



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-10

MEMORANDUM FOR ACS Research and Evaluation Steering Committee

From: Charles Nelson /*Signed*/
Chief
Social, Economic and Housing Statistics Division

Prepared by: Edward J. Welniak Jr., Amanda R. Noss and Kirby G. Posey
Income Statistics Branch
Social, Economic, and Housing Statistics Division

Subject: 2010 ACS Content Test Evaluation Report Covering Wages

Attached is the final 2010 ACS Content Test Evaluation Report Covering Wages. This report describes the results of the 2010 Content Test with a change to the wages and salary question. Final results indicate the test question wording will be implemented for the 2013 ACS. If you have any questions about this report, please contact Edward Welniak at (301)763-5533, Amanda Noss at (301)763-6675, or Kirby Posey at (301)763-5548.

Attachment: (2010 ACS Content Test Evaluation Report Covering Wages)

cc:
Todd Hughes (ACSO)
Debra Klein
Donna Daily
Jennifer Childs
Jennifer Tancreto (DSSD)
Tony Tersine
Mary Davis
Charles Nelson (SEHSD)

2010 ACS Content Test Evaluation Report Covering Wages

FINAL REPORT

Edward J. Welniak Jr.

Amanda R. Noss

Kirby G. Posey

Social, Economic, and
Housing Statistics Division



Intentionally Blank

TABLE OF CONTENTS

EXECUTIVE SUMMARY	iv
1. BACKGROUND	1
1.1 Motivation for the 2010 Content Test.....	1
1.2 Previous Testing or Analysis	1
1.3 Recommendations from Cognitive Testing	2
1.4 Recommendations from the Expert Review Panel	3
2. SELECTION CRITERIA	3
3. METHODOLOGY	3
3.1 Data Collection Methods	3
3.2 Sample Design	4
3.3 Methodology Specific to Wages.....	5
4. LIMITATIONS.....	5
5. RESEARCH QUESTIONS AND RESULTS	6
5.1 Response to the Content Test and Content Follow-Up	6
5.2 Is the response distribution of wages income comparable to the Current Population Survey’s Annual Social and Economic Supplement (ASEC) distribution of wages income?.....	7
5.3 Do the changes to the wages question raise the proportion of persons receiving wages income?	8
5.4 Do the changes to the wages question raise the estimate of wages income?.....	8
5.5 Do the changes to the wages question affect the response distribution, shifting the lower wage categories of the distribution higher?	8
5.6 Do the changes to the wages question result in the same or lower item missing data rates?.....	9
5.7 Do the changes to the wages question lower response error (i.e., bias) in the estimate of wages reciprocity and wages income?	10
5.8 Do the changes to the wages questions lower the estimate of poverty rate?	11
5.9 For each mode of data collection, do the changes to the wages question affect the item missing data rates, the estimates of reciprocity and wages income, or response error (i.e., bias)?	12

5.10 For each mail response stratum, do the changes to the wages question affect the item missing data rates, the estimates of reciprocity and wages income, or response error (i.e., bias)?	14
5.11 Does either question version elicit respondent or interviewer behaviors that may contribute to interviewer or respondent error?.....	15
6. SUMMARY	16
References	16
Acknowledgements	16
Appendix A: Tables	A-1
Appendix B: CATI and CAPI Versions of the Control and Test Questions	B-1
Appendix C: Flow of the Content Follow-Up Interview	C-1
Appendix D: Information Page	D-1
Appendix E: CFU Wording	E-1

LIST OF TABLES

Table 1. Content Test Response Rate Comparisons Between Control and Test Treatments	7
Table 2. Response Distribution CPS/ASEC, Control, and Test	7
Table 3. Reciprocity Rates, Control versus Test	8
Table 4. Mean and Median Estimates of Wages Income.....	8
Table 5. Shift in Distribution	9
Table 6. Item Missing Data Rates	10
Table 7. Net Difference Rates.....	11
Table 8. Poverty Estimates	12
Table 9a. Item Missing Data Rates for Reciprocity	12
Table 9b. Item Missing Data Rates for Amount	13
Table 9c. Reciprocity by Mode	13
Table 9d. Net Difference Rates for Reciprocity by Mode.....	14
Table 9e. Net Difference Rates for Amounts (Mail)	14
Table 10a. High Response Stratum- Net Difference Rate	15
Table 10b. Low Response Stratum- Net Difference Rate.....	15
Appendix Tables	
Table A-1. Net Difference Rates for Amounts (CATI/CAPI).....	A-1
Table A-2. Net Difference Rates for Amounts (CATI)	A-1
Table A-3. New Difference Rates for Amounts (CAPI).....	A-1
Table A-4. Reciprocity by Response Stratum	A-2
Table A-5. High Response Stratum Item Missing Rate.....	A-2
Table A-6. Low Response Stratum Item Missing Rate	A-2

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 help determine the content to be incorporated into production ACS in 2013 or 2014.

Research shows that respondents have difficulties remembering all the information read to them in a single, verbose question (Webster, 2006). If a question contains a long list of components or concepts for respondents to consider, respondents tend to focus on the last items in the list and forget the others when the list is presented orally. In the case of the ACS Computer Assisted Telephone Interview (CATI) and Computer Assisted Personal Interview (CAPI) wage and salary reciprocity question, respondents are asked if they received wages, salary, commissions, bonuses or tips. We believe respondents are focusing on reporting whether they received bonuses or tips and missing the reporting of wages and salary. We have anecdotal evidence to support this belief. While observing ACS interviews we noted that respondents report having a wage/salary job but report having no “wages, salary, tips, bonuses, or commissions” from that job. Therefore, changes were made to the CATI/CAPI questions only for this test. However, the analysis studies the impact on the full item since we do not publish ACS data by mode.

Methodology

The Content Test compared two versions of wages. 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 revised wages questions.

The control version asked...

“The next few questions are about income DURING THE PAST 12 MONTHS...
Did [<Name>/you] receive any wages, salary, tips, bonuses or commissions?”

[if yes] “What was the amount?”

The test version asked two separate questions however still keeping all components...

“The next few questions are about income DURING THE PAST 12 MONTHS, that is from <DATE> to <DATE>... “

“Did [<Name>/you] receive any wages or salary?”

<1> Yes

<2> No

“Did [<Name>/you] receive any tips, bonuses or commissions DURING THE PAST 12 MONTHS?”

<1> Yes

<2> No

Research Questions and Results

Is the response distribution of wages income comparable to the Current Population Survey’s Annual Social and Economic Supplement (ASEC) distribution of wages income?

Yes. The overall distribution of wages income for the test version is comparable to that of the CPS ASEC. However, formal comparisons were not made since the Content Test data were not edited or imputed, adjusted for nonresponse, nor raked to known population totals.

Do the changes to the wages question raise the proportion of persons receiving wages income?

No, the changes to the wages question do not raise the proportion of persons receiving wages income.

Do the changes to the wages question raise the estimate of wages income?

There are mixed results: the test version median estimate of wages income is significantly higher than the control version, but the test version mean estimate is not significantly higher than the control version mean.

Do the changes to the wages question affect the response distribution, shifting the lower wage categories of the distribution higher?

No, the changes to the wages question do not significantly affect the response distribution.

Do the changes to the wages question result in the same or lower item missing data rates?

No, the item missing data rates for both wages reciprocity and wages amount are significantly higher in the test version than the control. This is partially due to having a two part question.

Do the changes to the wages question lower response error (i.e., bias) in the estimate of wages reciprocity and wages income?

Yes, several wage categories had lower response error for the test version of amount than the control: Don’t Know/Refusals, \$3-\$2,499 and \$2,500 to \$19,999. But the response error for wage reciprocity for the test version was not significantly lower than the control.

Do the changes to the wages questions lower the estimate of poverty rate?

No. Changes to the wages question (in conjunction with the changes to the other income questions, Property Income and Public Assistance) do not significantly lower the estimate of the poverty rate

For each mode of data collection, do the changes to the wages question affect the item missing data rates, the estimates of reciprocity and wages income, or response error (i.e., bias)?

There are mixed results. For the CATI mode, wages reciprocity is significantly higher for the test version. For CATI/CAPI combined and CAPI, the absolute value of the net difference rates for wages reciprocity is significantly lower for the test version. The net difference rate is significantly lower for mail amount category \$40,000 - \$64,999 in the test version. Item missing data rates for wages reciprocity are significantly lower for mail for the test version but significantly higher for CATI/CAPI, CATI, and CAPI for the test version. Additionally, item missing data rates for wages amount are significantly higher for CATI/CAPI and CAPI for the test version.

For each mail response stratum, do the changes to the wages question affect the item missing data rates, the estimates of reciprocity and wages income, or response error (i.e., bias)?

For the high response stratum, absolute values of net difference rates for wages reciprocity, and wages amount for the Don't Know/Refusal category, are significantly lower for the test version. For the low response stratum, the item missing data rate for wages amount is significantly higher for the test version; but the absolute values of net difference rates are significantly lower for the test version in the Don't Know/Refusal and \$2,500 to \$39,999 categories. There are no other significant findings by mail response stratum.

Does either question version elicit respondent or interviewer behaviors that may contribute to interviewer or respondent error?

Results indicate that for the series as a whole the test performs better on interviewer behavior. For respondent behavior, the difference between the test and control series is not significant.

Recommendation

Health and Human Services (HHS), the sponsor for the change to the wages question, suggested to proceed with this change for 2013. There were several positive and negative results, however there were more positive results. The results are further discussed below. The goal was to capture more households with wages income and this was achieved with the test version.

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's 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 changes to the wages questions.

1.2 Previous Testing or Analysis

It was believed that respondents are focusing on reporting whether they received bonuses, tips, or commissions and missing the reporting of wages and salary. Often, if a question contains a long list of components or concepts for respondents to consider, respondents tend to focus on the last items in the list and forget the others when the list is presented

orally. We have anecdotal evidence to support this belief. While observing ACS interviews we noted that respondents report having a wage/salary job but report having no “wages, salary, tips, bonuses, or commissions from that job”. Therefore, changes were made to the CATI/CAPI questions only for this test. However, the analysis studies the impact on the full item since we do not publish ACS data by mode.

1.3 Recommendations from Cognitive Testing

Prior to conducting the 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. Multiple 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.

The main recommendation was to include the word “additional” as follows: Did [`<Name>/you`] receive any [if yes, fill with "additional"] tips, bonuses or commissions? If they answered yes to the first part, “additional” was to be used for this second portion. It was also suggested to indicate annual figures for each question. It was recommended to do so by adding “in the past 12 months” to each part of the questions. Many respondents wondered if they should calculate weekly, monthly or annual figures and this change could potentially avoid this confusion. The current ACS question asks about all of these types of earned income in a single question, and the Census Bureau had concerns about order effects caused by the presentation of the list of types of earnings. Therefore, the revised version was designed to ask the question in two steps: first ask about salary and wages, and then ask about additional earned income in bonuses, tips and commissions. In addition, a question on self-employment income was explicitly added to further separate different types of earnings and make sure respondents remembered to mention that income as well.

Both versions tested included an introduction that established the reference period: “The next few questions are about income DURING THE PAST 12 MONTHS...” Following that introductory statement, the current ACS question asks if the respondent received “any wages, salary, tips, bonuses, or commissions” and, if so, how much was received before taxes and other deductions. The alternative questions tested separating the different types of earnings. After the introductory statement questions ask if the respondent received any wages or salary, and if so, how much they received from all jobs before taxes and other deductions. Respondents are then asked whether they received any additional tips, bonuses or commissions during the past 12 months and if so, how much the person received from all jobs before taxes and other deductions.

For more information see (RTI International, “Cognitive Testing of the American Community Survey Content Test Items, 2009.”)

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 suggestion was to include “during the past 12 months” in the beginning of each question. This was to be stated within the question asking about wages or salary and for tips, bonuses and commissions. This “during the past 12 months” statement would also be restated in the amount questions following.

2. SELECTION CRITERIA

The research questions in sections 5.2 through 5.11 appear in order of importance for the decision of whether the test version of the question is better than the control question. The selection criteria below are also shown in order of importance to the decision.

The overall distribution of wages income for the test version should have been comparable to that of the CPS ASEC. An increase in wages income receipt and the amount of wages income received in the test version implies a positive change since this item was presumed to be underestimated. The lower part of the wages distribution should have shifted higher. The item missing data rates and response error (i.e., bias) were thought to be considered together when determining whether the test version performs better.

Since changes to the wages income question for the test version appear only in the CATI/CAPI instrument (and not in the mail questionnaire) we evaluated the following items together by response mode: item missing data rates; the estimates of wages reciprocity, wages income amount means and medians; and response error, as measured by net difference rates.

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 F 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 Wages

Only persons 15 or older were considered in the universe for the analysis, since all income questions are only asked of this universe. On the mail questionnaire, wages was determined if there was a Yes response in the reciprocity field or if a dollar amount greater than zero was in wages/salary or tips, bonuses and commissions.

The CFU question was not used as a direct method to ask respondents twice. The ASEC questions were instead used to make inferences and use as a “true measure”. See Appendix E for CFU question wording.

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 in order for the Census Bureau to contact them for a CFU interview. As a result, 18.4 percent of the respondents 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 about 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.

Changes to the wages questions were only made in CATI and CAPI but the report presents combined results for all modes. The sample was not designed to test by mode separately so significant results for the CATI/CAPI changes may have not been identified due to limited sample size.

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 show 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 cannot explain the decrease 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 Is the response distribution of wages income comparable to the Current Population Survey’s Annual Social and Economic Supplement (ASEC) distribution of wages income?

Table 2 shows the response distributions of the test and control versions compared to the 2010 CPS ASEC. Formal statistical comparisons were not made since the Content Test data was not edited or imputed, adjusted for nonresponse, nor raked to known population totals. One and two dollar amounts (\$1 and \$2) are sometimes CATI/CAPI keying errors so these amounts are tallied separately. Interviewers sometimes key a “1” or a “2” from the Yes/No fields into the amount field. This is less likely to occur as a keying error in the Mail forms.

The overall distribution of wages income for the test version is comparable to that of the CPS ASEC.

Category	ASEC Estimate (%)	Standard Error (%)	Test Estimate (%) (n=16,526)	Standard Error (%)	Control Estimate (%) (n=16,755)	Standard Error (%)
\$1 or \$2	0.0	NA	0.1	0.0	0.1	0.0
\$3 - \$2,499	5.4	NA	6.4	0.3	7.0	0.3
\$2,500 - \$19,999	27.3	NA	25.8	0.5	26.6	0.5
\$20,00 - \$39,999	28.5	NA	26.1	0.5	26.3	0.5
\$40,000 - \$64,999	21.6	NA	22.0	0.4	21.3	0.4
\$65,000 +	17.0	NA	19.6	0.4	18.6	0.5
Total:	100.0		100.0		100.0	

Source: U.S. Census Bureau, 2010 Current Population Survey, Annual Social and Economic Supplement.

5.3 Do the changes to the wages question raise the proportion of persons receiving wages income?

Table 3 shows reciprocity rates of persons receiving wages for the control and test groups and the difference between the test and control groups. A one-sided test was used to determine if the Test group has a statistically significant larger reciprocity proportion using an $\alpha = 0.10$.

The changes to the wages question did not significantly raise the estimate of persons receiving wages income. This was a goal of the content test. Since this was not achieved this is a negative results. Section 5.9 explains significant changes for reciprocity by mode.

	Test Estimate (%)	Standard Error (%)	Control Estimate (%)	Standard Error (%)	Test – Control (%)	Standard Error (%)	Significance
Reciprocity Rate	93.5	0.3	93.3	0.2	0.2	0.4	No

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

5.4 Do the changes to the wages question raise the estimate of wages income?

Table 4 shows median and mean estimates of wages income for the test and control groups and the difference between the test and control groups. A one-sided test was used to determine if the Test group has a statistically significant larger median using an $\alpha = 0.10$. The calculations showed there were mixed results. The test version median estimate of wages income is significantly higher than the control version, but the test version mean estimate is not significantly higher than the control version mean. This is capturing what was expected from the test version therefore a positive result which adds support to implement the test version.

Measure	Test Estimate	Standard Error	Control Estimate	Standard Error	Test - Control	Standard Error	Significance
Mean	\$45,701	\$2,490	\$42,713	\$681	\$2,988	\$2,622	No
Median	\$31,884	\$275	\$31,172	\$286	\$712	\$384	Yes

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

5.5 Do the changes to the wages question affect the response distribution, shifting the lower wage categories of the distribution higher?

The response distributions were compared between the control and test versions to determine if there were fewer \$0 amounts, \$1 or \$2 amounts, and amounts ranging from \$3 to \$2500 in the test panel and more amounts greater than \$2500 in the test panel. There was no expected wage category that should have increased due to the movement out of \$0 wages income. To test whether an overall categorical response distribution was

dependent on the question version (control or test), the Pearson’s chi square statistic, adjusted for the complex sample design, was calculated. In the event that the null hypothesis was rejected, the difference in the proportions between the control and test groups was also computed to determine if the two groups have significantly different proportions using a Bonferonni-Holm adjusted alpha controlling the family-wise error level of 0.10.

Table 5 shows that the changes to the wages question do not significantly affect the response distribution. Specifically, there was no increase in the “\$0” values for the test which would have been considered positive since changes to the question were made to decrease reports of no wages.

Table 5. Shift in Distribution							
Category	Test Estimate (%) (n=16,557)	Standard Error (%)	Control Estimate (%) (n=16,782)	Standard Error (%)	Test-Control (%)	Standard Error (%)	Significance
\$0	0.1	0.0	0.2	0.0	0.0	0.0	NO
\$1 or \$2	0.1	0.0	0.1	0.0	0.0	0.1	NO
\$3 - \$2,499	6.4	0.3	6.9	0.3	-0.5	0.4	NO
\$2,500-\$19,999	25.7	0.5	26.6	0.5	-0.9	0.7	NO
\$20,00 -\$39,999	26.0	0.5	26.3	0.5	-0.3	0.7	NO
\$40,000 -\$64,999	22.0	0.4	21.3	0.4	0.7	0.6	NO
\$65,000 +	19.6	0.4	18.6	0.5	1.0	0.6	NO
$\chi^2= 6.3$ with 6 degrees of freedom, not significant at the 10 percent level.							

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

5.6 Do the changes to the wages question result in the same or lower item missing data rates?

Table 6 shows that the item missing data rates for both wages reciprocity and wages amount are significantly higher in the test version. Statistical significance of differences was determined at the $\alpha = 0.10$ significance level using a one-sided test.

Much of this missing data is due to the two part question and how the test version was considered to be missing if one or both of the questions was not answered. The increase in the missing data rate for reciprocity appears to be due to a notably higher missing data rate for the second question in the test version, the one asking about tips, bonuses, and commissions, the components from the current question we thought respondents were having difficulty with originally. The missing data rate for reciprocity for this question is more than 3 times higher than the wages test question (4.6% vs. 1.4%).

	Test Estimate (%)	Standard Error (%)	Control Estimate (%)	Standard Error (%)	Difference Estimate (%)	Standard Error (%)	Test signif. less than control?
Reciency:	15.2 (n=20,151)	0.4	13.5 (n=20,353)	0.3	1.7	0.5	NO ¹
Amount:	13.7 (n=18,679)	0.4	12.9 (n=18,835)	0.4	0.8	0.6	NO ¹

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

¹ Test is significantly greater than control at the $\alpha = 0.10$ significance level using a one-sided test.

5.7 Do the changes to the wages question lower response error (i.e., bias) in the estimate of wages reciency and wages income?

Using data from the Content Test and CFU, net difference rates were compared between the control and test versions. A response was required in both survey measures to be included in this analysis. The net difference rate (NDR) provided an approximate measure of bias in the content test estimates when it was assumed that the reinterview provides a measure of “truth.”

CFU response (reinterview)	Content test response		
	Yes	No	Total
Yes	a	b	a+b
No	c	d	c+d
Total	a+c	b+d	n = a+b+c+d

$$ndr = estimatedvalue - truevalue \approx \frac{a+c}{n} - \frac{a+b}{n} = \frac{c-b}{n}$$

Note that the CFU used questions from the CPS Annual Social and Economic Supplement for the wages questions as well as the other income questions changed for the Content Test. The CFU was identical for the Control and Test versions. A negative NDR means that there is an overestimate of the true values while a positive ndr means there is an underestimate.

The difference in the absolute net difference rates ($| \text{test} | - | \text{control} |$) for wage reciency and the standard error on the difference was computed. A one-sided test was used to determine if the Test group had a statistically significant lower net difference rate than the Control group using an $\alpha = 0.10$. A negative ndr means that there is an overestimate of the true values while a positive ndr means there is an underestimate.

A net difference rate for each of the income ranges for the control and test groups was calculated. The difference in the absolute net difference rates ($| \text{test} | - | \text{control} |$) and the standard error on the difference was also calculated. A one-sided test was used to

determine if the Test group had a statistically significant lower net difference rate than the Control group for each of the income ranges using a Bonferonni-Holm adjusted alpha controlling the family-wise error level of 0.10.

Table 7 shows the categories with NDRs that are significantly lower for the test version of amount are DK/REF, \$3-\$2,499 and \$2,500 to \$19,999. This means that the test version significantly reduced the overestimate of wage amounts in these 3 categories. Since fewer respondents to the test version gave a DK/REF response, breaking up the question into two parts seemed to help respondents provide an amount. The NDR was not significantly lower for the test version of wages reciprocity.

Table 7. Net Difference Rates							
	Test Estimate (%)	Standard Error (%)	Control Estimate (%)	Standard Error (%)	Test - Control (%)	Standard Error (%)	Test signif. less than control?
Reciprocity:	(n=8,595)		(n=8,627)				
	15.9	0.7	17.0	0.8	-1.1	1.0	NO
Amount:	(n=6,247)		(n=8,164)				
DK/REF	- 7.7	0.9	-12.4	0.6	-4.7	1.0	YES
\$0	0.0	0.0	-0.1	0.1	0.1	0.1	NO
\$1 or \$2	0.0	0.0	0.0	0.0	0.0	0.0	NO
\$3 - \$2,499	0.7	0.3	1.8	0.3	-1.0	0.4	YES
\$2,500 -\$19,999	1.3	0.6	3.6	0.6	-2.3	0.9	YES
\$20,00 -\$39,999	2.3	0.6	3.5	0.5	-1.2	0.8	NO
\$40,000 -\$64,999	1.6	0.6	1.5	0.5	0.1	0.7	NO
\$65,000 +	1.7	0.4	2.1	0.4	-0.5	0.5	NO

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

5.8 Do the changes to the wages questions lower the estimate of poverty rate?

Crude estimates of the poverty rate, based on unedited data, between the control and test versions were compared. Since changes to the property income and public assistance questions were changed for the Content Test, those changes could have also lowered the poverty rate.

For the test panel, total income was used as defined in the poverty recode specification but amounts reported as tips, bonuses, and commissions were added to the total. The difference in the poverty rates (test – control) and the standard error on the difference was computed. A one-sided test was used to determine if the test group has a statistically significant lower poverty rate using an $\alpha = 0.10$.

Changes to the wages question (in conjunction with the changes to the other income questions: Property Income and Public Assistance) do not significantly lower the estimate of poverty rate.

Table 8. Poverty Estimate							
	Test Estimate (%)	Standard Error (%)	Control Estimate (%)	Standard Error (%)	Test-Control (%)	Standard Error (%)	Significance
Poverty:	32.2	0.4	31.5	0.5	0.8	0.7	NO

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

5.9 For each mode of data collection, do the changes to the wages question affect the item missing data rates, the estimates of reciprocity and wages income, or response error (i.e., bias)?

Table 9a shows that by mode item missing data rates were significantly higher for CATI and CAPI individually and combined for the test version. Much of this missing data is due to the two part question and how the test version was considered to be missing if one or both of the questions was not answered. The increase in the missing data rate for reciprocity appears to be due to a notably higher missing data rate for the second question in the test version, the one asking about tips, bonuses, and commissions, the components from the current question we thought respondents were having difficulty with originally. This also showed that for the item missing data rates for the test question were significantly lower for mail mode. This finding must be a random finding since the wages questions were identical for the test and control.

Table 9a. Item Missing Data Rates for Reciprocity							
Mode	Test Estimate (%)	Standard Error (%)	Control Estimate (%)	Standard Error (%)	Test-Control (%)	Standard Error (%)	Test signif. less than control?
Mail	24.1 (n=12,533)	0.6	25.1 (n=12,563)	0.5	-0.9	0.7	YES
CATI/CAPI	5.1 (n=7,618)	0.3	0.9 (n=7,790)	0.2	4.3	0.4	NO ¹
CATI	4.1 (n=2,425)	0.5	0.7 (n=2,474)	0.1	3.5	0.6	NO ¹
CAPI	5.4 (n=5,193)	0.4	0.9 (n=5,316)	0.2	4.4	0.5	NO ¹

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

¹ Test is significantly greater than control at the $\alpha = 0.10$ significance level using a one-sided test.

Table 9b shows that by mode item missing data rates for wages amount were significantly higher for the CATI/CAPI combined and CAPI only version. CAPI cases seem to be driving this higher missing data rate .

Mode	Test Estimate (%)	Standard Error (%)	Control Estimate (%)	Standard Error (%)	Test-Control (%)	Standard Error (%)	Test signif. less than control?
Mail	2.1 (n=11,406)	0.2	2.2 (n=11,427)	0.2	-0.2	0.3	NO
CATI/CAPI	26.3 (n=7,273)	0.8	24.1 (n=7,408)	0.8	2.2	1.2	NO ¹
CATI	24.4 (n=2,319)	1.5	22.5 (n=2,334)	1.2	1.8	1.9	NO
CAPI	26.8 (n=4,954)	1.0	24.4 (n=5,074)	1.0	2.3	1.5	NO ¹

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

Table 9c shows there were mixed results, when looking at wage reciprocity by mode. The CATI mode wage reciprocity was significantly higher on the test version, which part of the goal for the question changes as discussed earlier in this report. Mail, CATI/CAPI and CAPI alone showed no significant results.

Mode	Test Estimate (%)	Standard Error (%)	Control Estimate (%)	Standard Error (%)	Test-Control (%)	Standard Error (%)	Significance
Mail	91.5	0.3	91.6	0.3	-0.1	0.4	NO
CATI/CAPI	95.8	0.4	95.3	0.4	0.5	0.5	NO
CATI	95.9	0.5	94.7	0.6	1.1	0.8	YES
CAPI	95.7	0.4	95.4	0.4	0.3	0.6	NO

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

Table 9d illustrates that CATI/CAPI combined and CAPI alone have the NDR for wage reciprocity significantly lower for the test version. Statistical significance of differences is determined at the $\alpha = 0.10$ significance level using a one-sided test.

<i>Recipieny:</i>	Test Estimate (%)	Standard Error (%)	Control Estimate (%)	Standard Error (%)	Test - Control (%)	Standard Error (%)	Test signif. less than control?
Mail	11.5 (n=5,841)	0.7	11.6 (n=5,888)	0.7	-0.1	1.0	NO
CATI/CAPI	20.8 (n=2,754)	1.4	23.2 (n=2,739)	1.4	-2.4	1.8	YES
CATI	21.4 (n=986)	2.0	21.2 (n=988)	2.2	0.2	2.7	NO
CAPI	20.6 (n=1,768)	1.6	23.6 (n=1,751)	1.5	-2.9	2.0	YES

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

Net different rates are shown in Table 9e for the mail mode. For this family of one-sided hypothesis tests, the family-wise error rate has been controlled using the Bonferroni-Holm multiple comparison method at the $\alpha = 0.10$ level.

The net difference rate is significantly lower for mail amount category \$40,000 - \$64,999 in the test version. Besides this one change there were no more significant findings for mail mode. There is no explanation for this result since no changes were made to the mail version of the question.

Amount:	Test Estimate (%)	Standard Error (%)	Control Estimate (%)	Standard Error (%)	Test - Control (%)	Standard Error (%)	Test signif. less than control?
\$0	0.1 (n=4,533)	0.0	0.0 (n=4,531)	0.1	0.0	0.1	NO
\$1 or \$2	0.0	0.0	0.0	0.0	0.0	0.1	NO
\$3 - \$2,499	0.9	0.3	1.0	0.3	-0.1	0.4	NO
\$2,500 - \$19,999	-0.7	0.5	0.1	0.5	0.5	0.7	NO
\$20,00 - \$39,999	0.6	0.6	0.4	0.5	0.1	0.7	NO
\$40,000 - \$64,999	-0.1	0.5	-1.4	0.5	-1.3	0.6	YES
\$65,000 +	-0.7	0.3	-0.1	0.3	0.6	0.4	NO

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

See tables A-1 to A-3 for additional testing.

5.10 For each mail response stratum, do the changes to the wages question affect the item missing data rates, the estimates of recipieny and wages income, or response error (i.e., bias)?

For the high response stratum, the net difference rates for wages reciprocity and wages amount for the DK/Ref category are significantly lower for the test version than the control. See table 10a.

Table 10a. High Response Stratum- Net Difference Rate							
	Test Estimate (%)	Standard Error (%)	Control Estimate (%)	Standard Error (%)	Test-Control (%)	Standard Error (%)	Significance
Reciprocity:	14.8	0.8	16.6	0.9	-1.9	1.2	NO
Amount:							
DK/REF	9.2	1.0	13.3	0.7	-4.0	1.2	YES
\$0	0.0	0.0	0.1	0.1	0.0	0.1	NO
\$1 or \$2	0.1	0.0	0.0	0.0	0.0	0.0	NO
\$3 - \$2,499	0.7	0.4	1.9	0.4	-1.2	0.5	NO
\$2,500 - \$19,999	1.8	0.7	3.4	0.7	-1.6	1.0	NO
\$20,00 - \$39,999	2.6	0.7	3.8	0.6	-1.2	1.0	NO
\$40,000 - \$64,999	2.0	0.7	1.7	0.6	.03	0.9	NO
\$65,000 +	2.0	0.5	2.4	0.5	0.5	0.6	NO

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

For the low response stratum, the item missing data rate for wages amount is significantly higher for the test version; but the net difference rates are significantly lower for the test version in the DK/REF and \$2,500 to \$19,999 categories. See table 10b below. There are no other significant findings by mail response stratum.

Table 10b. Low Response Stratum- Net Difference Rate							
	Test Estimate (%)	Standard Error (%)	Control Estimate (%)	Standard Error (%)	Test-Control (%)	Standard Error (%)	Significance
Reciprocity:	19.2	0.9	18.3	0.9	0.9	1.1	NO
Amount:							
DK/REF	1.2	1.2	9.5	0.9	-8.3	1.4	YES
\$0	0.1	0.1	0.2	0.1	-0.2	0.2	NO
\$1 or \$2	0.0	0.0	0.0	0.0	0.0	0.0	NO
\$3 - \$2,499	0.8	0.4	1.3	0.4	-0.6	0.6	NO
\$2,500 - \$19,999	1.0	0.8	4.1	0.7	-3.1	1.1	YES
\$20,00 - \$39,999	1.0	0.8	2.3	0.7	-1.3	1.0	NO
\$40,000 - \$64,999	0.1	0.6	0.8	0.4	-0.7	0.7	NO
\$65,000 +	0.3	0.4	1.1	0.3	-0.8	0.5	NO

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

See tables A-4 to A-6 in Appendix A for additional testing.

5.11 Does either question version elicit respondent or interviewer behaviors that may contribute to interviewer or respondent error?

Results indicate that for the series as a whole the tester performs better on interviewer behavior, whereas for respondent behavior, the difference between the test and control

series is not significant. The behavior coding results were compared and derived from the CARI recordings between the control and the test versions.

6. SUMMARY

Ultimately the results found an increase in wage reciprocity for CATI mode. The NDR for CATI/CAPI combined and CAPI significantly decreased. The test version median estimates of wages income was significantly higher than the control. Several negative results include that item missing data rates for wages amounts and reciprocity increased overall. Missing data rates were significantly higher for the CATI/CAPI and CAPI only version. Item missing data rates were also higher for wage reciprocity overall as well. By mode item missing data rates were significantly higher for CATI/CAPI individually and combined. Based on these results it was recommended to implement the test version.

References

RTI International (August 12, 2009). "Cognitive Testing of the American Community Survey Content Test Items," Research Triangle Park, NC).

Webster, 2006. "Comparison of Income from the CPS and ACS," U.S. Census Bureau.

Acknowledgements

The authors would like to acknowledge Mary C. Davis and Padraic Murphy for their contributions to the statistical analysis of this report

Appendix A: Tables

Table A-1. Net Difference Rates for Amounts (CATI/CAPI)							
Category	Test Estimate (%)	Standard Error (%)	Control Estimate (%)	Standard Error (%)	Test - Control (%)	Standard Error (%)	Test signif. less than control?
Amount:							
\$0	0.1	0.1	0.2	0.1	0.1	0.1	NO
\$1 or \$2	0.0	0.0	0.0	0.0	0.0	0.0	NO
\$3 - \$2,499	2.6	0.9	1.3	0.6	1.2	1.0	NO
\$2,500 - \$19,999	9.5	1.7	1.6	1.0	8.0	2.0	NO
\$20,00 - \$39,999	6.2	1.5	2.5	1.0	3.7	1.8	NO
\$40,000 - \$64,999	3.7	1.8	0.8	0.8	2.9	2.0	NO
\$65,000 +	1.6	0.8	1.0	0.6	0.6	1.0	NO

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

Table A-2. Net Difference Rates for Amounts (CATI)							
Category	Test Estimate (%)	Standard Error (%)	Control Estimate (%)	Standard Error (%)	Test - Control (%)	Standard Error (%)	Test signif. less than control?
Amount:							
\$0	0.0	0.0	0.0	0.0	0.0	0.0	NO
\$1 or \$2	0.0	0.0	0.0	0.0	0.0	0.0	NO
\$3 - \$2,499	1.4	0.8	1.3	0.8	0.1	1.1	NO
\$2,500 - \$19,999	10.5	3.4	0.2	1.2	10.3	3.6	NO
\$20,00 - \$39,999	4.2	1.7	2.0	1.0	2.1	1.8	NO
\$40,000 - \$64,999	3.4	1.7	0.4	1.1	3.0	1.9	NO
\$65,000 +	1.6	1.1	1.7	0.9	-0.1	1.5	NO

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

Table A-3. Net Difference Rates for Amounts (CAPI)							
Category	Test Estimate (%)	Standard Error (%)	Control Estimate (%)	Standard Error (%)	Test - Control (%)	Standard Error (%)	Test signif. less than control?
Amount:							
\$0	0.1	0.1	0.2	0.2	-0.2	0.8	NO
\$1 or \$2	0.0	0.0	0.0	0.0	0.0	0.0	NO
\$3 - \$2,499	2.8	1.0	1.3	0.6	1.4	1.1	NO
\$2,500 - \$19,999	9.3	2.2	1.9	1.2	7.4	2.5	NO
\$20,00 - \$39,999	6.6	1.8	2.6	1.2	4.0	2.1	NO
\$40,000 - \$64,999	3.8	2.1	1.0	0.9	2.8	2.3	NO
\$65,000 +	2.3	1.0	1.0	0.7	1.4	1.2	NO

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

Table A-4. Reciprocity by Response Stratum							
Mode	Test Estimate (%)	Standard Error (%)	Control Estimate (%)	Standard Error (%)	Test-Control (%)	Standard Error (%)	Significance
High	93.5	0.3	93.3	0.3	0.2	0.5	NO
Low	93.6	0.3	93.5	0.3	0.0	0.4	NO

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

Table A-5. High Response Stratum Item Missing Rate							
Mode	Test Estimate (%)	Standard Error (%)	Control Estimate (%)	Standard Error (%)	Test-Control (%)	Standard Error (%)	Significance
Reciprocity:	15.9	0.5	14.7	0.4	1.2	0.6	NO
Amount:	11.0	0.5	10.5	0.5	0.6	0.7	NO

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

Table A-6. Low Response Stratum Item Missing Rate							
Mode	Test Estimate (%)	Standard Error (%)	Control Estimate (%)	Standard Error (%)	Test-Control (%)	Standard Error (%)	Significance
Reciprocity:	13.0	0.3	9.7	0.3	3.4	0.4	NO
Amount:	18.3	0.5	17.0	0.5	1.3	0.7	NO

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

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

CATI/CAPI Control Wording

The next few questions are about income DURING THE PAST 12 MONTHS...

Did [<Name>/you] receive any wages, salary, tips, bonuses or commissions?

<1> Yes

<2> No

If yes, How much did [<Name>/you] receive?

Report amount from all jobs before any deductions for taxes, bonds or other items.

\$_____ .00

CATI/CAPI Test Wording

The next few questions are about income DURING THE PAST 12 MONTHS, that is from <DATE> to <DATE>...

Did [<Name>/you] receive any wages or salary?

<1> Yes

<2> Nos

If yes, How much did [<Name>/you] receive in wages and salary from all jobs before taxes and

other deductions? \$_____ .00

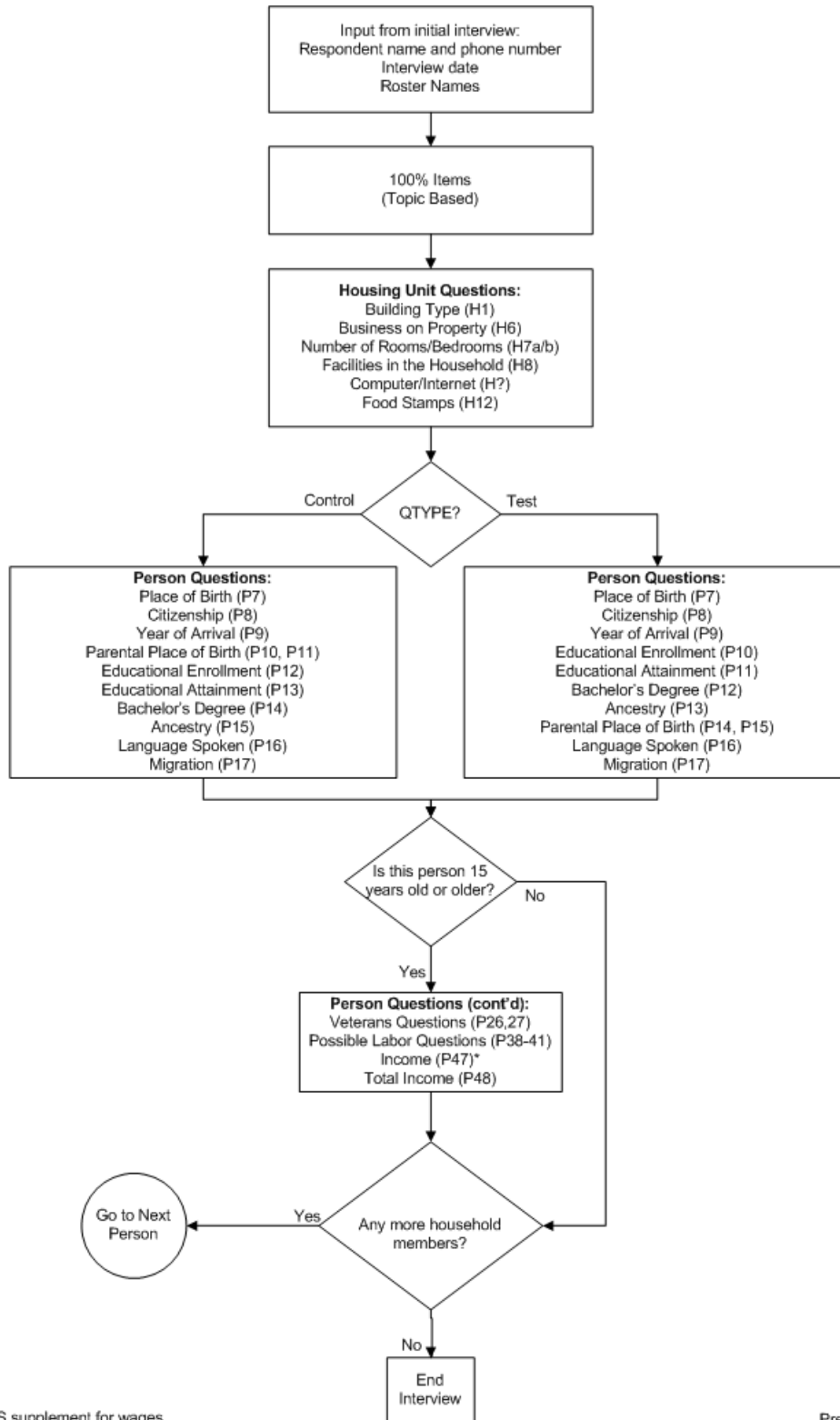
Did [<Name>/you] receive any tips, bonuses or commissions DURING THE PAST 12 MONTHS?

<1> Yes

<2> No

If yes, How much did [<Name>/you] receive in tips, bonuses, or commissions from all jobs before taxes and other deductions? \$_____ .00

Appendix C: Flow of the Content Follow-Up



* using CPS supplement for wages,
property income & cash public
assistance

Appendix D: Information Page

Test Design

Treatments	Two question versions with different wording for CATI/CAPI only.
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. The change to this question will only occur in the CATI and CAPI instruments, however all modes will be considered in the analysis. <i>CATI and CAPI interviews will be recorded using Computer-Assisted Recorded Interviewing (CARI) technology.</i>
Time Frame	Same schedule as the production September panel: mail out 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	Is the response distribution of wages income comparable to the Current Population Survey's Annual Social and Economic Supplement (ASEC) distribution of wages income?	Compare the response distribution of wages income between the test version and CPS ASEC. <i>Formal statistical comparisons cannot be made since the Content Test data will not have been edited or imputed, nor will there be adjustments for nonresponse or raking to known population totals.</i>
2	Do the changes to the wages question raise the estimate of persons receiving wages income?	Compare the estimate of persons receiving wages income between the control and test versions.
3	Do the changes to the wages question raise the estimate of wages income?	Compare the mean and median estimate of wages income between the control and test versions.
4	Do the changes to the wages question affect the response distribution, shifting the lower wage categories of the distribution higher?	Compare the response distributions between the control and test versions.

No.	Research Questions	Evaluation Measures
5	Do the changes to the wages question lower item missing data rates?	Compare the item missing data rates between the control and the test versions.
6	Do the changes to the wages question lower response error (i.e., bias) in the estimate of wages reciprocity and wages income?	Using data from the Content Test and CFU, compare net difference rates between the control and test versions (based on answers to more detailed content follow-up questions).
7	Do the changes to the wages question lower the estimate of poverty rate?	Compare a crude estimate of poverty rate, based on unedited data, between the control and test versions.
8	For each mode of data collection, do the changes to the wages question affect the item missing data rates, the estimates of reciprocity and wages income, or response error (i.e., bias)?	<p>For each mode (mail,CATI,CAPI), compare the item missing data rates, estimates of reciprocity and wages income, and response error (i.e., bias) between the control and the test versions.</p> <p><i>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).</i></p>
9	For each mail response stratum, do the changes to the wages question affect the item missing data rates, the estimates of reciprocity and wages income, or response error (i.e., bias)?	For each mail response stratum (high and low), compare the item missing data rates, estimates of reciprocity and wages income, and response error (i.e, bias) between the control and the test versions.
10	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.

Selection Criteria (In order of priority)

Research Question(s)	Criteria
1	The overall distribution of wages income for the test version should be comparable to that of the CPS ASEC.
2-4	An increase in wages income receipt and the amount of wages income received in the test version implies a positive change since this item is presumed to be underestimated. The lower part of the wages distribution should shift higher.
5,6	The item missing data rates and response error (i.e., bias) will be considered together when determining whether the test version performs better.

Supplemental Information

Research Question(s)	Criteria
7	Not part of the selection criteria. A crude estimate of poverty rate should show a decrease in the percentage number of households in poverty.
8-10	Not part of the selection criteria. These data are presented to give additional information regarding how the questions performed.

Appendix E: CFU Wording

ACS 47 a. wages, salary, commissions, bonuses, tips

CPS INSERT FOR WAGES, SALARY, COMMISSIONS...

CPS Q48aa

How much did (name/you) earn from this job before taxes and other deductions in the past 12 months?

- Enter dollar amount
- Enter 0 for none

CPS Q48a3

Does this amount include all tips, bonuses, overtime pay, or commissions (name/you) may have received from this employer in the past 12 months?

- 1 Yes (SKIP TO CPS Q49a)
- 2 No

CPS Q48aad

How much did (name/you) earn in tips, bonuses, overtime pay, or commissions from this employer in the past 12 months?

- Enter dollar amount

(NOTE: the next questions go with the concept of asking about longest job first)

CPS Q49a

Did (name/you) earn money from any other work (you/he/she) did in the past 12 months?

- 1 Yes
- 2 No (SKIP TO ACS Q47ba)

CPS Q49b1d

How much did (name/you) earn from all other employers before taxes and other deductions in the past 12 months?

- Enter dollar amount
- Enter "0" for None

CPS Q49b13

Does this amount include all tips, bonuses, overtime pay, or commissions (name/you) may have received from all other employers in the past 12 months?

- 1 Yes (SKIP TO ACS Q47ba)
- 2 No

CPS Q49B1A

How much did (name/you) earn in tips, bonuses, overtime pay, or commissions from all other

employers in the past 12 months?

- Enter dollar amount
