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

January 26, 2012

2012 AMERICAN COMMUNITY SURVEY RESEARCH AND EVALUATION REPORT MEMORANDUM SERIES #ACS12-RER-05

MEMORANDUM FOR ACS Research and Evaluation Steering Committee

From: David S. Johnson /***Signed***/
Chief, Social, Economic, and Housing Statistics Division

Prepared by: Tracy A. Loveless
Program Participation and Income Transfers Branch
Social, Economic, and Housing Statistics Division

Subject: Evaluation of the 2010 ACS Content Test Report Covering Food Stamps/SNAP

Attached is the final American Community Survey Research and Evaluation report for evaluation of the ACS Content Test Report Covering Food Stamps/SNAP. This report describes the results of changing the question wording of the food stamp question to include the new program name, SNAP (Supplemental Nutrition Assistance Program).

If you have any questions about this report, please contact Tracy Loveless at (301)763-3197 or John Hisnanick at (301)763-2295.

Attachment: (2010 ACS Content Test Evaluation Report Covering Food Stamps/SNAP)

cc:
ACS Research and Evaluation Steering Committee
Donna Daily (ASCO)
Todd Hughes
Debbie Klein
Dave Raglin
Jennifer Childs (CSM)
James Hartman (DSSD)
Jennifer Tancreto
Tony Tersine
John Hisnanick (SEHSD)
Charles Nelson

2010 ACS Content Test Evaluation Report Covering Food Stamps/SNAP

FINAL REPORT

Tracy A. Loveless
Social, Economic, and
Housing Statistics Division

United States™ Census Bureau

Intentionally Blank

TABLE OF CONTENTS

LIST OF TABLES

EXECUTIVE SUMMARY

Test Objective

In late August through mid-December 2010, the Census Bureau conducted a field test of new and revised content in the 2010 American Community Survey (ACS) Content Test. The results of that testing will determine the content to be incorporated into production ACS in 2013.

The food stamp program is now known as the Supplemental Nutritional Assistance Program (SNAP). A change in question wording is necessary to reflect the name change to ensure proper reporting of food stamp/SNAP receipt. Although states are encouraged to change their program name to SNAP, it is not required. Therefore, some states have changed their program name to SNAP, some states have chosen a different program name, and some states are still in the process of changing their program name. This variation across states adds to the complexity of data collection for this question.

Methodology

The Content Test compared two versions of the food stamp/SNAP question. The control version replicated the wording and response categories used in the current production ACS question. The test version included the following changes to the control version of the food stamp/SNAP question:

- Used the new program name, the Supplemental Nutrition Assistance Program (SNAP),
- added an instruction to exclude assistance from food banks.

Research Questions and Results

Do the changes to the food stamps question affect the estimate of households reporting receipt of food stamps?

No. There is no significant difference between the percent of households reporting receipt of food stamps in the test and control versions.

Do the changes to the food stamps question lower the item missing data rates?

No. There is no difference between the item missing data rates for the test and control versions.

Do the changes to the food stamps question improve the reliability of the data?

No. There is no difference in the gross difference rates or indexes of inconsistency between the test and control versions, suggesting that both question versions provide similar levels of data reliability. The indexes of inconsistency were low for both versions (Control 12.6 vs. Test 13.7), indicating a low inconsistency of response variability.

1. BACKGROUND

1.1 Motivation for the 2010 ACS Content Test

To evaluate proposed changes to the content of the American Community Survey (ACS), the Census Bureau conducted the 2010 ACS Content Test. The objective of the ACS Content Test, for both new and existing questions, was to determine the impact of changing question wording, response categories, and redefinition of underlying constructs on the quality of data collected.

Through the Office of Management and Budget (OMB) Interagency Committee on the ACS, subject matter experts from the Census Bureau and key data users from other federal agencies collaborated in identifying revised and new questions for inclusion in the Content Test. The suggested new and revised questions affected both the housing and detailed person sections of the ACS questionnaire.

In the housing section, the food stamps question was altered to reflect a name change for the food stamps program. In addition, a series of new questions were added related to household computer ownership and Internet subscription.

Several changes were made in the detailed person section. First, a change in data needs for the veteran series led to a revised set of response categories for the veteran status and period of military service questions. Second, the question wording of the cash public assistance income question was modified to address under-reporting of assistance on behalf of children and single payment recipients. Third, to simplify the income questions related to wages (wages, salary, commissions, bonuses, or tips) and property income (interest, dividends, rental income, royalty income or income from estates and trust), these questions were broken up into smaller questions for the Computer-Assisted Telephone Interviewing (CATI) and Computer-Assisted Personal Interviewing (CAPI) instruments only. Fourth, a set of new questions on parental place of birth were added to allow data users to divide the population into "first generation" (the foreign born), "second generation" (the children of immigrants), and "third or higher generation" (native born with no foreign-born parents).

To meet the test objective of the 2010 ACS Content Test, analysts evaluated changes to question wording, response categories, instructions, and examples relative to a control version of the question or another version for new questions. Specifically, this report discusses the food stamp question.

1.2 Previous Testing or Analysis

The 2006 ACS Content Test proposed adding the term "food stamp benefit card" and not asking for the total value of the food stamps received in the past 12 months to reduce the

under reporting of the receipt of food stamps. The results showed that changing the wording of the food stamp question, as well as removing the question for the value of the amount of food stamps received, significantly increased the proportion of households that reported receiving food stamps, nationally and in the high and low mail response strata that were used in testing. The question was changed for the 2008 ACS production based on the findings.

1.3 Recommendations from Cognitive Testing

Prior to conducting the 2010 ACS Content Test, the Research Triangle Institute (RTI), Westat, and Research Support Services (RSS) conducted cognitive interviewing, under contract, to assist in identifying a final set of questions for the field test. Two versions of each question topic were tested with the goal of choosing the best one for the revised questions and the best two for the new questions. The questions were pretested in the three modes used in the ACS data collection (paper, telephone interview, and personal interview) in English and Spanish. Cognitive interviews consisted of one-on-one interviews using the proposed questions in the context of the ACS survey. Survey methodologists also conducted respondent debriefings.

Based on the cognitive interviews conducted, the findings from the cognitive interviews do not strongly favor one version of the food stamp question over the other. The problems observed were almost evenly distributed across versions. In fact, when participants were asked which of the two versions they preferred, they were also evenly split as to whether Version 1 or 2 (see page B-1) was the better form of the question.

Many who preferred Version 2 criticized Version 1 for referring to SNAP before food stamps, since it is a name they did not recognize.

Those who preferred Version 1 often said they liked how this version makes it more clear that SNAP and food stamps are the same program. However, the problems and reactions of a number of participants who did not recognize SNAP (including many who assumed it was something other than food stamps) suggest that the question should <u>not</u> give emphasis to this new program name.

If all states have converted the food stamp benefit to an EBT (Electronic Benefit Transfer) card, cognitive testing contractors suggest that the following wording, which refers to a card and removes the reference to SNAP, would minimize misreporting. Note that a reference to WIC (Women, Infants and Children) and food banks was added to the instruction:

> <u>In the past 12 months</u>, did you or any member of this household receive a government issued food stamp card? Do NOT include WIC, the National School Lunch Program, or assistance from food banks.

If a reference to SNAP is preferred, the contractor suggested a slightly altered form of Version 2. While this version also refers to a card, it also makes it clear that SNAP and food stamps are the same thing:

> In the past 12 months, did you or any member of this household receive a government benefit card that can only be used to buy food? Include Food Stamps, now known as the Supplemental Nutrition Assistance Program (SNAP). Do NOT include WIC, the National School Lunch Program, or assistance from food banks.

Finally, if not all states have adopted EBT cards as the means of providing food stamps to beneficiaries, cognitive testing suggested removing the word "card" from the above. That is:

In the past 12 months, did you or any member of this household receive a government benefit card that can only be used to buy food? Include Food Stamps, now known as the Supplemental Nutrition Assistance Program (SNAP). Do NOT include WIC, the National School Lunch Program, or assistance from food banks.

1.4 Recommendations from the Expert Review Panel

Following the cognitive testing, an expert review panel, composed of government survey methodology experts, reviewed and added changes to the final question versions proposed to move forward from the cognitive testing into the field test. The proposed changes for each question topic were approved by the corresponding OMB interagency subcommittee responsible for initiating the research. The OMB provided final approval of the proposed changes.

The expert review panel recommended the following wording change to the question:

IN THE PAST 12 MONTHS, did you or any member of this household receive benefits from the Food Stamp Program or SNAP (the Supplemental Nutrition Assistance Program)? Do NOT include WIC or the School Lunch Program.

Yes
No

2. SELECTION CRITERIA

Before fielding the 2010 ACS Content Test, we identified the following criteria to determine which version of the question should move forward based on the results of the test.

The number of households reporting receipt of food stamps in the test version should be about the same as in the control version.

The item nonresponse rates and reliability measures will be considered together when determining which question version performs better.

3. METHODOLOGY

3.1 Data Collection Methods

The initial stages of the Content Test consisted of content determination, cognitive laboratory pretesting, and expert reviews for the purpose of developing alternate versions of question content. The field test portion of the ACS Content Test used the data collection methodology currently used in the production ACS (i.e., mail questionnaire, follow-up CATI, and follow-up CAPI) with an added reinterview conducted via a CATI instrument known as the Content Follow-Up (CFU). Additional data were collected on respondent and interviewer behavior during the field test via Computer Audio Recorded Interviewing (CARI) technologies for a subset of respondents during the CATI and CAPI follow-up modes of data collection.

The Content Test followed the same schedule and procedures for the mail, CATI, and CAPI operations as the September 2010 ACS production panel. Questionnaires were mailed to sampled households at the end of August 2010. The Content Test used an English-only mail form but the automated instruments (CATI, CAPI, and CFU) included both English and Spanish versions. Households not responding by mail and for which we had a phone number were contacted for a CATI interview during the month of October 2010. In November 2010, Census Bureau field representatives visited a sample of households that did not respond by mail or CATI to attempt a CAPI interview. The CAPI operations ended December 2, 2010.

The field test included a CATI CFU reinterview to collect additional measures for the study of response error. This operation started approximately two weeks after the initial mail out of questionnaires and ended two weeks after the end of the CAPI follow-up data collection operation. The CFU included all occupied households for which we received a response in the original interview and had a telephone number. A response was defined as a case where the household provided data through at least the first person's place of birth question for mail cases or at least a sufficient partial interview for CATI/CAPI interviews. The reinterview was conducted about 2 to 4 weeks after the original interview and with the original respondent when possible. Note that the CFU CATI interview was an abbreviated version of the original Content Test interview. The CFU instrument included the basic demographic section and only those questions preceding the questions being tested in the housing and the detailed person sections to provide context (see Appendix D for the flow of the CFU instrument).

The ACS Content Test did not include all of the production data collection operations and processes. First, while the Telephone Questionnaire Assistance program's toll-free number was available to Content Test respondents for assistance, the CATI instrument did not include content changes from the Content Test. Therefore, data collected from Content Test respondents via TQA CATI interview were not included in our analysis.

Second, since our objective was to study response error using unedited data, the Content Test excluded the Failed Edit Follow-up (FEFU) CATI operation and the edit and imputation data processes.

3.2 Sample Design

The 2010 Content Test consisted of a national sample of 70,000 residential addresses in the contiguous United States (the sample universe did not include Puerto Rico, Alaska, and Hawaii). The sample design for the Content Test was largely based on the ACS production sample design with some modifications to meet the test objectives. The modifications included adding an additional level of stratification by stratifying addresses into high and low mail response areas, over-sampling addresses from the low mail response areas to ensure equal response from both strata, and sampling units as pairs. The high and low mail response strata were defined based on ACS mail response rates at the tract-level. The paired sample selection formed pairs by first systematically sampling an address within the defined sampling strata and then pairing that address with the address listed next in the geographically sorted list. However, the pair was not likely comprised of neighboring addresses. One member of the pair was randomly assigned to the control group and the other member was assigned to the test group. Those addresses assigned to the test group received the revised ACS questions and the questions new to the ACS. The control group received the current questions on the production ACS as well as different versions of the new questions.

Another modification to the production ACS sample design included adding a third sampling stage. At the first stage, the production 2010 ACS first stage sample was used as the Content Test first stage sample. At the second stage, all housing units in the ACS first stage sample not selected in the production 2010 ACS second-stage sample were selected as the Content Test second-stage sample. In addition, any units that were selected to be in other operations (e.g., training, other tests, etc.) were not selected in the Content Test second stage sample. At the third stage, addresses were selected using a sampling method similar to the production ACS second stage sample design with the exception of adding the high and low mail response stratification.

3.3 Methodology Specific to the Food Stamps Question

The 2010 Content Test compared two versions of the food stamp question. The control version replicated the wording and response categories used in the current production ACS question.

The control version asked…IN THE PAST 12 MONTHS, did anyone in this household receive Food Stamps or a Food Stamp benefit card? *Include government benefits from the Supplemental Nutrition Assistance Program (SNAP). Do NOT include WIC or the national School Lunch program.*

The test version revises both the question and italicized instruction… IN THE PAST 12 MONTHS, did you or any member of this household receive benefits from the Food

Stamp Program or SNAP (Supplemental Nutrition Assistance Program)? *Do NOT include WIC, the School Lunch Program, or assistance from food banks.*

4. LIMITATIONS

Control and test CATI-CAPI workload assignments were not assigned using an interpenetrated experimental design. That is, interviewers were allowed to administer interviews for both control and test cases, in addition to production ACS cases. The potential risk of this approach is the introduction of a cross-contamination or carry-over effect due to the 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.

The CFU reinterview was not conducted in the same mode of data collection for households that responded by mail or 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.

Respondents needed to provide a telephone number in the original Content Test interview or the Census Bureau had to be able to find a telephone number for that unit through reverse address look-up to be included in the CFU interview. As a result, 18.4 percent of the responding households from the original interview were not eligible for the CFU reinterview.

We did not have the same respondent in the CFU that we had in the original interview for 9.1 percent of the CFU cases. This means that differences between the original interview and the CFU for these cases could be due in part to having different people answering the questions.

The Content Test does not include the production weighting adjustments for seasonal variations in ACS response patterns, nonresponse bias, and under-coverage bias. The CFU portion of the Content Test did include a unit nonresponse adjustment for those Content Test cases that responded to the Content Test, but failed to respond to the CFU. As a result, the statistics derived from the Content Test data do not provide the same level of inference as the production ACS to the entire population of housing units and persons in the contiguous United States.

5. RESEARCH QUESTIONS AND RESULTS

5.1 Response to the Content Test and Content Follow-Up

Table 1 shows the unit response rates for each of the modes of data collection and all modes combined (excluding CFU) by the control and test groups. The comparison between control and test shows that respondent participation was similar for both control and test for each of the modes of data collection and all modes combined, with the exception of the CATI mode. The test treatment produces a CATI rate of response that is 3 percentage points higher compared to that of the control. We are not able to explain the increase in response due to the test treatment for the CATI mode of data collection other than by random occurrence given that the conditions affecting unit response were equivalent between the test and control groups.

Table 1. Content Test Response Rate Comparisons Between the Control and Test Treatments

Mode	Test (%)	Standard Error (%)	Control (%)	Standard Error (%)	Test - Control (%)	Standard Error (%)	Significant
All Modes (CFU excluded)	95.4	0.2	95.7	0.2	-0.3	0.3	No
Mail	58.1	0.5	57.7	0.5	0.5	0.7	No
CATI	52.6	1.2	49.6	1.0	3.0	1.5	Yes
CAPI	90.4	0.5	91.5	0.5	-1.1	0.7	No
CFU	54.3	0.5	53.5	0.6	0.8	0.7	No

Source: U.S. Census Bureau, 2010 American Community Survey Content Test

5.2 Do the changes to the food stamps question affect the estimate of households reporting receipt of food stamps?

We compared the estimated percent of households reporting receipt of food stamps between the control and the test versions. Statistical significance between versions was determined using a t-test. An estimate from the test version that is the same as or higher than the estimate from the control version is acceptable, according to our criteria.

There was no significant difference between the percent of households reporting receipt of food stamps in the test and control versions.

Table 2. Difference in Receipt of Food Stamps between Test and Control

	Test	Standard Error (%)	Control	Standard Error (%)	Test – Control (%)	Standard Error (%)	Significant
Receipt of Food Stamps	12.1	0.3	12.5	0.3	-0.4	0.4	No

Source: U.S. Census Bureau, 2010 American Community Survey Content Test, September to December 2010

Note: Statistical significance of differences is determined at the $\alpha = 0.10$ significance level using a one-sided test.

5.3 Do the changes to the food stamps question lower the item missing data rates?

We compared the item missing data rates between the control and the test versions. Statistical significance between versions were determined using a t-test. An item missing data rate for the test version that is the same as or lower than the item missing data rate for the control version is acceptable.

The item missing data rate is the percent of eligible households that did not provide an answer to the food stamps question. All occupied households in the Content Test are eligible to answer this question. We used the following formula to calculate the item missing data rates.

$$\text{Item Missing Data Rate} = \frac{\text{\# of households that did not provide a response to the food stamps question}}{\text{Total \# of nonblank household records in the Content Test}} * 100$$

There was no difference between the item missing data rates for the test and control versions.

Table 3. Difference in Item Nonresponse Rates between Test and Control

	Test	Standard Error (%)	Control	Standard Error (%)	Test – Control (%)	Standard Error (%)	Significant
Item Nonresponse	2.5	0.1	2.5	0.1	-0.0	0.2	No

Source: U.S. Census Bureau, 2010 American Community Survey Content Test, September to December 2010

Note: Statistical significance of differences is determined at the $\alpha = 0.10$ significance level using a one-sided test.

5.4 Do the changes to the food stamps question improve the reliability of the data?

Using data from the Content Test and CFU, we answered this question by comparing the simple response variance, as measured by Gross Difference Rates (GDRs), and the index of inconsistency between the control and the test versions. For these calculations we only considered households that responded to the food stamps question in both the original interview (via mail, CATI, or CAPI) and the CFU interview.

Simple response variance measures the average variability, across respondents, between the responses to the food stamps question in the original interview and in the CFU. The GDR measures the gross rate of disagreement between the responses to the same question in the original interview and the reinterview. For example, for the food stamps question, disagreement occurs when a respondent answers "Yes" in the original interview and "No" during CFU, or "No" in the original interview and "Yes" during CFU. We used the following formula to calculate the GDRs.

$$\text{GDR} = \frac{\text{\# of households that provided a different response to the food stamps question in CFU compared to the original interview}}{\text{\# of households that responded to the food stamps question in both the original interview and CFU}}$$

The index of inconsistency is the percentage of the variance that is due to simple response variance for the given response category. It provides an estimate of the magnitude of response variability for a given item. 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. We used the following information and formula to calculate the index of inconsistency.

$$\text{Index of Inconsistency} = \frac{\text{\# of households that provided a different response in CFU compared to the original interview}}{\frac{1}{n}[(A \times D) + (B \times C)]}$$

Where
A = Total # of households that responded "Yes" in the original interview
B = Total # of households that responded "No" in the original interview
C = Total # of households that responded "Yes" in CFU
D = Total # of households that responded "No" in CFU
n = Total # of households that responded to both the original interview and CFU

There was no difference in the gross difference rate or index of inconsistency between the test and control versions, suggesting that they provide similar levels of data reliability. The index of inconsistency was low for both versions (Control 12.6 vs. Test 13.7).

Table 4. Difference in Reliability between Test and Control

	Test	Standard Error (%)	Control	Standard Error (%)	Test – Control (%)	Standard Error (%)	Significant
Gross Difference Rates	2.7	0.3	2.7	0.2	0.0	0.4	No

Source: U.S. Census Bureau, 2010 American Community Survey Content Test, September to December 2010
Note: Statistical significance of differences is determined at the $\alpha = 0.10$ significance level using a one-sided test.

5.5 For each mode of data collection, do the changes to the food stamps question affect the estimate of recipiency, item missing data rate, or reliability of the data? (informational purposes only)

We answered this question by comparing the estimates of food stamps recipiency, item missing data rates, and reliability measures as defined above between the control and the test versions for each mode (mail, CATI, CAPI). Statistical significance between versions was determined using a t-test.

Note that comparisons across modes of data collection could not be made since measurable differences cannot be attributed strictly to the mode of data collection. Observed differences across modes may also be due to mode specific respondent characteristics and reinterview mode effects (CFU was conducted by telephone only). Specifically, respondents self-select into a mode, such that the mail universe has different characteristics from the CATI and CAPI universes.

There was no difference in the estimate of recipiency, item missing data rate, or reliability of the data. See Tables A1-A4 in Appendix A.

5.6 For each mail response stratum, do the changes to the food stamps question affect the estimate of recipiency, item missing data rate, or reliability of the data? (informational purposes only)

We answered this question by comparing the estimates of food stamps recipiency, item missing data rates, and reliability measures as defined above between the control and the test versions for each mail response stratum (high and low). Statistical significance between versions was determined using a t-test.

The differences in the estimate of recipiency, item missing data rate, or reliability of the data were not statistically different in the high mail response stratum. However, for the low mail response stratum, the difference in the gross difference rates and difference in Index of Inconsistency were statistically different. See Tables A5-A8 in Appendix A.

5.7 Does either question version elicit respondent or interviewer behaviors that may contribute to interviewer or respondent error? (informational purposes only)

We answered this question by comparing the behavior coding results derived from the CARI recordings between the control and the test versions.

For respondent behavior, the test and control performed similarly. Interviewers were reading the test version of the question as worded less frequently than the control. Analysis of the behavior coder notes revealed that interviewers were truncating the test question version as well as dropping the term “SNAP.” The test version dramatically reduced the rate of standard interviewer behavior compared to the control version (73% for control vs. 34% for test).

5.8 For the Hispanic and Black population subgroups, do the changes to the food stamps question affect the estimate of recipiency, item missing data rate, or reliability of the data? (informational purposes only)

We answered this question by comparing the estimates of food stamps recipiency, item missing data rates, and reliability measures as defined above between the control and the test versions for the Hispanic and Black subgroups separately.

Note that this test was not designed to study differences across panels by race/ethnicity breakdowns with statistical precision, as this was not a stated goal of the test. Therefore, these results are provided for informational purposes only.

The differences in the estimate of recipiency, item missing data rate, or reliability of the data in the Black population subgroup were not statistically significant. For the Hispanic population subgroup, the difference in receipt of food stamps between test and control and differences in gross difference rates between test and control were statistically significant. See Tables A9-A12 in Appendix A.

6. SUMMARY

The content test results indicate that changing the wording of the food stamp question to include SNAP showed no impact to the item missing data rate and reliability of the data. There was no difference between the percent of households overall reporting receipt of food stamps in the test and control versions. Results suggested no difference between the test version and the control version for item nonresponse and reliability. Therefore, the recommendation is to use the test version of the question.

Acknowledgements

I would like to thank the following Census Bureau staff for their valuable contributions and assistance to the development and analysis of this project: Mary Frances Zelenak and Michelle Ruiter, Decennial Statistical Studies Division, and Joanne Pascale, Center for Survey Measurement.

References

Hisnanick, J., Loveless, T.,Chesnut, J. (2007) "Evaluation Report Covering receipt of Food Stamps, U.S. Census Bureau, Washington.

Pascale J. and Goerman P. (2010) "ACS 2010 Content Test Behavior Coding Report" U.S. Census Bureau, Washington.

RTI International (2009) "Cognitive Testing of the American Community Survey Content Test Items" RTI International, Research Triangle Park.

United States Department of Agriculture, Food and Nutrition Service http://www.fns.usda.gov/snap/snap.htm

Appendix A: Tables

Table A1. Difference in Receipt of Food Stamps between Test and Control by Mode

Mode	Test	Standard Error (%)	Control	Standard Error (%)	Test – Control (%)	Standard Error (%)	Significant
Mail	8.30	0.26	8.52	0.29	0.22	0.35	No
CATI	10.66	0.66	12.01	0.75	-1.35	0.98	No
CAPI	18.58	0.75	18.94	0.64	-0.36	0.94	No

Source: U.S. Census Bureau, 2010 American Community Survey Content Test, September to December 2010

Note: Statistical significance of differences is determined at the $\alpha = 0.10$ significance level using a two-sided test.

Table A2. Difference in Item Missing Data Rates between Test and Control by Mode

Mode	Test	Standard Error (%)	Control	Standard Error (%)	Test – Control (%)	Standard Error (%)	Significant
Mail	3.3	0.2	3.2	0.2	0.0	0.3	No
CATI	0.2	0.1	0.2	0.1	-0.1	0.2	No
CAPI	1.6	0.3	1.8	0.3	-0.2	0.4	No

Source: U.S. Census Bureau, 2010 American Community Survey Content Test, September to December 2010

Note: Statistical significance of differences is determined at the $\alpha = 0.10$ significance level using a two-sided test.

Table A3. Difference in Gross Difference Rates between Test and Control by Mode

Mode	Test	Standard Error (%)	Control	Standard Error (%)	Test – Control (%)	Standard Error (%)	Significant
Mail	1.7	0.2	1.5	0.2	0.2	0.3	No
CATI	2.4	0.5	3.0	0.5	-0.6	0.7	No
CAPI	4.4	0.6	4.4	0.7	-0.0	0.9	No

Source: U.S. Census Bureau, 2010 American Community Survey Content Test, September to December 2010

Note: Statistical significance of differences is determined at the $\alpha = 0.10$ significance level using a two-sided test.

Table A4. Difference in Index of Inconsistency between Test and Control by Mode

Mode	Test	Standard Error (%)	Control	Standard Error (%)	Test – Control (%)	Standard Error (%)	Significant
Mail	12.8	1.6	11.1	1.3	1.6	2.2	No
CATI	12.3	2.5	14.3	2.3	-2.0	3.3	No
CAPI	15.2	2.4	14.1	2.0	1.1	3.1	No

Source: U.S. Census Bureau, 2010 American Community Survey Content Test, September to December 2010

Note: Statistical significance of differences is determined at the $\alpha = 0.10$ significance level using a two-sided test.

Table A5. Difference in Receipt of Food Stamps between Test and Control by Mail Response Stratum

Stratum	Test	Standard Error (%)	Control	Standard Error (%)	Test – Control (%)	Standard Error (%)	Significant
High	8.94	0.39	9.41	0.37	-0.47	0.52	No
Low	21.52	0.42	21.67	0.40	-0.15	0.53	No

Source: U.S. Census Bureau, 2010 American Community Survey Content Test, September to December 2010

Note: Statistical significance of differences is determined at the $\alpha = 0.10$ significance level using a two-sided test.

Table A6. Difference in Item Missing Data Rates between Test and Control by Mail Response Stratum

Stratum	Test	Standard Error (%)	Control	Standard Error (%)	Test – Control (%)	Standard Error (%)	Significant
High	2.3	0.2	2.4	0.2	-0.1	0.3	No
Low	2.8	0.2	2.8	0.2	0.1	0.2	No

Source: U.S. Census Bureau, 2010 American Community Survey Content Test, September to December 2010

Note: Statistical significance of differences is determined at the $\alpha = 0.10$ significance level using a two-sided test.

Table A7. Difference in Gross Difference Rates between Test and Control by Mail Response Stratum

Stratum	Test	Standard Error (%)	Control	Standard Error (%)	Test – Control (%)	Standard Error (%)	Significant
High	1.8	0.29	2.1	0.28	-0.4	0.42	No
Low	5.6	0.4	4.4	0.4	1.2	0.6	Yes

Source: U.S. Census Bureau, 2010 American Community Survey Content Test, September to December 2010

Note: Statistical significance of differences is determined at the $\alpha = 0.10$ significance level using a two-sided test.

Table A8. Difference in Index of Inconsistency between Test and Control by Mail Response Stratum

Stratum	Test	Standard Error (%)	Control	Standard Error (%)	Test – Control (%)	Standard Error (%)	Significant
High	12.2	2.1	12.6	1.7	-0.4	2.6	No
Low	16.7	1.2	13.3	1.0	3.4	1.8	Yes

Source: U.S. Census Bureau, 2010 American Community Survey Content Test, September to December 2010

Note: Statistical significance of differences is determined at the $\alpha = 0.10$ significance level using a two-sided test.

Table A9. Difference in Receipt of Food Stamps between Test and Control by Hispanic and Black subgroups

Mode	Test	Standard Error (%)	Control	Standard Error (%)	Test – Control (%)	Standard Error (%)	Significant
Hispanic	21.96	0.98	18.62	0.89	3.34	1.25	Yes
Black	26.52	0.84	28.66	1.25	-2.14	1.58	No

Source: U.S. Census Bureau, 2010 American Community Survey Content Test, September to December 2010

Note: Statistical significance of differences is determined at the $\alpha = 0.10$ significance level using a two-sided test.

Table A10. Difference in Item Missing Data Rates between Test and Control by Hispanic and Black subgroups

Mode	Test	Standard Error (%)	Control	Standard Error (%)	Test – Control (%)	Standard Error (%)	Significant
Hispanic	2.1	0.2	1.8	0.2	0.2	0.3	No
Black	3.5	0.4	3.0	0.3	0.5	0.6	No

Source: U.S. Census Bureau, 2010 American Community Survey Content Test, September to December 2010

Note: Statistical significance of differences is determined at the $\alpha = 0.10$ significance level using a two-sided test.

Table A11. Difference in Gross Difference Rates between Test and Control by Hispanic and Black subgroups

Mode	Test	Standard Error (%)	Control	Standard Error (%)	Test – Control (%)	Standard Error (%)	Significant
Hispanic	6.9	0.9	4.9	0.7	2.1	1.2	Yes
Black	5.2	0.8	5.9	1.2	-0.7	1.5	No

Source: U.S. Census Bureau, 2010 American Community Survey Content Test, September to December 2010

Note: Statistical significance of differences is determined at the $\alpha = 0.10$ significance level using a two-sided test.

Table A12. Difference in Index of Inconsistency between Test and Control by Hispanic and Black subgroups

Mode	Test	Standard Error (%)	Control	Standard Error (%)	Test – Control (%)	Standard Error (%)	Significant
Hispanic	20.8	2.8	15.6	2.2	5.1	3.8	No
Black	13.5	2.1	13.8	2.9	-0.3	3.6	No

Source: U.S. Census Bureau, 2010 American Community Survey Content Test, September to December 2010

Note: Statistical significance of differences is determined at the $\alpha = 0.10$ significance level using a two-sided test.

Appendix B: Images of the Mail Versions of the Control and Test Questions

Figure B-1. Control Version of the food stamp/SNAP question:

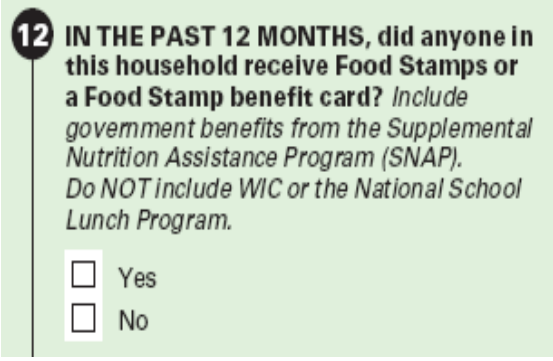
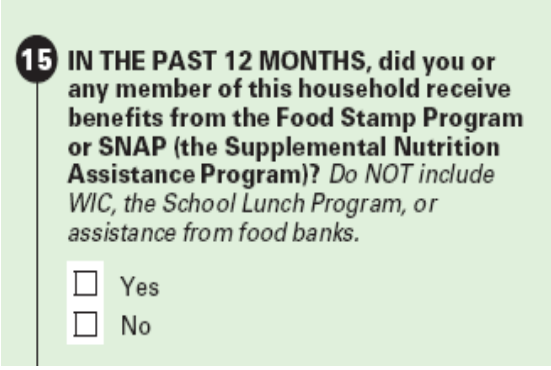


Figure B-2. Test Version of the food stamp/SNAP question:



Appendix C: CATI and CAPI Versions of the Control and Test Questions

Control Version

IN THE PAST 12 MONTHS, did anyone in this household receive Food Stamps or a Food Stamp benefit card? In some states the Food Stamps program may be known as the Supplemental Nutrition Assistance Program (SNAP).

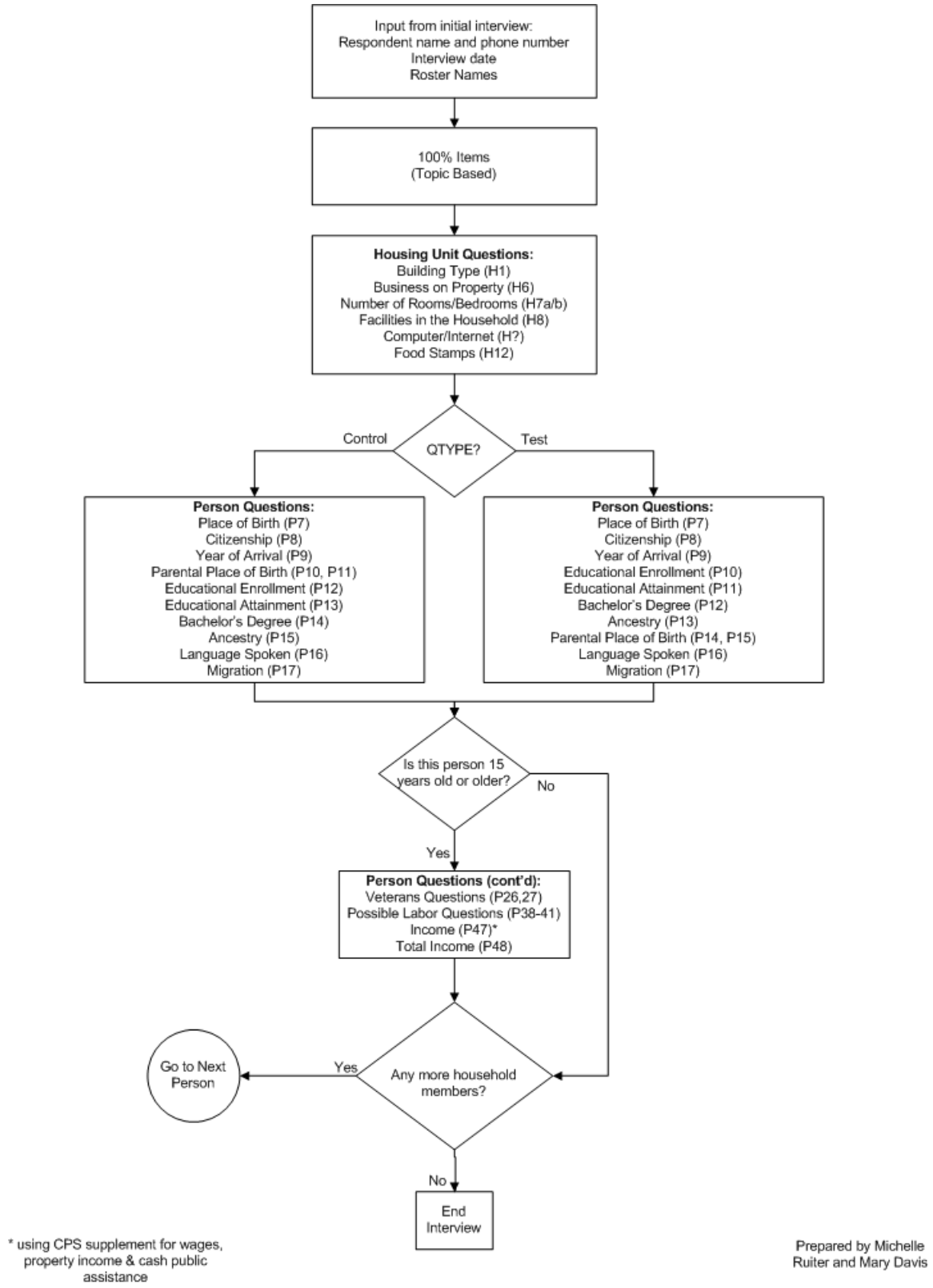
- o 1. Yes
- o 2. No

Test Version

IN THE PAST 12 MONTHS, did you or any member of this household receive benefits from the Food Stamp Program or SNAP, the Supplemental Nutrition Assistance Program? Do NOT include WIC, the School Lunch Program, or assistance from food banks.

- o 1. Yes
- o 2. No

Appendix D: Flow of the Content Follow-Up



Appendix E: Information Page

Test Design

Treatments	Two question versions with different wording (see page 3).
Sample Size	35,000 households per treatment (70,000 total)
Sample Design	Similar to production ACS with an additional level of stratification into high and low mail response areas.
Modes	Mail, CATI, and CAPI, with a CATI content follow-up (CFU) of all households. *CATI and CAPI interviews will be recorded using Computer-Assisted Recorded Interviewing (CARI) technology.*
Time Frame	Same schedule as the production September panel: mailout in late August, CATI in October, CAPI in November. CFU goes from mid-September to mid-December.

Research Questions & Evaluation Measures

No.	Research Questions	Evaluation Measures
1	Do the changes to the food stamps question affect the estimate of households reporting receipt of food stamps?	Compare the estimate of households reporting receipt of food stamps between the control and the test versions.
2	Do the changes to the food stamps question lower the item missing data rates?	Compare the item missing data rates between the control and the test versions.
3	Do the changes to the food stamps question improve the reliability of the data?	Using data from the Content Test and the Content Follow-up (CFU), compare the simple response variance and the index of inconsistency between the control and the test versions.
4	For each mode of data collection, do the changes to the food stamps question affect the estimate of recipiency, item missing data rate, or reliability of the data?	For each mode (mail,CATI,CAPI), compare the item missing data rates, estimates of food stamps recipiency, and reliability measures between the control and the test versions. *Comparisons across modes of data collection cannot be made since measurable differences cannot be attributed strictly to the mode of data collection. Observed differences across modes may also be due to mode specific respondent characteristics and reinterview mode effects (CFU only).*

No.	Research Questions	Evaluation Measures
5	For each mail response stratum, do the changes to the food stamps question affect the estimate of recipiency, item missing data rate, or reliability of the data?	For each mail response stratum (high and low), compare the item missing data rates, estimates of food stamps recipiency, and reliability measures between the control and the test versions.
6	Does either question version elicit respondent or interviewer behaviors that may contribute to interviewer or respondent error?	Compare the behavior coding results derived from the CARI recordings between the control and the test versions.
7	For the Hispanic and Black population subgroups, do the changes to the food stamps question affect the estimate of recipiency, item missing data rate, or reliability of the data?	For the Hispanic and Black subgroups separately, compare the item missing data rates, estimates of food stamps recipiency, and reliability measures between the control and the test versions. *Note: This test was not designed to study differences across panels by race/ethnicity breakdowns with statistical precision, as this was not a stated goal of the test. Therefore, these results will be provided for informational purposes only.*

Selection Criteria (In order of priority)

Research Question(s)	Criteria
1	The number of households reporting receipt of food stamps in the test version should be about the same as in the control version.
2, 3	The item nonresponse rates and reliability measures will be considered together when determining which question version performs better.

Supplemental Information

Research Question(s)	Criteria
4-7	Not part of the selection criteria. These data are presented to give additional information regarding how the questions performed.

Question Wording

Current ACS Wording	Content Test Wording
Q.15 IN THE PAST 12 MONTHS, did anyone in this household receive Food Stamps or a Food Stamp benefit card? *Include government benefits from the Supplemental Nutrition Assistance Program (SNAP). Do NOT include WIC or the National School Lunch Program.* ☐ Yes ☐ No	**Q.15 IN THE PAST 12 MONTHS, did you or any member of this household receive benefits from the Food Stamp Program or SNAP (the Supplemental Nutrition Assistance Program)?** *Do NOT include WIC, the School Lunch Program, or assistance from food banks.* ☐ Yes ☐ No