REDUCING RESPONDENT BURDEN:

Evaluating the Progress of the SIPP-EHC

Rachael M. Walsh, PhD and Jason M. Fields, PhD
U.S. Census Bureau
Social, Economic, and Housing Statistics Division
Longitudinal Research, Outreach, and Evaluation Branch
4600 Silver Hill Road
Washington, DC 20233
rachael.walsh@census.gov, jason.m.fields@census.gov

Disclaimer: The views expressed on statistical, methodological, technical, or operational issues are those of the author and not necessarily those of the U.S. Census Bureau.[1]

Abstract: Under the current redesign plan, the Survey of Income and Program Participation is incorporating an event history calendar (SIPP-EHC). One goal of this redesign is to reduce respondent burden by interviewing annually. Using data from both the 2010 and the 2011 SIPP-EHC field tests, differences in respondent burden are evaluated based on number of question asked as well as interview length. Section timers based on audit trail files are used to evaluate topical sections and sequences where significant changes were made. The design of both the 2010 and 2011 instruments produce skewed individual interview lengths given the first and last person interviewed. In approximately 70% of the households interviewed, this is the same individual thus creating an interesting dynamic that will be accounted for in the evaluation and detailed through a thorough description of the sample. The implications of this research will be used to guide the development of the 2012 SIPP-EHC instrument.

Introduction

The survey community continues to struggle with declining response rates, which is further complicated when conducting longitudinal research (Fitzgerald et al, 1998; Taylor et al, 2005; Watson, 2003; Westat, 2001). Longitudinal surveys rely heavily on respondent compliance and willingness to participate for an extended period of time. Respondent compliance and willingness to participate have been found to be influenced by many interrelated components such as incentives, interviewer experience and continuity, respondent identification with the study, the survey topic, and the interview experience in prior waves (Watson and Wooden, 2009). Decreases in response rates and increases in the restrictions to limit respondent burden has led data collection institutions to revisit current practices, as well as survey design. As a longitudinal survey, the Survey of Income and Program Participation (SIPP) is no exception.

The SIPP is a longitudinal, nationally representative survey of U.S. households. As its title implies, the SIPP collects data on all sources of household income which is then used to demonstrate the dynamic picture of income and wealth distribution as it changes over time. Each SIPP panel begins as a household survey, drawing a sample of between 30,000 and 65,000 households. All initial household members 15 years of age and over are then followed for a period of three to four years.

---

[1] All comparative statements in this report have undergone statistical testing, and, unless otherwise noted, all comparisons are statistically significant at the 5% significance level.

The SIPP is currently undergoing redesign to incorporate an Event History Calendar (EHC) as part of an annual data collection design. Traditionally, the production SIPP is collected through three interviews per year for a duration of three to four years[2]. The redesigned SIPP-EHC will minimize one element of respondent burden by reducing the interviews from three to one per year. To accomplish this, the EHC will be used as a memory aide for the annual reference period (Belli, 1998; Callegaro, Belli, Serrano, and Palmer 2007; Freedman, Thornton, Camburn, Alwin, and Yound-DeMarco 1988), and the core and topical module data previously collected through several interviews per year will be done in one annual interview[3]. As such, the interview length is of particular interest as a measure of respondent burden.

The burden placed on respondents by federal surveys is measured in terms of the type of information collected in a manner comprehendible to respondents through a questionnaire that is of reasonable length (OMB, 2012). As the SIPP-EHC has taken all of its content from the current SIPP interview, the content has been approved by the rigorous OMB process. Additionally, the portion of respondent burden that focuses on the relevancy of the survey content and comprehension of the question wording has already been addressed. However, the merging of the core data collection with the Topical Module content into one annual survey may result in interview length being cumbersome. The purpose of this research is to assess the final element of respondent burden defined by OMB—the length of the questionnaire—in the SIPP-EHC. As the survey is still in the redesign phase, this research also aims to determine whether the changes made to the survey instrument between 2010 and 2011 improved upon the survey by reducing the interview length. To achieve these goals, the following research questions will be addressed:

1. Was respondent burden reduced from 2010 to 2011 by decreasing the number of questions asked of respondents and therefore the interview length;
2. How did the length of specific sections of the instrument change as a result of instrument and training modifications;
3. Did accepting person non-response work as an unsanctioned method to reduce interview length;
4. Where can improvements continue to be made?

Prior to answering these questions, it is necessary to first explain the EHC interviewing, as well as the incorporation of the EHC, into the SIPP.

EHC Interviewing

When expanding the reference period from four to twelve months, memory recall is a concern for data quality. EHC interviewing asks respondents to report information with respect to interrelated topics, referencing the calendar year to improve recall (Freedman, et al, 1988). Different studies show varying results with respect to the length of the reference period, with some finding accurate reporting of detailed information as far back as five years (Freedman, et al, 1988; Caspi, et al, 1996). The EHC method of interviewing relies on memory cues to aid respondents in accessing the autobiographical memory, specifically top-down, sequential, and parallel cuing (Belli, Shay, and Stafford, 2001).

These memory cues rely on the importance of the events in the memory, as well as the time sequencing of occurrence, while helping respondents recall specific events in relation to other coinciding topics (Belli, 1988). Top-down cuing allows the respondent to begin with the most memorable events and progress down to those memories of lesser import (Barsalou, 1988; Conway, 1996). As the EHC is essentially a calendar, sequential cuing probes the autobiographical memory to recall events in a time sequence. Parallel cuing is perhaps the most unique attribute of the EHC in that related events can be used to reconcile gaps in the recollection of events (Baddeley, Lewis, and Nimmo-Smith 1978; Loftus and Marburger 1983). The flexibility and conversational nature of EHC interviewing enables the respondent to tell a narrative of his/her life during the reference period, which provides a

[2] When referring to the SIPP questionnaire that is currently in the field, the term "production" will be used, as that is the version that is currently being used to produce public-use data.
[3] Not all of the Topical Module data that are currently collected in the SIPP are being collected by the SIPP-EHC.

more accurate picture of events and enhances data quality (Brown and Schopflocher, 1998; Schank and Abelson, 1995).

The SIPP-EHC is not solely an EHC interview. Some of the information historically collected by the SIPP either cannot or should not be collected through the EHC format. As such, the SIPP-EHC has both conventional scripted questions, as well as a calendar portion that utilizes the EHC interviewing method. For the purposes of the SIPP-EHC, the calendar portion of the interview contains the following topics: landmark events, residency, marital history, educational enrollment, employment, program participation, and health insurance coverage. Based on the EHC interviewing framework, the SIPP-EHC has undergone three field tests to date and will continue to be tested annually until it replaces the production SIPP in 2014.

The SIPP-EHC Project

The redesign has include several field tests, the first of which was a paper and pencil interview (PAPI) conducted in 2008. As a PAPI will not be replacing the production SIPP, this research focuses on the subsequent field tests conducted in 2010 and 2011. The first field test of a computer assisted personal interview (CAPI) version of the SIPP-EHC was conducted in 2010. At the request of the survey sponsors, the sample for both the 2010 and 2011 field tests purposefully oversampled the low-income stratum to ensure the program participation questions were adequately tested.

The 2010 SIPP-EHC was fielded in six of the twelve Regional Offices with a sample size of 7,982 households, and an 82% household response rate. These household interviews produced 14,738 individual interviews, of which 11,058 were adult interviews. The person non-response rate in 2010 was 11.3%, which is significantly higher than that of the wave 1 geographically and economically match sample from the 2008 production SIPP, where person non-response was 3.2%. Several issues reduced the quality of the data collected in this test, as well as contributed to lengthy interviews.

During training, it was stressed to interviewers that, “This is only a test.” Interviewers were instructed to field the survey to the best of their ability, while remembering that is was only a test. During a debriefing conference with the regional offices, it was brought to light that when the field budget was reached, interviewers were instructed to complete what cases they could, without opening additional cases. While the response rate for the test was equivalent to that of the production SIPP, the person-level non-response was approximately 8% higher in the 2010 SIPP-EHC test than in production SIPP (available through internally accessible data only). When interviews became cumbersomely long, interviewers were instructed to collect what information they could, and then accept person-level non-response for the remaining household members. Additionally, this was the first CAPI version of any SIPP instrument to utilize the Blaise software. As such, there were several “bugs” in the survey instrument, contributing to lengthy interviews, dissatisfied respondents, and higher person non-response rates.

Based on the 2010 field test, the 2011 SIPP-EHC underwent several changes, including the sample from which the interviews were drawn. In 2011, the SIPP-EHC interviewing was done in all 12 Regional Offices, though the sample size was reduced to 4,051 households. The household response rate increased to 85%, obtaining 7,127 individual interviews. Of the total number of individual interviews obtained, 5,345 were adult interviews, and the person non-response rate was 9.8%. The 2011 interviewer training instructed interviewers to collect data as though the SIPP-EHC were any other survey. While the 2011 person non-response rate is an improvement from the 11.3% in 2010, it is still significantly higher than the economically and geographically matched sample for wave 1 of the production SIPP, which was 3.3%.

The 2011 SIPP-EHC instrument included several key modifications that were attempts to reduce respondent burden through shorter interview length, and fewer questions. To reduce the number of questions asked of respondents, an income screener was incorporated and subsequent fields were prefilled in both marital history and residency sections of the instrument. Flow and functionality were enhanced. Additionally, changes were made to utilize reverse-time reporting.

When asking respondents to access time-period specific memory, it is beneficial to begin with the current date and work back through the reference period (Belli, 1998; Martyn & Belli, 2002). As such, in 2011, the EHC was

redesigned to extend the reference period from December of the reference year to the interview month. The time sequencing of questions was modified such that respondents were asked to report about the present, then backward through time to the beginning of the reference period. Not only was this change supposed to aid in the flow of spell collection and completion, but also to enable slightly more scripted transitions from one topic to the next. Census Bureau interviewers are more accustomed to the scripted dialogue, and the interviewer training relies on scripted interviewing when instructing interviewers in data collection processes.

An income screener was created based on family relationships. For the purposes of the income screener only, a family was determined based on the household reference person—that is the person who rents/owns the sample unit. The spouse and children of the household reference person did not have to answer the income screener or program participation questions because the household reference person provided the information for the family. An entire family could be screened out based on annual and then monthly income above 200% of the poverty threshold. To ensure no one was inappropriately screened out, a third question was asked about a range of social welfare program receipt during the reference period. Once screened out, several social welfare program participation questions were taken off path for the entire family.

Data collected from the household reference person were copied onto subsequent household member's calendars for both residency and marital history. If the reference person reported living with any of the members on the household roster, then the residency information provided was copied. Additionally, in households where the reference person reported being married, with the spouse present, the marital history information provided by the household reference person was copied to the spouse's calendar. The copying of previously collected data reduces the number of questions asked of respondents such that individuals with copied data were only required to provide left censored information. While this reduced the number of questions asked of subsequent household members, it further skews the interview length of the household reference person.

With respect to flow and functionality, the SIPP-EHC was substantially improved in that information was recorded in the demographics section of the interview was used to change question text and pre-fill certain responses. This was an attempt to make unnecessary questions verifications by the interviewer instead of re-asking information that had already been provided. Taking these modifications into account, the purpose of this research is to evaluate the progress of the redesign of the SIPP. To do so, internally available data from the 2010 and 2011 SIPP-EHC field tests were used, with a descriptive methodological approach.

Data & Methods

To assess the progress of the redesign of the SIPP, as well as the modifications made to the SIPP-EHC from 2010 to 2011, both the 2010 and 2011 SIPP-EHC test data were accessed[4]. The sample characteristics are displayed in Table 1. To use these data to answer the proposed research questions, descriptive statistics and Wald tests for statistically significant differences were used where applicable.

Number of Questions—Subsequent interview questions are dependent on previously provided information. For questions to come on path for an individual, previously collected data must make the questions relevant. For example, if there are no children who identify the current respondent as a guardian, it does not make sense to ask the respondent child care questions. As such, in this individual's interview, the child care questions would not come on path. The number of questions that come on path for an individual can be summed using unedited data. Questions that were not asked as a result of not being on path for a respondent all receive the same default value, making this portion of the analysis possible.

The household level data collected were added to the total of the first and/or last respondent where applicable. Additionally, the adult well-being and food security questions are only asked of the owner/renter (the household reference person) of the sample unit, so those questions were added to the household reference person's total. This provides an accurate measure of the number of questions asked of everyone in the household, accounting for all questions in the SIPP-EHC interview. This portion of the analysis provides the answer to the first research question,

[4] Please refer to http://www.census.gov/quality/standards/standarde2.html.

assessing the respondent burden through the number of questions asked of each respondent. A Wald test for statistically significant differences between the number of questions asked in 2010 and that in 2011 was used.

Interview Length—To address the second research question—changes in the length of each section of the interview—audit trail files were used to generate and then decompose interview length. Blaise software provides an audit trail file, which is a file that records every keystroke made by the interviewer while in the case. Each keystroke has a time and date stamp. These files were used to divide the interview into section timers for a more detailed analysis. Section timers generate both individual and household interview length, enabling assessment of interview length at both levels, while also permitting the evaluation of each section of the instrument. This also addresses the last research questions—what sections can be improved—by identifying lengthy sections of the instrument, which may need to be adjusted in subsequent iterations of the survey.

The section timers for 2010 and 2011 can be compared, though are subject to a few alterations. The commuting and work schedule questions were moved into the EHC in the 2011 test. To create a comparative measure, the EHC timer was added to the commuting and work schedule timer from 2010. The income screener questions and the dependent care questions were not asked in 2010 but are included here for evaluative purposes.

The availability of section timers also enables the interview decomposition by the content. Some household level information is only asked once, at either the beginning or the end of the interview. As such, the roles of the first and last respondent are more substantial than those who are interviewed in the interim, resulting in an increased interview length for the first and last respondent. Failure to parse out these differences would result in skewed findings. Additionally, knowing which respondents experience the most burden in terms of interview length may help identify ways in which the instrument can be modified to improve the interview experience for all respondents. The specificity provided by audit trail files enables the decomposition of the individual household members into the following, mutually exclusive categories: first respondent, last respondent, both, and neither.

The first respondent—The first person interviewed in the household is responsible for the majority of the information that is provided with respect to the household composition and the sample address. This includes the coverage, roster, and demographics for every member of the household. Once this information is provided, then the first respondent moves into his/her own personal interview.

The last respondent—The last person interviewed is responsible for providing the interviewer with contact information as well as feedback. The contact information collected consists of information about both the household as well as information about one to three non-household member who could be contacted if the household should relocate between waves. As the SIPP-EHC is still in the design phase, the last respondent was also asked to provide opinions about the interview, including ways in which the survey could be improved. While this portion of the interview is not quite as substantial as the introductory portion provided by the first respondent, it is still additional burden that is taken on by this particular respondent.

Once the personal interview is complete for the last respondent in the 2010 instrument, that same person was then asked to complete the contact information as well as the feedback section. In 2011, however, the interviewer was instructed to ask for the household reference person before moving into the final portions of the interview. As such, the last respondent had to be coded differently for 2010 and 2011. The section timers generated through the audit trail files provides the capability to parse out this information based on the line number of the person responding to each section. As such, this was not a threat to the validity of the study.

Both the first and last respondent—In several households the first and last respondent is the same person. In 2010, 72% of households saw the first person also serving as the last person to be interviewed, while this number increased to 73% in 2011. As such, this person's interview length would significantly skew the mean interview length if not addressed. This information can be helpful when assessing the mean interview length at the person level to determine if the collection of the household level information was key to reducing the interview length for these individuals, or if the change in interview length was attributable to the collection of individual information.

Neither first nor last respondent—As seen with the person serving as both the first and last respondent, those households with multiple respondents are going to have people who are neither the first nor the last respondent. As such, these respondents would only respond to the individual portion of the interview and have substantially shorter

interview lengths than either the first or the last respondent. This information was used to determine whether the unsanctioned method of accepting person non-response to minimize interview length was effective.

Person non-response—There are several elements that need to be addressed with respect to person non-response. How and why these respondents differ from those who did respond is outside the scope of this paper. In the 2010 SIPP-EHC, we know from conversations with staff who fielded the survey that the perception of the SIPP-EHC as a lengthy interview resulted in the unsanctioned acceptance of person non-response[5]. This was a frequently used method by interviewers as a means of reducing the interview length while still being able to transmit the case as a "completed" interview[6]. This unsanctioned practice of accepting person non-response was thought by interviewers to shorten household interview length. This research hypothesizes that this, however, was not the case, and the unsanctioned acceptance of person non-response did not in fact reduce the average household interview length.

A simulation was done to test this hypothesis in which all non-responding persons were given the mean interview length for those who were neither the first nor the last respondent (20 minutes in 2010 and 12 minutes in 2011), and the interview length for the household was recalculated. The interview length of an individual who was neither the first nor the last respondent was used, given someone who was a non-respondent could not have served in either role. One person households were excluded from this simulation given in order for person non-response to occur, more than one person must reside at the sample unit. Removing one-person households also increases the mean household interview length.

This simulation demonstrates the difference between the length of a completed household interview and that of a sufficient partial household interview. In both the 2010 and the 2011 test, a sufficient partial interview would be accepted if one respondent in the household completed the calendar portion of the individual interview. The completion of the calendar portion of the interview was determined to be a sufficient partial because it contains the majority of the information with respect to income and program participation. While this is merely a simulation, and each household interview presents its own unique set of data collection difficulties, a generalization may be beneficial to present to interviewers, minimizing the amount of person non-response accepted in subsequent waves of data collection.

Adult Interviews—For the purposes of this research, only those considered as adults remained in sample. The instrument considers respondents over the age of 14 to be adults. Members of the household aged 14 and under do not respond for themselves but proxy information is obtained on his/her behalf by a knowledgeable adult household member. Child interviews are substantially shorter in both length and content. While these interviews could be a measure of the burden to the adult providing the proxy, that is not within the scope of this research and will be done later[7].

Findings

Table 2 displays the number of questions respondents were asked. The number of questions asked significantly increased by eight from 2010 to 2011, despite the addition of the income screener[8]. In both surveys, the number of questions to which the respondent replied with a "Don't Know" or "Refuse" was only three questions, indicating the data reported are useable answers and not another form of non-response.

Tables 3a and 3b display the interview length by section timer. Table 3a displays the average length for the entire adult sample while Table 3b restricts the timers to those who were on path for each section. The modifications to the 2011 SIPP-EHC decreased the length of the following sections: front, calendar, well-being, and length of time

---

[5] The 2011 interviewer training stressed this was not acceptable, resulting in a pattern of person non-response similar to production interviewing (Walsh, 2012).
[6] Interviewers consider a case to be complete when it reaches a point at which it can be transmitted, or a sufficient partial.
[7] Child interviews were included in the household interview length and for the person non-response simulation.
[8] The number of questions on path for an individual does not account for the copying of data because this information still had to be verified by the interviewer.

necessary to switch respondents. The same modifications to the 2011 SIPP-EHC increased the following sections: fertility, annual programs, assets, health care utilization, disability, and the respondent identification policy. There was not a statistical difference in the length of the child roster, child care, and the back of the instrument. The individual interview length decreased by 10 minutes, and the household interview length by 24 minutes. A change in respondents is not reflected in the individual interview length as it cannot be assigned to any one person, however, it is included in the household interview length.

The individual interview length is decomposed by the type of respondent depicted in Table 4. In all cases, the interview length decreased significantly from 2010 to 2011. Table 4 demonstrates the additional commitment from the first and last respondent. In both 2010 and 2011, despite the additional content, the respondent serving as both the first and the last respondent had a decreased interview length than the respondents who served as only the first respondent.

The results from the person non-response simulation are displayed in Table 5, decomposed by household size. In 2010, interviewing everyone in the household would increase the average household interview length by 7.3 minutes, which was the only statistically significant difference seen in this simulation. The decomposition could not be included for households exceeding nine people, given the insufficient sample size; however, none of the simulated interview lengths were statistically different from the average interview lengths. The largest difference seen was in nine person households, where the simulated interview length was 39 minutes longer than the actual interview length.

In 2011, none of the simulated interview lengths were statistically different than the actual interview lengths. As with the 2010 sample, the decomposed simulation could not be included for households exceeding eight people given the insufficient sample size; however, none of the simulated interview lengths were statistically different from the average interview lengths. The most substantial increase was seen in households with eight members, where the simulated interview length was 17 minutes longer than the actual interview length.

Given the significant strides made in 2011, clearly the efforts made by the Census Bureau between the fielding of the 2010 and the 2011 instruments has been beneficial. However, improvements can continue to be made. Some sections of the interview increased in length, as did the overall number of questions asked of respondents. The implications of these results can be used to continue to improve the SIPP-EHC.

Implications

The SIPP-EHC redesign efforts have substantially improved the interview, reducing the respondent burden. From 2010 to 2011, despite the addition of content to the instrument and increase in number of questions asked, the interview length decreased significantly. While some sections of the instrument were actually longer in 2011 than in 2010—fertility, annual programs, health care utilization, and disability—most of the lengthy sections saw significant reductions. The calendar length was significantly reduced, as was front section of the interview.

The calendar portion of the interview most likely benefited from the functionality and flow modifications. These include the timeline resolution to enhance respondent memory recall, as well as the copying of marital history and residency information. Additionally, the use of information provided in the demographics section of the interview to pre-populate fields within the calendar likely contributed to the significant decrease in the length of this section. The enhancements to the functionality of the calendar certainly contributed to the 13 minute reduction seen in this section from 2010 to 2011.

The length of the front section of the interview—that is the rostering and demographics for the entire household and information about the sample unit—saw a significant decrease. The burden of providing this information falls to the first respondent, who consequently also has the longest interview. While this portion of the interview was changed only minimally, the decrease in the length of this section is reflected in the significant decrease in the length of the person serving as the first respondent and those who serve as both the first and last respondent. While it is probable that the person serving as both the first and the last respondent provided proxy information for other household members, this research was not specifically designed to test that theory.

Nor was this research designed to test the differences between the length of the first, last, both the first and the last, and neither the first nor the last respondents. However, this research suggests further research is necessary. Several possible factors could be resulting in the person serving as the first respondent only having lengthier interviews than the person serving as both the first and the last respondent, one of which is the time required to switch respondents. It is not yet clear from the current state of data analysis at which screen in the interview the interviewer is choosing to switch respondents, though this information can be obtained from the audit trail files with further investigation.

We do know switching respondents can take approximately 17 to 18 minutes (refer to Table 3a). To clarify, the time spent switching respondents displayed in Tables 3a and 3b is a function in the instrument when one individual interview is completed and the interviewer is asking for another individual on the household roster to interview. The instrument, when choosing the last respondent, does not perform this same function. In 2010, the last person who completed an individual interview was then the last respondent. In 2011, however, it was preferred that the interviewer ask these questions of the household reference person, though not a requirement. Further work is needed to determine whether the changing of respondents in the end of the instrument is resulting in the delay.

The length of time dedicated to switching respondents—that is moving from one personal interview to the next within the household—is taking considerable time. While this was reduced from 2010 to 2011, this section of the interview is still taking, on average, 17 minutes per household. If the household member serving as both the first and the last respondent is in fact providing proxy responses for other individuals in the household, then this 17 minutes may be avoided. Additional research is necessary to determine if this is in fact the case in households with proxy interviews, generating a new mean for this particular section of the interview for those households without proxy interviews. If this is in fact the case, then the actual time spent switching respondents could be even greater than 17 minutes, and this would be a key place to focus interviewer training to reduce the household interview length.

Additionally, proxy interviews could be contributing to this decrease in time spent switching respondents. It may be possible that the person serving as both the first and the last respondent is also providing proxy interviews for other household members, thus reducing the overall household interview length through the reduction of time spent switching respondents throughout the interview. Again, additional research is necessary to compare cases where this is occurring. The inclusion of the length necessary to proxy for child interviews should also be incorporated in this analysis.

The person non-response simulation, though not a perfect measure, demonstrates that the average household interview length is not reduced through person non-response. While this simulation can be used to demonstrate to interviewers that it is not an effective way in which household interview length can be reduced, it does not take into consideration the situations interviewers are facing in the field. All Americans have busy lives, and it is growing increasingly more difficult to gain access to households to conduct personal interviews (Kiezebrink, et al, 2009). Before assuming high person non-response rates are the fault of the interviewer, additional research is necessary, comparing the interview length of households with person non-response, the interviewer effect in these households, and so on. Regardless of the outcome of future research, during training, interviewers need to see that person non-response is not decreasing the average household interview length.

Based on the outcomes of this research, several improvements are being made to the SIPP-EHC before the 2012 field test. The fertility, child care, and disability blocks of the instrument have all undergone extensive modifications. Additionally, the 2012 SIPP-EHC is the first Wave 2 version of the SIPP-EHC instrument to be tested. With the incorporation of dependent data, other sections of the instrument were also improved, including assets and annual programs. This research will thus be replicated with the 2012 data, with an additional emphasis placed on the use of the dependent data.

The results from this research are encouraging. The positive results of the Census Bureau's efforts to reduce respondent burden in the SIPP-EHC are clearly present as both the interview length and the person non-response are decreasing. Efforts on the part of the Census Bureau and those working to reduce respondent burden are continuing and this analysis should be duplicated with the data from the 2012 test. Additional research is necessary to find the attributes of the survey that are contributing to the variation in interview length. In an ongoing effort to produce the highest quality data in the most efficient manner, the Census Bureau is also conducting research with respect to interviewer effect, regional effect, and the effect of interviewer training (Walsh, 2012). Further research should also examine the differences in the quality of the data between the production instrument, and the new EHC method of

data collection. In anticipation of the switch from the SIPP to the SIPP-EHC as the production instrument in 2014, these early results are encouraging; however, it is also clear that there is much left to do.

References

Baddeley, A. D., Lewis, V., and Nimmo-Smith, I. (1978). When Did You Last...? In Practical Aspects of Memory, ed. Michael M. Gruneberg, Peter E. Morris, and R. N. Sykes, pp. 77-83. London: Academic Press.

Barsalou, L.W. (1988). The Content and Organization of Autobiographical Memories. In Remembering Reconsidered: Ecological and Traditional Approaches to the Study of Memory, ed. Ulric Neisser and Eugene Winograd, pp. 193-243. New York: Cambridge University Press

Belli, R.F. (1998) The Structure of Autobiographical Memory and the Event History Calendar: Potential Improvements. *Quality of Retrospective Reports in Surveys Memory*, 6(4): 383 – 406.

Belli, R.F., Shay, W.L., and Stafford, F.P. (2001). Event History Calendars and Question List Surveys: A Direct Comparison of Interviewing Methods. *Public Opinion Quarterly*, 65(1), 45-74.

Brown, N.R., and Schopflocher, D. (1998). EventC lusters:A n Organizationo f Personal Events in Autobiographical Memory. *Psychological Science*, 9:470-75.

Callegaro, M., Belli, R., Serrano, E. & Palmer, D. (2007). Cultural variability in event history calendar and convention questionnaire interviews: A verbal behavior analysis. In *2006 Proceedings of the American Statistical Association,* 61st Annual conference of the American Association for Public Opinion Research, Alexandria, VA: American Statistical Association.

Caspi, A., Moffitt, T.E., Thornton, A., Freedman, D., Amell, J.W., Harrington, H., Smeijers, J. and Silva. P.A. (1996). The Life History Calendar: A Research and Clinical Assessment Method for Collecting Retrospective Event-History Data. *International Journal of Methods in Psychiatric Research*, 6:101-14.

Conway, M.A. (1996). Autobiographical Knowledge and Autobiographical Memories. In Remembering Our Past: Studies in Autobiographical Memory, ed. David C. Rubin, pp. 67-93. New York: Cambridge University Press.

Fitzgerald, J., Gottschalk, P. and Moffitt, R. (1998), 'An Analysis of Sample Attrition in Panel Data: The Michigan Panel Study of Income Dynamics', *The Journal of Human Resources*, vol. 33, no. 2 (Spring), pp. 251-299.

Freedman, D., Thronton, A., Camburn, D., Alwin, D., & Young-DeMarco, L. (1988). The life history calendar: a technique for collecting retrospective data. *Sociological Methodology,* 18, 37-68.

Kiezebrink,K., Crombie, I.K., Irvine, L., Swanson,V., Power,K., Wrieden,W.L., and Slane, P.W. (2009). Strategies for achieving a high response rate in a home interview survey. *BMU Medical Research Methodology*, 9: 9-46.

Loftus, E.F., and Marburger, W. (1983). Since the Eruption of Mt. St. Helens, Has Anyone Beaten You Up? Improving the Accuracy of Retrospective Reports with Landmark Events. *Memory and Cognition*, 11:114-20.

Martyn, K.K. and Belli, R.F. (2002). Retrospective Data Collection Using Event History Calendars. *Nursing Research*, 51(4): 270-274.

Office of Management and Budget (OMB). (2012). Office of Management and Budget Standards and Guidelines for Statistical Surveys. Viewed at: http://www.whitehouse.gov/sites/default/files/omb/inforeg/statpolicy/standards_stat_surveys.pdf

Schank, Roger C., and Robert P. Abelson. 1995. "Knowledge and Memory: The Real Story." In Advances in Social Cognition (vol. 7), ed. Robert S. Wyer, pp. 1-85. Hillsdale, NJ: Erlbaum.

Taylor, M.F. (ed.) with Brice, J., Buck, N. and Prentice-Lane, E. (2005), *British Household Panel Survey User Manual Volume A: Introduction, Technical Report and Appendices.* Colchester: University of Essex.

Walsh, R. (2012). Interviewer Effect on Conversational Interviews: Results from the 2010 & 2011 SIPP-EHC. To be presented at the Population Association of America Annual Meeting, San Francisco, CA.

Watson, D. (2003), 'Sample Attrition Between Waves 1 and 5 in the European Community Household Panel', *European Sociological Review*, vol. 19, no. 4 (September), pp. 361-378.

Watson, N. and Wooden, M. (2004), 'The HILDA Survey Four Years On', *The Australian Economic Review,* vol. 37, no. 3 (September), pp. 343-349. http://www.iser.essex.ac.uk/files/survey/ulsc/methodological-research/mols-2006/scientific-social-programme/papers/Watson.pdf

Westat, in association with Mathematica Policy Research (2001), *Survey of Income and Program Participation Users' Guide* (3rd ed.), US Department of Commerce, US Census Bureau, Washington DC.

Table 1. Sample Characteristics.

		2010	2011
Households		7982	4051
	Timers Sample	5055	2562
Individuals			
	Total	14738	7127
	Adults	11058	5345
	Non-Respondents	1246	560
	Remaining Adult Sample	9812	4785

Table 2. Number of Questions Asked of Respondents

	2010			2011		
	Mean	Min	Max	Mean	Min	Max
Questions Asked	312.053	189	491	319.995*	178	445
Don't Know	2.094	0	67	2.177	0	69
Refuse	0.982	0	102	1.125	0	123
Total Non-Response	3.076	0	102	3.302	0	124
N		9812			4785	

*indicates statistically significant difference between 2010 and 2011 based on 95% confidence intervals.

Table 3a. SIPP-EHC Section Timers—Adult Sample

	2010 SIPP-EHC		2011 SIPP-EHC		Wald Test
	Minutes	N	Minutes	N	
Front	25.314	5055	22.395	2562	*
Income Screener[1]	--	--	0.982	4785	--
Calendar(+Commuting)[2]	17.944	9812	5.380	4785	*
Fertility	0.511	9812	0.556	4785	**
Dependent Care[3]	--	--	0.097	4785	--
Child Roster	0.050	9812	0.046	4785	***
Child Care	0.320	9812	0.314	4785	***
Annual Programs 1	0.866	9812	1.040	4785	**
Annual Programs 2	1.010	9812	1.337	4785	**
Assets 1	2.116	9812	2.197	4785	**
Assets 2	1.570	9812	2.656	4785	**
Health Care Utilization	2.361	9812	2.690	4785	**
Disability	0.578	9812	0.695	4785	**
Well Being	1.758	9812	1.609	4785	*
RIP	0.178	9812	0.193	4785	**
Back	8.648	5081	8.926	2583	***
Switching Respondents	18.212	5055	17.049	2562	*
Personal Interview	42.313	9812	32.036	4785	*
Household Interview	100.836	5055	76.720	2562	*

[1] This was a new block added in 2011.

[2]This block was moved into the EHC in 2011.

[3] This block was added in 2011.

*indicates 2010 significantly longer; ** indicates 2011 significantly longer; *** indicates no statistically significant difference at the 0.05 level.

Table 3b. SIPP-EHC Section Timers—On Path Only

	2010 SIPP-EHC		2011 SIPP-EHC		Wald Test
	Minutes	N	Minutes	N	
Front	25.334	5051	22.404	2561	*
Income Screener[1]	--	--	1.317	3596	--
Calendar (+Commuting)[2]	18.162	9694	5.483	4695	*
Fertility	0.524	9575	0.574	4634	**
Dependent Care[3]	--	--	0.276	1684	--
Child Roster	0.254	1923	0.249	2844	***
Child Care	2.260	1391	2.215	679	***
Annual Programs 1	0.940	9031	1.070	4651	**
Annual Programs 2	1.038	9544	1.374	4687	**
Assets 1	2.162	9604	2.255	4662	**
Assets 2	1.631	9448	2.698	4710	**
Health Care Utilization	2.422	9566	2.747	4686	**
Disability	0.595	9518	0.709	4690	**
Well Being	3.134	5504	2.838	2713	*
RIP	0.265	6614	0.282	3274	**
Back	8.829	4977	9.102	2533	***
Switching Respondents	18.361	5014	17.123	2551	*
Personal Interview	42.360	9801	32.057	4782	*
Household Interview	100.836	5055	76.750	2561	*

[1] This was a new block added in 2011.

[2]This block was moved into the EHC in 2011.

[3] This block was added in 2011.

*indicates 2010 significantly longer; ** indicates 2011 significantly longer; *** indicates no statistically significant difference at the 0.05 level

Table 4. Individual Interviews by Respondent Type.

	2010		2011		
Interview Type	Length	N	Length	N	Wald Test
First Person	68.64	1482	51.64	708	*
Last Person	35.43	1268	25.01	562	*
Both	43.21	3716	33.27	1962	*
Neither	20.62	3340	12.26	1554	*
All Respondents	38.35	9806	28.20	4786	*

*indicates 2010 significantly longer at the 0.05 level.

Table 5. Person Non-Response Simulation

	2010				2011			
Household Size	N	Interview Length	Simulated Length	Difference	N	Interview Length	Simulated Length	Difference
2	996	92.212	95.701	3.489	677	71.820	73.429	1.609
3	692	108.220	113.958	5.738	411	87.043	90.715	3.672
4	600	128.415	136.124	7.709	325	103.137	106.738	3.601
5	341	147.219	157.673	10.454	228	113.244	118.301	5.057
6	179	163.007	176.620	13.613	100	130.740	139.372	8.632
7	81	175.186	194.963	19.777	45	144.574	155.620	11.046
8	35	182.533	207.675	25.142	22	143.951	160.808	16.857
9	18	196.240	235.098	38.858	†	†	†	†
All Household	2956	118.539	125.813	7.274*	1819	92.534	96.255	3.721

*indicates statistically significant difference between average household interview length and the simulated household interview length based on 95% confidence intervals.

Note: A household size of 1 was not included in this simulation because there could not be person non-response to simulate in a one person household. Household sizes exceeding 9 were not included due to insufficient sample size.

† Insufficient sample size.