Ikeda, Jeongsoo Kim, Jason Lee, Carla Medalia, Jennifer Ortman, Joanne Pascale, Elizabeth Poehler, Megan Rabe, Nicole Scanniello, Anthony Tersine, Marina Vornovitsky, David Waddington, and Nathan Walters.

8. REFERENCES

Barnett, J. C., & Vornovitsky, M. S. (2016). *Health Insurance Coverage in the United States: 2015*. Washington, DC: U.S. Government Printing Office.

- Bhaskar, R., Noon, J., O'Hara, B., & Velkoff, V. (2016). *Medicare Coverage and Reporting: A Comparison of the Current Population Survey and Administrative Records*. Retrieved January 10, 2017 from https://www.census.gov/srd/carra/16_12_Medicare_Cov_and_reporting_Comparison_of_CPS_and_Adms_Recs.pdf.
- Biemer, P. (2011). *Latent Class Analysis of Survey Error*. New York: Wiley.
- Boudreaux, M., Call, K. T., Turner, J., Fried, B., & O'Hara, B. (2015). Measurement Error in Public Health Insurance Reporting in the American Community Survey: Evidence from Record Linkage. *Health Services Research, 1973–1995*. Retrieved June 25, 2015, from U.S. Census Bureau Official Web Site: http://www.census.gov/hhes/www/hlthins/publications/ACSUundercount_WorkingPaper_0.pdf
- Boudreaux, M., Call, K., Turner, J., & Fried, B. (2014, May). *Estimates of Direct Purchase from the ACS and Medicaid Misreporting: Is there a link?* Retrieved June 25, 2015, from State Health Access Data Assistance Center: <http://shadac.org/publications/estimates-direct-purchase-acs-and-medicaid-misreporting-there-link>
- Boudreaux, M., Ziegenfuss, J. Y., Graven, P., Davern, M., & Blewett, L. A. (2011, February). Counting Uninsurance and Means-Tested Coverage in the American Community Survey: A Comparison to the Current Population Survey. *Health Services Research, 46, Number 1*(February 2011). Retrieved June 25, 2015, from <http://www.hsr.org/hsr/abstract.jsp?aid=46431245331>
- Cohen, R. A., Martinez, M. E., & Zammitti, E. P. (2016). *Health Insurance Coverage: Early Release of Estimates from the National Health Interview Survey, January - March 2016*. National Center for Health Statistics: National Health Interview Survey Early Release Program.
- Dusch, G., & Meier, F. (2012, June 13). *2010 Census Content Reinterview Survey Evaluation Report*. Retrieved February 11, 2013, from 2010 Census: http://www.census.gov/2010census/pdf/2010_Census_Content_Reinterview_Survey_Evaluation_Report.pdf
- Flanagan, P. (1996). *Survey Quality & Response Variance (Unpublished Internal Document)*. U.S. Census Bureau: Demographic Statistical Methods Division. Quality Assurance and Evaluation Branch.
- Flanagan, P. (2001). *Measurement Errors in Survey Response*. Baltimore, MD: University of Maryland Baltimore County.
- Holm, S. (1979). A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics, 6, 2, 65-70*. Retrieved January 11, 2017, from <https://www.jstor.org/stable/4615733>
- Lynch, V., Kenney, G. M., Haley, J., & Resnick, D. M. (2011, September 27). *Improving the Validity of the Medicaid/CHIP Estimates on the American Community Survey: The Role of Logical Coverage Edits*. Retrieved June 25, 2015, from U.S. Census Bureau Official Web Site: <http://www.census.gov/hhes/www/hlthins/publications/Improving%20the%20Validity%20of%20the%20Medicaid-CHIP%20Estimates%20on%20the%20ACS.pdf>

- Mach, A., & O'Hara, B. (2011). *Do people really have multiple health insurance plans? Estimates of Nongroup Health Insurance in the American Community Survey*. Retrieved June 25, 2015, from U.S. Census Bureau Official Web Site:
http://www.census.gov/hhes/www/hlthins/publications/Multiple_Coverage.pdf
- O'Hara, B. (2010). *Is there an undercount of Medicaid participants in the 2006 ACS Content Test?* Washington, D.C.: U.S. Census Bureau. Retrieved June 25, 2015, from
https://www.census.gov/acs/www/Downloads/library/2010/2010_OHara_01.pdf
- Parsons, V., Moriarity, C., Jonas, K., & al., e. (2014). Design and Estimation for the National Health Interview Survey. *Vital Health Statistics 2(165)*. National Center for Health Statistics.
- Pascale, J., Rodean, J., Leeman, J., Cosenza, C., & Schoua-Glusberg, A. (2013). Preparing to Measure Health Coverage in Federal Surveys Post-Reform: Lessons from Massachusetts. *Inquiry: The Journal of Health Care Organization, Provision, and Financing*, pp. 106-123.
- Rao, J., & Scott, A. (1987). On Simple Adjustments to Chi-Square Tests with Sample Survey Data. *The Annals of Statistics*, 15, 1, 385-397. Retrieved on January 31, 2017 from
<http://projecteuclid.org/euclid.aos/1176350273>
- SAS Institute, Inc. (2009). *SAS/STAT(R) 9.2 User's Guide*. Cary, NC: SAS Institute, Inc.
- Stapleton, M., & Steiger, D. (January 2015). *Cognitive Testing of the 2016 American Community Survey Content Test Items: Summary Report for Round 1 and Round 2 Interviews*. Rockville, Maryland: Westat.
- State Health Access Data Assistance Center. (2005). *Reclassifying Health Insurance Coverage for the Indian Health Service in the Current Population Survey: Impact on State Uninsurance Estimates*.
- Steiger, D., Anderson, J., Folz, J., Leonard, M., & Stapleton, M. (2015). *Cognitive Testing of the 2016 American Community Survey Content Test Items: Briefing Report for Round 3 Interviews*. Washington, D.C.: Westat.
- U.S. Census Bureau. (2006). *Current Population Survey Design and Methodology Technical 66*.
- U.S. Census Bureau. (2007, March 30). *Subjects Planned for the 2010 Census and American Community Survey: Federal Legislative and Program Uses*. Retrieved June 25, 2015, from U.S. Census Bureau Official Web Site:
http://www.census.gov/acs/www/Downloads/operations_admin/Final_2010_Census_and_American_Community_Survey_Subjects_Notebook.pdf
- U.S. Census Bureau. (2014). *American Community Survey Design and Methodology (January 2014)*. Retrieved June 3, 2015, from
https://www.census.gov/acs/www/Downloads/survey_methodology/acs_design_methodology_report_2014.pdf
- U.S. Census Bureau. (2016). *2015 Planning Database Tract Data [Data file]*. Retrieved from
https://www.census.gov/research/data/planning_database/2015/.
- U.S. Census Bureau. (2017). *National Population Totals Tables: 2010-2016*. Retrieved from
<https://www.census.gov/data/tables/2016/demo/popest/nation-total.html>
- United States Congress. (2010, March 23). *PUBLIC LAW 111-148, 111TH CONGRESS, An Act Entitled The Patient Protection and Affordable Care Act*. Retrieved June 25, 2015, from U.S. Government Publishing Office Official Web Site:
<http://www.gpo.gov/fdsys/pkg/PLAW-111publ148/pdf/PLAW-111publ148.pdf>

Wikipedia contributors. (2015, April 21). *Holm-Bonferroni method*. (Wikipedia, The Free Encyclopedia) Retrieved June 5, 2015, from Wikipedia.org:
http://en.wikipedia.org/w/index.php?title=Holm%E2%80%93Bonferroni_method&oldid=657461463

Appendix A: Extra Tables

Table A-1. Unit Response Rates by Designated High (HRA) and Low (LRA) Response Areas

Mode	Test Interviews	Test Percent	Control Interviews	Control Percent	Test minus Control	P-Value
Total Response	19,400		19,455			
HRA	7,556	94.3 (0.4)	7,608	94.5 (0.3)	-0.2 (0.6)	0.72
LRA	11,844	91.5 (0.3)	11,847	91.0 (0.3)	0.5 (0.5)	0.29
Difference		2.7 (0.5)		3.5 (0.5)	-0.7 (0.7)	0.33
Self-Response	13,131		13,284			
HRA	6,201	59.7 (0.7)	6,272	60.6 (0.7)	-0.9 (0.9)	0.31
LRA	6,930	33.2 (0.4)	7,012	33.6 (0.4)	-0.4 (0.6)	0.55
Difference		26.5 (0.8)		27.0 (0.8)	-0.5 (1.2)	0.66
Internet	8,168		8,112			
HRA	4,119	39.6 (0.6)	4,048	39.1 (0.6)	0.5 (0.8)	0.51
LRA	4,049	19.4 (0.3)	4,064	19.5 (0.3)	0.1 (0.4)	0.87
Difference		20.2 (0.6)		19.6 (0.7)	0.6 (0.9)	0.52
Mail	4,963		5,172			
HRA	2,082	20.0 (0.4)	2,224	21.5 (0.4)	-1.5 (0.6)	0.02*
LRA	2,881	13.8 (0.3)	2,948	14.1 (0.3)	-0.3 (0.4)	0.43
Difference		6.2 (0.5)		7.4 (0.4)	-1.1 (0.7)	0.11
CATI	872		880			
HRA	296	9.0 (0.5)	301	9.6 (0.6)	-0.6 (0.8)	0.44
LRA	576	7.9 (0.4)	579	8.0 (0.3)	-0.1 (0.5)	0.85
Difference		1.1 (0.6)		1.6 (0.7)	-0.5 (0.9)	0.58
CAPI	5,397		5,291			
HRA	1,059	82.2 (1.0)	1,035	82.7 (0.9)	-0.5 (1.3)	0.69
LRA	4,338	85.8 (0.5)	4,256	85.0 (0.4)	0.8 (0.7)	0.23
Difference		-3.7 (1.1)		-2.3 (1.0)	-1.3 (1.5)	0.36

Source: U.S. Census Bureau, 2016 American Community Survey Content Test

Note: Standard errors are in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test at the $\alpha=0.1$ level. The weighted response rates account for initial sample design as well as CAPI subsampling.

Appendix B. Dual Coverage Edit Checks

The test version of the CATI/CAPI instrument also included edit checks that were triggered if respondents reported two types of insurance coverage (see Section 1.3). As the control treatment did not have the same edit checks as the test treatment, these edit checks were not included in our decision criteria. Table 39 reports the frequencies of each of the edit checks. As observed in Table 12, 1,815 CATI responses and 9,869 CAPI responses included a “Yes” or “No” for at least one part of the Health Insurance Coverage question. Edit checks for multiple coverage types were triggered between 1 and 158 times in CATI interviews (for employer-based and Indian Health Service coverage and employer-based and Medicare coverage, respectively). They were triggered between 20 and 118 times for CAPI interviews (for the same two categories).

Table B-1. Counts of Dual-Coverage Edit Triggers and Final Responses, Test Treatment

Categories	CATI Triggered	CATI Remained “Yes”	CAPI Triggered	CAPI Remained “Yes”
Employer & Direct	50	30	59	27
Employer & Indian	1	1	20	11
Employer & Other	12	3	34	10
Employer & Medicare	158	148	118	86
Employer & Medicaid	30	24	81	45
Employer & VA	22	21	33	22
Employer & Military	10	8	38	28
Medicare & Other	29	21	47	34
Medicaid & Direct	26	16	87	45
Medicaid & Other	6	2	44	24
Direct & Other	9	8	32	7

Source: U.S. Census Bureau, 2016 American Community Survey Content Test

The second column for each interview mode lists whether both types of health insurance coverage were reported (i.e. “Yes” to both types of coverage) even after the edit check was triggered. In general, results suggested that the edit check prompted some respondents to change their coverage report, but, in most cases, most still reported the same coverage combination that triggered the check.