

Impact of the 2012 Computer Audio Recorded Interviewing Application on Survey of Income and Program Participation Event History Calendar Response Rates and Item-Level Responses¹

Robyn Sirkis

US Census Bureau, 4600 Silver Hill Road, Washington DC 20233

Abstract

The Computer Audio Recorded Interviewing (CARI) system is a laptop software application that seamlessly records the verbal exchange of pre-specified questions between the interviewer and the respondent without disrupting the normal interview process. After the interview, any recordings obtained are digitally stored and can be reviewed later for quality assurance purposes. CARI was incorporated into the 2012 Survey of Income and Program Participation-Event History Calendar (SIPP-EHC) instrument. The purpose of this research is to determine the impact on response rates and measure the quality of estimates due to the addition of CARI to the survey instrument. Distributions and item nonresponse rates were compared within the 2012 CARI SIPP-EHC test and with the 2011 SIPP-EHC test. Cooperation rates were calculated to determine the percent of respondents who agreed to be recorded. Statistical models were constructed to determine if certain characteristics of respondents, geographic regions or interviews predict CARI consent propensity. Finally, we examined interviewer effects.

Key Words: Computer Audio Recorded Interviewing, Quality Assurance, SIPP-EHC

1. Introduction

The use of the Computer Audio Recorded Interviewing (CARI) application as a tool to ascertain interview authenticity and enhance data quality is quickly growing in survey programs. The U.S. Census Bureau conducted the 2012 CARI Survey of Income and Program Participation-Event History Calendar (SIPP-EHC) Field Test to determine if the deployment of CARI had a significant impact on response rates and data quality. It is essential that the U.S. Census Bureau determine the impact of CARI on data quality indicators. The deployment of CARI would be a major asset for survey programs using automated computer assisted survey instruments to meet quality assurance objectives.

¹ Disclaimer: This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress. Any views expressed are those of the authors and not necessarily those of the U. S. Census Bureau.

2. Background

2.1 CARI Application

CARI is a laptop software application, developed by the Research Triangle Institute (RTI) that is integrated into automated computer assisted survey instruments. With the respondent's consent, CARI seamlessly records the verbal exchange between the interviewer and the respondent without disrupting the normal interview process. Quality assurance staff can evaluate the likelihood that the exchange between the interviewer and the respondent is authentic by reviewing portions of the interview (Arceneaux, 2007). CARI potentially meets several needs to field interviewing including detecting interview fabrication, evaluating interviewer performance and providing feedback, identifying questionnaire problems, and recording information to open-ended questions (Herget, Biemer, Morton, and Sand, 2001). An additional benefit to incorporating CARI into a quality assurance program is the potential to reduce reinterview efforts and costs. Reinterview is a second interview conducted at a household to monitor interviewer work performance and identify response inconsistencies. Ultimately, CARI has the capability of lowering costs for determining the authenticity of field interviews by reducing field observations and verifications.

Previous tests have proven the capability of the CARI technology. The U.S. Census Bureau implemented the 2006 Health and Wellness Study (HWS) Field Test to demonstrate the feasibility of using CARI. Overall, CARI did not have an impact on system performance and audio quality. In general, the respondents were typically receptive to the introduction of CARI while the interviewers had mixed reactions (Arceneaux, 2007).

The 2010 American Community Survey (ACS) Content Test relied on CARI recordings from the automated computer assisted instruments to aid with decisions on questionnaire design by analyzing respondent and interviewer behaviors provided by the use of audio and image files. The U.S. Census Bureau and RTI collaboratively developed the CARI Interactive Data Access System (Nguyen, Thissen, and Siege, 2010) which allows users to review and monitor audio files.

2.2 Re-engineering SIPP

The U.S. Census Bureau is currently re-engineering the SIPP. SIPP is a nationally representative and longitudinal survey that collects information on topics such as income, participation in government transfer programs, employment, and health insurance coverage. The SIPP sample is a multistage-stratified sample of the U.S. civilian non-institutionalized population. Each respondent is interviewed every four months over the course of the panel or time period. Each four-month interview period is called a wave. Interviews began in September 2008 and will continue through December 2013 (U.S. Census Bureau, 2009).

The basic goals of the re-engineered SIPP-EHC were to create a method of collecting SIPP data that allowed for more accurate information to be reported, reduce respondent burden, and be cost-effective. Another objective was to create an instrument that was easier for interviewers to navigate while collecting SIPP data from the respondents. A

key component of the SIPP-EHC was to use a twelve-month recall period instead of the SIPP four-month recall period while utilizing event history methods (Fields, 2011).

Several tests were conducted to assess the data quality difference between the EHC and the traditional SIPP data collection methods beginning in 2008. In the 2011 test, the sample consisted of approximately 4,000 addresses in the lower income areas. Interviewing was administered across twenty states covering twelve Regional Offices (ROs) using the wave 1 Computer Assisted Personal Interviewing (CAPI) instrument. Data collection occurred mainly from January through March 2011. The 2012 SIPP-EHC test included 3,416 households that were interviewed in the 2011 SIPP-EHC test. Interviewers administered the SIPP-EHC interviews using a wave 2 instrument. Data collection occurred mainly from May through June 2012. The CARI SIPP-EHC Field Test occurred during the same time-period and is described in the next section. It is important to note that the samples for these tests are not generalizable to the population.

3. Methodology

3.1 CARI SIPP-EHC Survey Design

The wave 1 sample for the 2012 CARI SIPP-EHC Field Test consisted of 1,321 sample addresses selected from a high concentration of lower income areas chosen from the SIPP 2000 sample redesign frame. Interviews were administered across ten states covering six ROs. For each of these cases, this was the first time these respondents were contacted to complete a SIPP-EHC interview. Interviewers conducting the 2012 CARI SIPP-EHC also conducted 2012 SIPP-EHC interviews. Interviewers conducting the CARI SIPP-EHC interviews were provided with self-study training materials.

The 2012 CARI SIPP-EHC test recorded portions of the respondent's interview upon their consent. The questions chosen to be recorded included a wide range of topics in the household and person interview sections. There are two consent screens, one at the beginning of the household interview section and the other at the beginning of the person interview section. The interviewer asked the respondent to provide verbal consent to have their interview recorded for quality control purposes. After the introduction screen, the household respondent was asked the CARI consent question giving their verbal consent to record their responses to the household section and their own person interview. If the respondent agreed to be recorded the application was executed and a specific set of questions was automatically recorded during the interview. Then, all remaining eligible respondents 15 years old and older were asked the CARI consent question before the beginning of the person interview. The interview proceeded as normal without activating the CARI application if the respondent did not provide consent. If the interviewer exited and then re-entered the instrument the CARI consent question was asked to the respondent again to record the interview. For proxy interviews, the respondent was not asked the consent question prior to the interview if the consent question came on path during their own interview or during a proxy interview for another respondent. A proxy respondent completed the person interview if the respondent was less than 15 years old. When the respondent decided to leave the interview and a proxy respondent continued the interview, the interviewer manually turned off the recording before continuing the interview with the proxy respondent. The interviewer also turned off the recording when the respondent changed their mind and decided to not be recorded during the interview. The interview proceeded as normal but no new recordings were stored for the respondent.

3.2 Response Rates and Other Data Quality Indicators

The primary objective of the 2012 CARI SIPP-EHC test was to determine if the deployment of CARI had any significant impact on the response rates at the national level. Household response rates were compared among the 2012 CARI SIPP-EHC test, 2012 SIPP-EHC test and 2011 SIPP-EHC test. The wave 2 response rate is provided for the 2012 SIPP-EHC test. Response rates were also compared at the interviewer experience level. Experienced interviewers were defined as those who conducted a SIPP-EHC interview in 2010 and/or 2011 for this evaluation. The households and persons included in the analysis were those restricted to the states where data was collected in the 2012 CARI SIPP-EHC test.

In addition, other data quality indicators were compared among the tests. These included the noninterview eligible household rate, household refusal rate, temporary ineligible household rate, permanent ineligible household rate, telephone household rate, personal visit household rate, person proxy rate, and person noninterview rate. A household was considered to be noninterview eligible for this analysis when the interviewer was unsuccessful in obtaining an interview with an eligible sample household. Examples include households who refused to participate in the survey or the occupants were temporarily away. A household was considered to be temporary ineligible when the household was not eligible for an interview during the current interview period. Examples include ready to be occupied vacant units or a unit that was under construction. Additionally, a household was permanent ineligible for this analysis when the household no longer qualified for a SIPP-EHC interview. Examples include demolished structures and when the unit was converted to a permanent business or storage space. The household may also be contacted to complete an interview by personal visit or telephone. A person noninterview occurred when the interviewer obtained an interview with a household but did not obtain information for one or more adult sample persons.

The person noninterview rate was also calculated at the interviewer experience level. Returning and new household members were included in the person data quality indicators. Household members who moved out of the household in a prior interview (wave 2+) were excluded from the analysis. Proxy respondents have typically been used when the target respondent was not available. However, studies have shown that proxy reporting is most likely based on dispositional information rather than self-reporting (Schwarz and Wellens, 1994). However, having a lower proxy rate and a higher person noninterview rate may not be preferable. It is required that a proxy respondent completes the interview for a household member less than 15 years old. Therefore, proxy rates were only calculated for eligible respondents 15 years or older. The personal visit household rate should be higher for tests involving wave 1 interviews and lower for tests with wave 2+ interviews. It is desirable to have low noninterview eligible household rates and person noninterview rates. Out of the three tests, only the 2012 SIPP-EHC test included cases involving movers. Only those cases where the mover was not located were included in some of the rates below. The household (hh) and person rates are defined below.

HH Response Rate =

$$\frac{I + P}{I + P + A + D} \times 100$$

Noninterview Eligible HH Rate =

$$\frac{A}{I + P + A + D} \times 100$$

HH Refusal Rate =

$$\frac{\text{Refusals}}{I + P + A + D} \times 100$$

Temporary Ineligible HH Rate =

$$\frac{TI}{I + P + A + TI + PI + D} \times 100$$

Permanent Ineligible HH Rate =

$$\frac{PI}{I + P + A + TI + PI + D} \times 100$$

Telephone HH Rate =

$$\frac{T}{T + PV \text{ (Including I, P, A)}} \times 100$$

Personal Visit HH Rate =

$$\frac{PV}{T + PV \text{ (Including I, P, A)}} \times 100$$

Person Type Z Noninterview Rate =

$$\frac{Z}{\text{Eligible Persons Age 15+}} \times 100$$

Person Proxy Rate =

$$\frac{\text{Proxy}}{\text{Self + Proxy (Including Eligible Persons Age 15+)}} \times 100$$

where, I= complete interviews, P=partial interviews, A = noninterview eligible hh, D= mover not located, TI=temporary ineligible hh, PI=permanent ineligible hh, T=telephone interviews, PV=personal visit interviews, Self=self interviews, Proxy= proxy interviews, Z=person noninterviews

A difference of proportions test (using the t-test) was conducted under the null hypothesis that the response rates and other data quality indicators across two tests are the same/equal. The household response rates and person noninterview rates were compared within the 2012 CARI SIPP-EHC test by interviewer experience level. The statistical testing involved choosing the appropriate method (Satterthwaite or Pooled) depending on whether there were unequal or equal variances. The Satterthwaite method was used when the variances were unequal as opposed to the Pooled method which assumed equal variances. All statements in this report have undergone statistical testing at the 0.10 alpha level. The data are subject to error arising from a variety of sources.

3.3 Distributions and Item Nonresponse Rates Analysis

An additional objective of the test was to determine if there was a significant impact on distributions and item nonresponse rates. The topics covered in the analysis included coverage, demographics, annual programs, tax returns, assets, health and utilization expenses, disability, and adult well-being. Questions were chosen on the basis of variety and adequate sample sizes. There were 99 questions chosen for the distribution and item nonresponse rate analysis. Additionally, 15 questions were chosen for the distribution analysis where the interviewer answered the questions. The number of questions chosen from the topics is shown in the results section. The following comparisons were included in the distribution analysis.

- 2012 CARI SIPP-EHC test vs. 2011 SIPP-EHC test (CARI/2011)
- 2012 CARI SIPP-EHC test: Yes versus No consent (CARIY/N)
- 2012 CARI SIPP-EHC test vs. 2012 SIPP-EHC test (CARI/2012)

The 2012 CARI SIPP-EHC test was compared to the 2011 SIPP-EHC test since this was the first time the interviewers administered SIPP-EHC interviews to the respondents in the 2011 SIPP-EHC test. In terms of the second comparison, it was important to determine if the responses to the SIPP-EHC questions were affected based on whether the respondents thought they were being recorded during the interview. The CARI/2012 comparison was only used to evaluate the neighborhood questions, which were completed by the interviewer and not asked of the respondent. The household consent question response was used for the household section questions as well as the person demographics questions. The household respondent provided answers to questions such as age, race, Hispanic origin, education, marital status, and place of birth for the entire household. In this analysis, the household person was considered to be the same respondent that completed the person one interview (according to the order of the household roster). For the remaining eligible respondents, the person consent was used for the person interview questions.

Instances occurred where the respondent quickly changed their response to the consent question. On other occasions the respondent completed their interview over the course of more than one day or interviewer visit. Therefore, the consent question was asked more than once to the respondent. Due to these circumstances, it was decided to take the consent response when the household respondent completed the verify address question for the household interview portion. The verify address question immediately follows the consent question. The consent response was taken for the person interview if the respondent started the event history calendar. The respondent was excluded from the household or person portion of the distribution analysis if the consent response changed after these cut-off points. Proxy interviews were included in the distribution and item nonresponse rate analysis. The proxy respondent was considered to have provided consent if the respondent agreed to be recorded during their proxy interview, another proxy interview, or their own interview and did not change their consent response.

To assess the data quality, the distribution of responses for select questions across various topics was compared for the 2012 and 2011 SIPP-EHC tests. For both binary and multiple level categorical variables, the Fishers Exact test and Pearson's Chi-Square test were used to determine if the distributions for each comparison combination were significant at the 0.10 alpha level. The null hypothesis is that the proportions are equal. When some of the expected values are small, the Fishers Exact test is more accurate than the Chi-Square test of independence. Binary variables were created from the multiple categorical variables if there was a significant result in the comparison. The purpose was to determine if one or more specific categories in the multiple categorical variables caused a significant result in the comparison. The Kolmogorov-Smirnov two sample test was utilized for the continuous variables (ex. number of days sick). The null hypothesis is that the two samples or groups come from a common distribution. Continuous variables were also converted into categorical variables to determine if there were differences due to placing outliers into categories. Only complete and partial interviews were included in the distribution analysis.

Item nonresponse rates were calculated for the 2012 CARI SIPP-EHC test, 2012 SIPP-EHC test, and 2011 SIPP-EHC test for each question. Additionally, item nonresponse rates were calculated separately for those respondents who agreed to the consent question and those who did not agree to be recorded during the interview. The Fishers Exact test and Pearson's Chi-Square test were used to determine if the item nonresponse rates for each comparison combination were significant at the 0.10 alpha level. Item nonresponse

was defined as a don't know or refusal response to a question. Respondents that were not asked the question were excluded from the item nonresponse rate analysis. Only complete and partial interviews were included in the analysis. The coefficient of variation (cv) was calculated for each item nonresponse rate, which is the standard deviation, divided by the mean. Non-significant results for the majority of the questions in the distribution and item nonresponse rate analysis would indicate that there was not an impact on data quality due to the addition of CARI to the SIPP-EHC instrument.

3.4 Cooperation Measures

CARI cooperation measures were calculated to determine if respondents were receptive to the introduction of CARI in the SIPP-EHC instrument. The cooperation rates were calculated at the national and interviewer experience level. A respondent was considered as having provided consent for these measures upon agreement to be recorded for the entire interview and never altering their response. The respondent was considered to have not provided consent when the respondent did not agree to be recorded but then instantly changed their mind and provided a yes response to the consent question. When the household respondent agreed to be recorded for the entire interview and then changed their mind during their person interview the respondent was considered to have not provided consent for both sections (vice versa). This is different from the distribution analysis where it was necessary to determine if the responses to the SIPP-EHC questions were affected based on whether the respondents thought they were being recorded during the interview. The household (hh) and person level cooperation rates are defined below.

<p>Person Asked Cooperation Rate =</p> $\frac{\text{Number of persons who were asked the consent question and agreed to be recorded for the entire interview}}{\text{Number of persons asked consent question}} \times 100$	<p>Person Cooperation Rate =</p> $\frac{\text{Number of persons who agreed to be recorded for the entire interview}}{\text{Number of complete and proxy interviews}} \times 100$
<p>Full CARI HH Cooperation Rate =</p> $\frac{\text{Number of HH where all respondents agreed to be recorded for the entire interview}}{\text{Number of HH asked consent question}} \times 100$	<p>Partial CARI HH Cooperation Rate =</p> $\frac{\text{Number of HH where some respondents agreed to be recorded for the entire interview}}{\text{Number of HH asked consent question}} \times 100$
<p>Full CARI Refusal HH Rate =</p> $\frac{\text{Number of HH where all respondents did not agree to be recorded for the entire interview}}{\text{Number of HH asked consent question}} \times 100$	<p>HH Respondent Cooperation Rate =</p> $\frac{\text{Number of HH respondents who agreed to be recorded for the entire interview}}{\text{Number of HH respondents asked consent question}} \times 100$

Person level asked cooperation measures were calculated at the interviewer level in order to determine if the interviewers had an impact on whether the respondents agreed or did not agree to be recorded during the interview. Low cooperation measures would result in

fewer recordings to verify the authenticity between the interviewer and respondent during the interview.

3.5 CARI Consent Propensity Models

Person and household level logistic regression models of CARI consent propensity scores were created to evaluate which characteristics of the respondents or housing units affected CARI consent. All models take on the general form:

$$\text{logit}(p_i) = \log\left(\frac{p_i}{1-p_i}\right) = \alpha + \beta_1 x_1 + \dots + \beta_k x_k + \varepsilon$$

where p_i is the consent propensity of the i^{th} person and $(x_1 \dots x_k)$ is the vector of covariates.

Only those respondents asked the consent question were included in the statistical models. In addition, proxy interviews were excluded from the models. Potential predictor variables included characteristics of the respondent (e.x. sex, race, age, education level), the household (e.x. number of people in the household, housing unit conditions), and the interviewer experience level. The covariates with significant differences in the distribution analysis were used in the initial statistical models. In addition, only main effects were considered in the potential statistical models. Issues relating to sample size, missing values, and collinearity among predictor variables, was taken into account when choosing an appropriate model. Potential models were evaluated based on model fit statistics such as the R-squared value, which refers to the fraction of variance explained by the model. The odds ratios were examined for those covariates left in the statistical model. The odds ratio is the probability of an occurrence of an event to that of nonoccurrence. It accesses the strength of association and the potential impact of confounding variables. Automatic selection methods such as forward selection or backward elimination were used to aid in choosing a parsimonious model. Identifying similar characteristics for those that provided consent may be informative when considering how to implement CARI in future survey programs.

4. Findings

4.1 Response Rates and Other Quality Indicators

The primary purpose of the 2012 CARI SIPP-EHC test was to determine if CARI had any significant impact on response rates. Prior to calculating the response rates, the number of interviews was tallied at the household and person level as shown in Table 1.

Table 1: Number of Interviews-Restricted to States in CARI SIPP-EHC Test

Description	2012 CARI SIPP-EHC Test	2012 SIPP-EHC Test	2011 SIPP-EHC Test
Total Number Households	1,321	1,886	2,443
Total Interviewed Households	725	1,209	1,622
Total Interviewed Persons	2,050	3,576	4,485
Total Interviewed Persons (new or returning household members)	2,044	3,471	4,468

Household data quality indicators for the 2011 and 2012 tests are shown in Table 2.

The 2012 CARI SIPP-EHC test national household response rate was 68.85 percent compared to 79.70 percent in the 2012 SIPP-EHC test and 85.37 percent in the 2011 SIPP-EHC test. The difference of proportions tests assuming unequal variances found that the 2012 CARI SIPP-EHC test household response rate was significantly lower than the 2011 SIPP-EHC test and 2012 SIPP-EHC test household response rates. There were 169 interviewers who contacted households that were a part of the 2012 CARI SIPP-EHC test sample. Approximately 40 percent of those interviewers had experience administering SIPP-EHC interviews prior to 2012. The response rates were approximately 69 percent for the experienced and inexperienced interviewers. The difference of proportions test assuming equal variances indicated that the household responses rates were not significantly different between the experienced and inexperienced interviewers within the 2012 CARI SIPP-EHC test.

The 2012 CARI SIPP-EHC test household refusal rate was 15.57 percent. However, the household refusal rates were less than 10 percent for the 2012 SIPP-EHC test and 2011 SIPP-EHC test. The difference of proportions tests assuming unequal variances found that the 2012 CARI SIPP-EHC test household refusal rate was significantly higher than the 2011 SIPP-EHC test and 2012 SIPP-EHC test household refusal rates. The results of the difference of proportions test had the same results for the household refusal rates and the household noninterview eligible household rates.

The 2012 CARI SIPP-EHC test temporary ineligible household rate was 14.31 percent, compared to 16.74 percent in the 2011 SIPP-EHC test. The difference of proportions test assuming unequal variances found that the 2012 CARI SIPP-EHC test temporary ineligible household rate was significantly lower than the 2011 SIPP-EHC test. The 2012 CARI SIPP-EHC test permanent ineligible household rate was 5.98 percent compared to 5.49 percent rate in the 2011 SIPP-EHC test. The difference of proportions test found that the permanent ineligible household rate was not significantly different between the 2012 CARI SIPP-EHC test and 2011 SIPP-EHC test.

The 2012 CARI SIPP-EHC test household personal visit rate was 86.10 percent compared to 86.82 percent in the 2011 SIPP-EHC test and 79.95 percent in the 2012 SIPP-EHC test. The difference of proportions tests assuming equal variances indicated that the personal visit rate and telephone rate was not significantly different between the 2012 CARI SIPP-EHC test and 2011 SIPP-EHC test. However, the difference of proportions test assuming unequal variances showed that the 2012 CARI SIPP-EHC test personal visit rate was significantly higher than the 2012 SIPP-EHC test personal visit rate. Additionally, the difference of proportions test assuming unequal variances showed that the 2012 CARI SIPP-EHC test telephone rate was significantly lower than the 2012 SIPP-EHC test telephone rate.

Table 2: Household Data Quality Indicators (Percent)

Description (Rates)	2012 CARI SIPP-EHC Test	2012 SIPP-EHC Test	2011 SIPP-EHC Test
Response	68.85	79.70	85.37
Noninterview Eligible	31.15	19.45	14.63
Refusal	15.57	9.56	9.84
Telephone	13.90	20.05	13.18
Personal Visit	86.10	79.95	86.82

The 2012 CARI SIPP-EHC test person noninterview rate was 12.63 percent compared to 9.88 percent in the 2012 SIPP-EHC test and 10.84 percent in the 2011 SIPP-EHC test. The difference of proportions tests assuming unequal variances found that the 2012 CARI SIPP-EHC test person noninterview rate was significantly higher than the 2011 SIPP-EHC test and 2012 SIPP-EHC test person noninterview rates. There were 151 interviewers who obtained complete or partial 2012 CARI SIPP-EHC interviews. Approximately 37 percent of those interviewers had experience administering SIPP-EHC interviews prior to 2012. Inexperienced interviewers had a person noninterview rate of 13.49 percent compared to 10.27 percent for experienced interviewers. The difference of proportions test assuming unequal variances found that the person noninterview rate for inexperienced interviewers was significantly higher than person noninterview rate for experienced interviewers within the 2012 CARI SIPP-EHC test.

The proxy rate was 31.61 percent for the 2012 CARI SIPP-EHC test compared to 32.06 percent for the 2011 SIPP-EHC test and 38.46 percent for the 2012 SIPP-EHC test. The difference of proportions test assuming equal variances indicated that the proxy rate was not significantly different between the 2012 CARI SIPP-EHC test and the 2011 SIPP-EHC test. However, the difference of proportions test assuming unequal variances indicated the 2012 CARI SIPP-EHC test proxy rate was significantly lower than the 2012 SIPP-EHC test proxy rate.

4.2 Distributions and Item Nonresponse Rates Analysis

An additional research objective was to determine if there was a significant impact on the distributions and item nonresponse rates between the tests. As a result of the distribution analysis, significant differences existed in 42 out of 99 questions in the CARI/2011 comparison. This was one of the primary comparisons since both tests involved interviewers administering SIPP-EHC interviews to respondents for the first time. Some of the differences can potentially be explained by question wording, question structure, universes, answer categories, seasonality, and rare response categories. Specifically, modifications in question and answer categories existed from one test to the next. Seasonality differences occurred for topics such as tax returns due to the different time periods. There may be a chance that few households or respondents had a specific characteristic in one test but not the other. Several questions from the disability and adult wellbeing topics had structural question changes between 2011 and 2012. For example, the question layout and the way the interviewers entered the response was different from one year to the next.

Significant differences existed in 39 out of 99 questions in the CARIY/N comparison. Some of these differences occurred between the distributions due to rare response categories and small sample sizes. Table 3 shows the percentage of questions significant in each topic in the CARI/2011 and CARIY/N comparisons. The annual programs and assets topics had less of an impact than the other topics in the CARI/2011 comparison. The annual programs had less of an impact than the other topics in the CARIY/N comparison. The fifth column includes those significant in both comparisons. When only considering these two comparisons, there were 18 questions with significant results in just the CARI/2011 comparison, as opposed to 15 questions with significant results in just the CARIY/N comparison. Generally, converting the nine continuous variables into categorical variables to account for outliers did not seem to change whether the distributions had a significant result for each comparison.

Table 3: Number Questions Significant in Distribution Analysis

Topic	Questions Analyzed	Percent Significant		Percent
		CARI/2011	CARIY/N	Significant Both Comparisons
Coverage	4	50.00	0.00	0.00
Demographics	14	50.00	57.14	35.71
Annual Programs	25	32.00	20.00	16.00
Tax Returns	4	75.00	25.00	25.00
Assets	27	33.00	48.15	22.22
Health and Utilization	7	42.86	28.57	14.29
Disability	9	44.44	77.78	44.44
Adult Well Being	9	66.67	33.33	33.33

The distribution of interviewer responses for select questions for the neighborhood topic was compared within the 2012 CARI SIPP-EHC test and for the CARI/2012 comparison. Out of the 15 questions, three had significant results in only the CARI/2012 comparison but not the CARIY/N comparison. Four different questions had significant results in the CARIY/N comparison while one had significant results in both comparisons.

The differences in the item nonresponse rates were small between the CARI/2011, CARI/2012, and CARIY/N comparisons. Some differences occurred between the item nonresponse rates in the 2012 CARI SIPP-EHC test and 2011 SIPP-EHC test. Overall, 73 percent of the questions had higher item nonresponse rates for the 2011 SIPP-EHC test while 19 percent had higher item nonresponse rates for the 2012 CARI SIPP-EHC test. Additionally, significant differences occurred in 43 out of 99 questions in the CARI/2011 comparison. All 27 questions chosen from the assets topic had higher item nonresponse rates in the 2011 SIPP-EHC test. It is feasible that the same respondent replied with don't know or refusal answers for the entire section.

Additional differences occurred in the item nonresponse rates between the 2012 CARI SIPP-EHC test and the 2012 SIPP-EHC test than the previous comparison. In general, 53 percent of the questions had higher item nonresponse rates for the 2012 CARI SIPP-EHC test as opposed to 40 percent for the 2012 SIPP-EHC test. However, significant differences occurred in only 15 out of 99 questions in the CARI/2012 comparison. Approximately 84 percent (21 questions) in the annual programs topic had higher item nonresponse rates in the 2012 CARI SIPP-EHC test. The majority of the questions in the assets topic had higher item nonresponse rates for the 2012 SIPP-EHC test (93 percent).

Additional distinctions occurred between the item nonresponse rates for those respondents who agreed to be recorded versus those that did not provide consent. Nearly 45 percent of the questions resulted in higher item nonresponse rates for those that did not provide consent as opposed to 41 percent for those that agreed to be recorded during their interview. Although, significant differences occurred in only 6 out of 99 questions in the CARIY/N comparison. The majority of the questions in the assets topic had higher item nonresponse rates for those that provided consent (67 percent).

The coefficient of variation (cv) was less than 10 percent for the 2012 CARI SIPP-EHC test, 2012 SIPP-EHC test, and 2011 SIPP-EHC tests. The CVs were also less than 10 percent for those respondents that agreed to be recorded versus those who did not provide consent. Over 75 percent of the questions for each test including within the 2012 CARI SIPP-EHC test (yes and no) had CVs less than 5 percent.

4.3 Cooperation Measures

Household and person level cooperation rates were calculated at the national and interviewer experience level. It is important to note that the respondent was considered to have not provided consent if at any point during the interview a no response was offered to the consent question. Person level and household respondent cooperation rates were between 41 and 44 percent. The household respondent cooperation rate was 41.44 percent. The person asked cooperation rate was 41.68 percent and person cooperation rate (includes proxy interviews) was 43.58 percent. The cooperation rates were drastically different for the experienced interviewers (24-25 percent) from the inexperienced interviewers (48-51 percent). The difference of proportions tests assuming unequal variances found that the three cooperation rates for experienced interviewers were significantly lower than the cooperation rates for inexperienced interviewers within the 2012 CARI SIPP-EHC test.

Household cooperation measures were calculated to determine if the majority of the respondents in the household were providing the same response to the consent question. The full CARI household cooperation rate was 42.36 percent while the full CARI refusal rate was 57.64 percent only including households where one person was asked the consent question.

The full CARI household cooperation rate was 36.23 percent while the full CARI refusal household rate was 53.14 percent when more than one respondent was asked the consent question in the household. The partial CARI cooperation rate was 10.63 percent. This implies that there may be a household effect meaning that the majority of respondents provided the same response to the consent question within the household.

Finally, the full CARI household cooperation rate was 40.61 percent while the full CARI refusal household rate was 56.35 percent regardless of how many respondents were asked the consent question in the household.

Ultimately, this means that in most households all respondents either agreed or did not agree to be recorded for the entire interview.

Interviewer level cooperation measures were calculated to determine if the 151 interviewers had an impact on whether the respondents agreed or did not agree to be recorded during the entire interview. Interviewers were placed into categories based on the number of respondents who were asked the consent question. Then the interviewers were placed into additional categories depending on the percentage of respondents who agreed to be recorded for the entire interview divided by the number of persons who were asked the consent question. Below are important findings from the analysis.

- 13 out of 36 interviewers asked 10 or more respondents the consent question and the person asked cooperation rate was between 0 and 20 percent.
- 15 out of 30 interviewers asked between six and nine respondents the consent question and the person asked cooperation rate was between 0 and 20 percent.
- 14 out of 36 interviewers asked 10 or more respondents the consent question and the person asked cooperation rate was between 61 and 100 percent.
- 8 out of 30 interviewers asked between six and nine respondents the consent question and the person asked cooperation rate was between 61 and 100 percent.

It seems as though many interviewers did not get any respondents to agree to be recorded during the entire interview. Although there were interviewers who managed to ask a decent number of respondents the consent question while still having a high percent asked cooperation rate.

4.4 CARI Consent Propensity Models

Person and household level logistic regression models of CARI consent propensity scores were created to evaluate which characteristics of the respondents or housing units affected CARI consent. The R-squared and predictive power statistics were fairly low (R-squared less than 0.12) possibly due to the small sample size in the 2012 CARI SIPP-EHC test.

The covariates left in the final person level statistical model were social security income payments, earned income tax credit, regular non-interest earning checking account, disability, and education (less than a high school diploma or GED). The covariates that contributed most to the model were education (less than a high school diploma or GED) and regular non-interest earning checking account. Additionally, the odds of agreeing to the consent question for those respondents who had at least a high school diploma or GED were twice as large as for the respondents where the highest level of education was less than 12th grade (no high school diploma or GED). The covariates left in the final household statistical model were interviewer experience level, neighborhood (yards), interview in language other than English, and household size. The covariates that contributed most to the model were interviewer experience level and interview in language other than English. Furthermore, the odds of agreeing to the consent question were approximately three times as large for those respondents asked the consent question by inexperienced interviewers than respondents asked the consent question by experienced interviewers. Additionally, the odds of agreeing to the consent question were three times as large for those respondents who completed the interview in English than respondents who completed the interview in a language other than English. The majority of the covariates were chosen based on the results of the distribution and cooperation analysis results.

5. Conclusions and Recommendations

The 2012 CARI SIPP-EHC test household response rates were significantly lower than the 2012 SIPP-EHC test and 2011 SIPP-EHC test household response rates. However, there may be an explanation for some of these differences. The 2012 tests were conducted at the same time utilizing the same interviewers. It is plausible that the 2012 SIPP-EHC test took priority over the 2012 CARI SIPP-EHC test. Additional analysis could include determining if the interviewers started contacting households in one test before the other test. It is important to note that the respondents who were apart of the 2012 SIPP-EHC test had previously completed a SIPP-EHC interview. Furthermore, the CARI application was only briefly mentioned in the RO training sessions. Interviewers were provided with CARI self-study training materials. Finally, the RO realignment occurred in 2012, which involved closing six ROs and redefining some of the positions. The response rates could be improved if there is additional emphasis on CARI during training sessions and possibly allowing more time and resources to conduct the interviews.

An additional objective was to determine if there was a significant impact on the distributions and item nonresponse rates between the tests. Significant distribution differences existed in the CARI/2011 and CARIY/N comparisons. However, some of the differences can potentially be explained by question wording, question structure, universes, answer categories, seasonality, and rare response categories. In terms of item nonresponse rates there were some differences between the tests. The cooperation rates for the experienced interviewers were significantly different from the inexperienced interviewers. Improving the cooperation rate would involve implementing training procedures that stressed the importance of interviewers asking the consent question exactly as worded during the interview. Overall, there needs to be more testing on the impact on CARI on data quality, which includes examining the results from the 2013 SIPP-EHC test.

6. Limitations

To give proper interpretation to the results it is important to keep in mind limitations in the development of the evaluation methodology. Straightforward comparisons did not exist between the 2012 CARI SIPP-EHC test and the other tests. Some of the questions were modified from the 2011 to the 2012 SIPP-EHC tests. Seasonality could have an impact on the comparisons between the 2011 and 2012 tests. Respondents who were a part of the 2012 SIPP-EHC test had familiarity completing the SIPP-EHC the prior year. All interviewers asked the respondents the consent question in the 2012 CARI SIPP-EHC test. A more appropriate experiment would have a sample of households selected to be asked the consent question and others would not be asked the consent question.

Many assumptions were established when deciding whether to consider a respondent as having agreed to be recorded. For the cooperation measures, the respondent was deemed as having not offered consent even if their response changed from no to yes immediately during the start of the interview. Respondents were excluded from the distribution analysis, if the consent changed after the defined cut-off points in the evaluation. It is possible that the interviewer could have had an effect on the respondent changing their consent response. Sometimes the cases were reassigned to other interviewers after the initial interviewers already contacted the respondents. Therefore, the respondents could have changed their consent response due to the interviewer. This would need to be explored in future analyses. Finally some of the questions were not included in the data quality analyses including the distribution and statistical modeling due to small sample sizes.

7. Further Research

The CARI technology was used in conjunction with the 2013 SIPP-EHC. The 2012 CARI SIPP-EHC and 2012 SIPP-EHC interviews were combined into the 2013 SIPP-EHC. The plan was to only ask some respondents the consent question. The 2013 SIPP-EHC utilized the CARI system for coaching, behavioral coding, and quality assurance. Additional evaluations will be conducted comparing the 2012 CARI SIPP-EHC to the 2013 SIPP-EHC data quality indicators (Federal Registrar, 2012).

8. Acknowledgements

The work in this article resulted from the contributions of many individuals from the U.S. Census Bureau. The author wishes to thank Candice Barnes, Rachel Bray, Josue Delarosa, Jason Fields, Kenisha Staine, Alicia Walker, Rachael Walsh, members from the Income Survey Branch and Income Survey Programming Branch, and individuals from other divisions/branches involved in the 2012 CARI EHC evaluation planning and data collection stages. I wish to thank Candice Barnes and Tracy Mattingly for their useful suggestions and comments on the paper.

References

- Arceneaux, T. (2007). "Evaluating the Computer Audio-Recording Interviewing (CARI) Household Wellness Study (HWS) Field Test", 2004 Joint Statistical Meetings: Proceedings of the Survey Research Methods Section. American Statistical Association.
- Fields, J. (2011). "The Evolution of the SIPP-EHC". Retrieved from <http://www.dism.ssri.duke.edu/pdfs/Fields.pdf>.
- Nguyen, M., Thissen, M.R., and Siege, B. (2010), "Development of an Integrated CARI Interactive Data Access System for the US Census Bureau", 2010 International Blaise User's Conference, Baltimore, Maryland.
- Schwarz, N., and Wellens, T. (1994). "Cognitive Dynamics of Proxy Responding: The Diverging Perspectives of Actors and Observers". Retrieved from <http://www.census.gov/srd/papers/pdf/sm9407.pdf>.
- Sirkis, R., Snyder, C., and Staine, K. (2012). "Evaluation Plan for the 2012 Computer Audio-Recorded Interviewing (CARI) Survey of Income and Program Participation Event History Calendar (SIPP-EHC) Test".
- U.S. Census Bureau (2009). SIPP Users' Guide Sample Design and Interview Procedures. Retrieved from http://www.census.gov/sipp/usrguide/ch2_nov20.pdf.
- Federal Registrar Notice (2012). Retrieved from <http://www.gpo.gov/fdsys/pkg/FR-2012-09-06/html/2012-21947.htm>.