"I Don't Remember": Effects of Recall Period on Reported Job and Program Participation Duration

Renée Ellis and Rebecca Chenevert

U.S. Census Bureau[1]

renee.ellis@census.gov

Abstract
In 2006 an initiative was taken to re-design the Survey of Income and Program Participation (SIPP), a major longitudinal survey. This was done in order to reduce costs and respondent burden as well as improve data collection. As part of this re-design effort, the reference period was extended from four months to one year and an Event History Calendar (EHC) was introduced. A major concern in making this change was that the longer recall period would result in fewer reports of spells of program participation and employment, especially during the early months of the recall period. Several studies are ongoing to evaluate the effectiveness of the re-design; this paper seeks to study the effect of the recall period on reported job duration and program participation. Careful evaluation of how well the data in surveys are collected is important for ensuring the accuracy of these empirical studies, and recall is an important component of collecting quality data.

As a result of testing the new survey instrument, a natural experiment has been created where households were randomly assigned to the current production version of the survey or the re-designed survey. Using two years of field test data and the current SIPP, we will exploit the experimental nature of this design to look at the effect of the recall period on the reported job duration and duration of program participation (in particular Food Stamp (SNAP) or WIC participation. This involves carefully mapping the data to a common standard. We can also compare different uses of the EHC for the two years of field testing. We will also examine differences by sex and marital status. We use a proportional hazards model to estimate whether the recall period influences the reported duration of jobs and program participation, and we find that there is little evidence of differences in recall over the two periods.

[1] This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress. The views expressed on methodological or operational issues are those of the authors and not necessarily those of the U.S. Census Bureau.

1. Introduction

The Survey of Income and Program Participation (SIPP) is a major longitudinal survey. The purpose of the survey is to provide information about the income, program participation, and determinants of program participation of individuals and households in the United States. The SIPP offers detailed information on income and transfers on a sub-annual basis, and allows for the evaluation of the effectiveness of federal, state, and local government programs. The survey is currently administered every four months, with a panel lasting approximately three to four years.[2]

In an effort to reduce costs, reduce respondent burden, and improve data quality, the Census Bureau has undertaken a re-engineering of SIPP. The new version of the survey is intended to collect the same quality of information and to maintain the ability to capture intra-year transitions. The new version of the survey, which includes an Event History Calendar (the new survey will be referred to as SIPP-EHC in this paper), is scheduled to replace the current survey instrument in 2014, and will be administered yearly, with a panel lasting approximately three to four years.

The issue we are addressing in this paper is that the longer recall period could potentially result in fewer accounts of events lasting a short period, particularly if the event is early in the year. For example, whether we see fewer changes in jobs or program receipt. However, two iterations of test data have been collected. As a consequence of how the sample was chosen for the test data, we are essentially left with a natural experiment; two groups which should maintain the same properties as potential SIPP respondents were randomly assigned to the current version of SIPP with a four month recall period, or the SIPP-EHC with a one year recall period.[3]

Many studies evaluating the changes in the survey instrument are underway at the Census Bureau. The purpose of this paper is to evaluate whether the changes in the survey instrument have changed the reported duration of a job or participation in WIC (the Special Supplemental Nutrition Program Women, Infants and Children) or SNAP (food stamps) in order to evaluate potential issues with recall. There is a small subtlety to this question; we could compare the differences in reported duration across the two versions of the survey instrument to each other, or we could compare both versions to some third objective measure of "truth," such as administrative records. While there is merit to both approaches, here we employ the former. Similar evaluations with comparisons to administrative records are also underway.

This evaluation is of importance to social scientists that rely on survey data for many related questions; for examples, see Hu and Taber (2011), Kuhn and Sweetman (1999), and Pearce (1990). This work also contributes to a literature on survey evaluation including Brown and Light (1992) and Pierret (2001). We consider appropriate uses of survey data and evaluate how the method of collection may influence variables of interest to social scientists. It is of particular use to those who use the SIPP and are interested in comparability of the redesigned survey with other panels.

Overall, we find very little difference in the reported (completed) duration of a job or spell of program participation on average. However, conditional on duration, there is some evidence that suggests SIPP-EHC respondents are less likely to report a job or spell of program participation ending relative to SIPP respondents with the shorter recall period. There is also evidence that in the first iteration of test data that males and females responded differently to the different survey instruments when reporting jobs, although this is not true in the second test. Similarly, marital

[2] All data are subject to error arising from a variety of sources, including sampling error, nonsampling error, model error and any other sources of error.

All comparative statements in this report have undergone statistical testing, and, unless otherwise noted, all comparisons are statistically significant at the 10 percent significance level.

[3] This analysis is similar to that of Pierret (2001) although the analysis here has comparable samples while his analysis uses two reports from the same individual.

status was important in the first test but not in the second, so it seems that the second round of SIPP-EHC test data is more comparable to the production SIPP data.

The section below details the two surveys and the similarities and differences between them. Section Three discusses how we made the datasets comparable and the methods used for evaluation. Section Four presents results, and Section Five concludes.

2. Background

2.1. SIPP

As mentioned, the SIPP is a major longitudinal survey focused on collecting information relating to income and program participation of individuals and households. Because of the focus on government programs, the survey oversamples areas with high poverty. However, weights are provided and nationally representative (and some state representative) statistics can be calculated. The first data collection for SIPP began in 1983, and the survey underwent a major redesign in 1996. In its current form, the SIPP has a set of "core" questions which are asked every four-month wave and "topical modules" which focus on different topics and are not administered with each wave. Some of these topical modules are repeating (for example, information on assets is collected once a year or every three waves) while others are only administered once (such as marital history).[4]

Every ten years after the decennial census, several SIPP sample designations are created with the particular properties desired of a SIPP sample. Each sample designation is also divided into 101 reduction groups; these reduction groups are constructed in such a way that their inclusion or exclusion from the sample should not alter these desired sample properties. With the creation of the 2008 panel of SIPP, there was a desire to have state-representative poverty rates for the twenty most populous states. In order to accommodate this request, the 2008 panel of SIPP was created by including one complete sample designation and supplementing it with reduction groups from some of these states from a second sample designation. However, as will be described below, the creation of the sample for SIPP-EHC was different so this will be taken into account in the analysis which follows.

SIPP interviews are conducted in person, with Field Representatives (FRs) using laptops to record responses (this is known as computer assisted personal interviewing, or CAPI). Follow-up interviews may be conducted over the phone. Within a panel, respondents are assigned to be in one of four different rotation groups. These rotation groups determine when the interview will take place and what the calendar months of the reference period are. For example, wave 1 interviews for rotation group 1 took place in September of 2008 so the reference period included May through August of 2008. For rotation group 2, the interviews took place a month later so their four month reference period included June through September. This pattern is repeated for rotation groups 3 and 4. The first rotation group was interviewed again as part of wave 2 in January of 2009, so the reference period for wave two included September through December of 2008. Therefore, when matching SIPP waves to a calendar year we must include four waves of data.[5]

2.2. Re-Engineered SIPP

Due to a projected budget shortfall in 2006, the Census Bureau opted to eliminate the SIPP program. However, this decision was met with strong protest, so Congress decided to fund the continuation of SIPP conditional on re-designing the survey to lower costs. A key factor in this redesign is the shift away from a model where respondents are interviewed every four months to one where respondents are interviewed once per year. In order to maintain the intra-year transitions and short spells which are an important part of SIPP, an Event History Calendar was

[4] For topical module listings for each panel, see http://www.census.gov/sipp/top_mod/top_mods_chart.html. For further information on SIPP, see http://www.census.gov/sipp/technical.html.

[5] For more documentation on how SIPP data are collected, see http://www.census.gov/sipp/technical.html.

introduced to help respondents with recall. An Event History Calendar presents respondents with a calendar and asks them to report "landmarks" that occur throughout the year which can help them remember. This also allows for a more flexible and conversational interview. For example, if a respondent is having a difficult time remembering when they started a job, the FR is able to probe with questions such as "Was it before or after your birthday in March?" This can help respondents retrieve more accurate information. While landmark events can be anything that helps a respondent put a date on an event, they are usually events such as weddings, birthdays, holidays, and trips.

Two iterations of testing the new SIPP-EHC survey have been completed; the first was administered in early 2010 with the reference period of calendar year 2009, and the second was administered a year later with 2010 as the reference year. Both surveys were "wave 1 interviews," meaning the 2011 respondents were not the same respondents as 2010 being re-interviewed. In early 2012 there are scheduled re-interviews of the 2011 respondents. Because the structure of core and topical module questions is no longer preserved, the survey has been reorganized in order to accommodate the addition of topical module questions to the standard interview. However, the same topic areas are covered and many of the same questions are maintained.

In order to focus the testing on groups of interest, the sample for the 2010 field test was modified from a traditional SIPP sample. Two SIPP sample designations were used, so that the properties of a SIPP sample are maintained for the new survey. However, in 2010 the only addresses that were visited included those in the high poverty strata, self representing PSUs, and a particular set of states served by certain regional offices. In the 2011 data collection, the sample was expanded to all regional offices and more states, but the sample was still restricted similar to what was done in 2010. There are two important points to note here. First is that this sample is very particular, and therefore none of the sample properties calculated in this paper should be interpreted as nationally representative. Second, because the SIPP sample for the 2008 panel was not restricted in the same ways as the SIPP-EHC sample, we must employ the same sample restrictions to the SIPP data in order to make comparisons between the surveys. This is discussed below.

3. Data Construction

In order to make valid comparisons between the surveys, we need to have comparable samples and data. The following section discusses differences between SIPP and SIPP-EHC that needed to be reconciled in order to have comparable data. This data construction involved multiple steps including the selection of the subset of SIPP respondents, the use of edited versus unedited data, selection of job for respondents with multiple jobs, selection of spell of program participation for respondents with multiple spells, and construction of begin and end date variables. At several points during data construction there was more than one viable choice; for example, we could choose to analyze a single job spell or multiple job spells for each respondent. In these situations we chose the option that would make the cleanest possible analysis of the data and was most clearly related to our question of interest. Although there were many valid ways to construct these data, we believe our construction is as accurate a comparison as possible given the differences in data collection and structure.

3.1 Data and sample selection

As a part of data processing, the SIPP data are "edited" to maintain consistency as well as impute responses for individuals who did not give a valid response for some items on the survey (e.g., maintaining consistency could include reconciling responses given by two spouses or maintaining temporal consistency). However, as SIPP-EHC is still in the design phase, only basic demographic data have been edited. Therefore, we use unedited SIPP data for non-demographic variables to maintain comparability. See the next section for a description of the variables used in our analysis.

Because the 2008 SIPP panel is an ongoing nationally representative sample and the SIPP-EHC tests were limited samples in select low income strata, the SIPP data needed to be subset to be comparable. We compared the 2010 and 2011 field test data to the subset of SIPP data from the 2008 panel during the respective calendar years; this subset

was restricted to respondents who lived in one of the geographies specified at the time their household was rostered.[6] An argument could be made for comparing SIPP-EHC data to the subset of households which satisfied the sample restrictions in early 2010 and 2011, respectively; however, because people and households move, this method would be problematic so we included households which satisfied the sample restrictions at the time rostering took place.[7]

In SIPP and SIPP-EHC, records exist for individuals that were in the household at the time of interview but did not complete an interview. In SIPP-EHC, job spell and program participation information is not available for these individuals. They did not meet the sample criteria (described below in Section 3.2) so they were not included in the sample. In SIPP these respondents have imputed spells if they were available in previous waves. Because these respondents lack reported data for the reference year, they were removed from the comparison sub-sample.

3.2 Variables and Variable Construction

Analysis required construction of variables for jobs, WIC, and SNAP. Construction of variables and analysis for jobs differed from construction of variables and analysis for programs. For this reason many of the following sections are divided into a section discussing jobs and a section discussing program participation. The labels for these sections include an A for jobs and a B for programs.

3.2.A Variables and Variable Construction for Jobs

Both SIPP and SIPP-EHC are output as person-month files and include data on multiple jobs. This means that duration analysis must be done carefully; we cannot just look at duration for "the" job for an individual, since there may be multiple jobs held concurrently or over the course of the calendar year. Therefore, we chose to employ a particular strategy. We defined the main job in January of 2009 in SIPP and the 2010 SIPP-EHC Field test, and we defined the main job in January of 2010 in SIPP and the 2011 SIPP-EHC Field test. The main job was defined by the job where the most hours were worked in January. Because we use unedited data, we treat responses of "don't know," "refuse," and "hours vary" (SIPP only) as if they worked zero hours; in other words, a job with such a response will only be considered the main job if there is no other job with positive hours reported. We focus on the job held in January because a major concern that arises from redesigning the survey stems from recall issues; we want to ensure that we are capturing jobs and information about those jobs that occur early in the calendar year. We then follow the job through calendar year 2009 or 2010 (respectively) and consider whether the probability of the job being reported as ending each month differs between the two surveys, conditional on the current duration.

There were also some changes that were made to the variables themselves to make a valid comparison of beginning and ending dates. In order to reduce respondent burden, in SIPP individuals employed at the time of their first interview are asked the year that the job began. If the job began more than two years prior to the reference period, no further information about the start date is requested. If the job began more recently, the month and possibly the day the job began are asked.[8] In the 2010 Field Test of SIPP-EHC individuals were asked the complete start date for all jobs. In 2011 the question changed to asking start week for any jobs that began during the reference period, but

[6] In the 2010 test, the geographies were limited to California, Connecticut, Illinois, Louisiana, Massachusetts, Maryland, New York, Rhode Island, Texas, and Wisconsin served by the Boston, Chicago, Dallas, Los Angeles, Philadelphia, or New York Regional Offices. In 2011, the test sample was expanded to include Arizona, California, Connecticut, Florida, Georgia, Illinois, Indiana, Louisiana, Maryland, Massachusetts, Michigan, Missouri, New Jersey, New York, Ohio, Oklahoma, Pennsylvania, Tennessee, Texas, and Virginia, and all 12 regional offices.

[7] See "An Initial Evaluation of the 2010 Field Test of Re-Engineered SIPP" for more information about how individuals and households could differ due to differences in the time of rostering.

[8] The start day is asked for those jobs that began within the interview year. However, this analysis is conducted at the monthly level so this difference in data collection is not a concern here.

still asked start date for jobs beginning prior to the reference period. In order to measure job duration in months, we have assigned a value of '6' for the beginning month of any job that began prior to 2006, whether the job was reported in SIPP or SIPP-EHC. Also, as a crude imputation, any individual who reported a start year for a job but responded "don't know" or "refuse" for the start month was also assigned a value of '6' for start month. Those jobs which had no value for start year were dropped in both SIPP and SIPP-EHC due to issues of evaluating left censored observations. As a robustness check for this restriction, the same models were run only using jobs that began in 2006 or later; results follow similar patterns as reported in Tables 2 and 3.

In the current SIPP survey, those who report being self employed are taken down a different path to an entirely different set of questions than those who work for an employer. Data are reported for up to two jobs and up to two businesses. In SIPP-EHC many of the questions are different for the self employed, but they are still treated as holding a job. The SIPP-EHC allowed a respondent to record up to five jobs or businesses in 2010 and seven in 2011. The differences in this section of the survey are not problematic here, however. Both SIPP and SIPP-EHC ask for the regular hours worked at a job or business, industry, occupation, and end dates in the same way for both jobs and businesses. Therefore, we do not make adjustments to these variables.

As discussed above, the sample was limited to respondents that had a job or business in January of the calendar year. The sample size for calendar year 2009 was 63,682 person-months representing 6,747 individuals. Roughly 36 percent of these individuals were SIPP respondents and 64 percent were SIPP-EHC respondents. The sample size for calendar year 2010 was 49,370 person-months, representing 5,420 individuals. This sample included 62 percent SIPP respondents and 38 percent SIPP-EHC respondents.

In addition to beginning and ending dates of respondent's main job, analysis includes: a variable indicating if a respondent was a participant in SIPP or SIPP-EHC, a month variable indicating the month of end of job, a variable indicating the end of the interview reference period (also known as a seam month, described below), and two demographic control variables: sex and marital status.

Because start and end dates for jobs tend to be clumped at the beginning and end of the four month reference period in SIPP, a variable was created to control for this tendency. This is a "seam month" indicator. In SIPP this variable is set to '1' in the month that ends the respondents reference period (this varies by rotation group- see above information on SIPP in background section 2.1) and '0' in other months. In SIPP-EHC this variable is 1 in month 12 and 0 in other months.

We include two demographic control variables in this analysis because respondent burden may vary by demographic characteristics, and this may in turn impact recall or failure to include job spells. We use edited versions of these variables because there are edited versions in both SIPP and SIPP-EHC. Edited variables have been checked for consistency. The two demographic variables used, sex and marital status, are asked identically on both surveys so we assume that they are directly comparable. There are other demographic variables that could be valuable for this analysis such as household size and age, but they are either not directly comparable or are not edited and therefore will be included in future analysis.

3.2.B Variables and Variable Construction for Program Participation

There are some differences between the questions on jobs and the questions on program participation that resulted in slightly different approaches to variable construction. First, because of changes made to the universes of who could respond to program questions the SIPP-EHC 2011 data was not as readily comparable to SIPP 2008 as the data from SIPP-EHC 2010. Although we have plans to analyze the SIPP-EHC 2011 program data in the future, this paper only includes 2010 analysis.

Second, start dates for a main job held in January but started before the calendar year were available. This allowed us to overcome left censoring of jobs. This was not the case for WIC or SNAP. To avoid left censoring we include only respondents that reported starting SNAP or WIC during the 2009 calendar year.

This created a very small sample, in order to conduct meaningful analysis we combined SNAP and WIC participants into one sample. We examine only the first spell of participation in a program. A respondent was only included if they had a begin date. If a participant had both SNAP and WIC participation the SNAP participation was used. WIC participation tends to have fewer breaks than SNAP because of less frequent recertification. Because we were interested in transitions out of programs we chose to use a SNAP report if available and add WIC cases only if respondents did not report receiving SNAP.

The sample size of the combined SNAP/ WIC participants for calendar year 2009 was 756 person-months representing 164 individuals. Roughly 48 percent of these individuals were SIPP respondents and 52 percent were SIPP-EHC respondents. Because of the small sample size we were unable to include month variables. We were able to include the seam variable, marital status, and sex. See the above description in jobs for the construction of these variables

3.3 Methods

This study uses duration models to analyze differences in the reporting of transitions in a 1 year period. For jobs, this includes two sets of comparisons: SIPP 2008 Waves 2-5 is compared to SIPP-EHC 2010 and SIPP 2008 Waves 5-8 is compared to SIPP-EHC 2011. These cover calendar years 2009 and 2010, respectively. For program participation, analysis only includes calendar year 2009. This is because of differences in collection of program participation data in SIPP-EHC 2010 that have not yet been made comparable to SIPP 2008 data. Start time is the year and month the job or participation in program began, end date is the year and month the job or participation in program ended or the end of the calendar year, and failure is the end of the job or participation in program. This allows us to compare the probability that a respondent in either survey reports an ending, conditional on it having lasted until January of the respective calendar year.

It is expected that there is some measurement error associated with survey data. However, if we think that this error is the same in both the SIPP and the SIPP-EHC, then we know that we can essentially think of the surveys in the same way going forward. The concerns that we are addressing here are twofold. First, we are concerned that a job or program participation spell which was over early in the year may be less likely to be reported in SIPP-EHC, and therefore fewer spells with short durations may be recorded. For this reason we focus on the main job in January and the first reported spell of participation. The second concern is that there may be more bias towards reporting a later ending date in the SIPP-EHC than in SIPP due to the longer reference period. The proportional hazards model can give us some insight into these questions. Both of these concerns may lead to the hazard of ending a spell being lower in SIPP-EHC than in SIPP.

The regression model we use is a proportional hazards or Cox model. We can define the hazard function as $\lambda\ t\ = \lim_{\Delta t \to 0} \Pr(t \le T + \Delta t | T \ge t)\ /\Delta t$ where T is the failure time (end of the job or program participation, in this case). We can similarly define the survivor function $S\ t\ = \Pr\ T > t\ $, and the c.d.f. $F\ t\ = 1 - S\ t\ = {}_0^t f\ s\ ds$ where $f\ t\ = dF(t)/dt$. We can then see that the hazard function can also be defined as $\lambda\ t\ = \frac{f(t)}{S(t)}$. In the proportional hazards model we make the simplifying assumption that the hazard function can be broken down into two pieces, $\lambda\ t|x\ = \lambda_0\ t, \alpha\ \phi\ x, \beta\ $ where $\lambda_0\ t, \alpha\ $ is called the baseline hazard. Therefore, the proportional hazards model looks at the likelihood of failure as a ratio of the hazard rates in two different states; in this study, we are most concerned with the hazard of a job or participation in program ending when the respondent is a member of the SIPP-EHC observation group relative to the SIPP. This model is flexible because it allows us to study the effect of a particular variable (indicator for SIPP-EHC) on the probability of reported failure without have to specify an underlying structure of the hazard model. In other words, the baseline hazard $\lambda_0\ t, \alpha\ $ is not estimated but is assumed constant for SIPP and SIPP-EHC respondents. The partial likelihood function is the joint product of $\Pr(T_j = t_j | j \in R(t_j))$ over k ordered failure times. Coefficients are estimated by minimizing the log partial likelihood function, given by $\ln L_p\ \beta\ = \ {}_{i=1}^{N} \delta_i\ \ln\phi\ x_i, \beta\ - \ln\ \ {}_{l \in R(t_i)} \phi(x_l, \beta)\ $ where δ_i is an indicator for

whether the observation is right censored and *R(t)* is the risk set at time *t*. We further assume that the specification of the function $\phi(x, \beta)$ is exponential to make the likelihood function more tractable.

3.3.A Methods- Jobs

For each of the two comparisons four duration models were run: 1) a baseline model that compares the duration of transitions in SIPP to the durations in SIPP-EHC; 2) A model that includes indicator variables in the seam months, corresponding to the last month of the reference period for each survey and rotation group; 3) A model that includes interaction effects with month of reported end of job in order to assess possible memory effects of a longer recall period; and 4) A model that includes both these interactions and the seam indicator variable. We also consider several models with sex and marital status as well as their interactions with month indicators and survey indicators to allow for the possibility that recall may be different across genders or that marital status may be important. Marital status could be important in one of two ways; having a spouse present at the interview may cause better recall (two heads are better than one, in a sense) or it could cause the interview to last longer, and possibly decrease the respondents' willingness to cooperate.

3.3.B Methods- Programs

Similar models were run for the combined WIC and SNAP samples for the 2009 SIPP-EHC. Because of sample size, the month variables were not included in the analysis. The following four models were run 1) a baseline model that compares the duration of transitions in SIPP to the durations in SIPP-EHC; 2) A model that includes indicator variables in the seam months, corresponding to the last month of the reference period for each survey and rotation group; 3) A model that includes interaction effects with sex and a seam indicator; and 4) A model that includes interaction effects with marital status and a seam indicator.

4. Results

4.A Results Jobs

In calendar year 2009, 732 respondents ended a job. Below, Table 1 shows the average durations of jobs that ended for each year of observation; this does not account for censoring or any of the controls discussed above.

Table 1: Durations for Jobs and spells of WIC or SNAP receipt that Ended[9]

		Average Duration (in months)	Std Dev of Duration
Jobs Calendar Year 2009	SIPP	46.7	65.0
	SIPP-EHC	41.3	71.4
Jobs Calendar Year 2010	SIPP	41.6	74.4
	SIPP-EHC	41.3	64.9
WIC/SNAP Calendar Year 2010	SIPP	4.1	2.5
	SIPP-EHC	3.9	2.6

While this is encouraging, we have a small sample of jobs that are reported to have ended. Researchers are not only interested in the overall duration of a job, but also in the probability that a job ends conditional on it having lasted as long as it has. Using this information, we can study job tenure and look at causal relationships for jobs ending or not

[9] Source: U.S. Census Bureau, Survey of Income and Program Participation (SIPP), 2008 Panel Waves 2-8; SIPP – Event History Calendar (SIPP-EHC), 2010 Panel; SIPP-EHC, 2011 Panel. For information on sampling and nonsampling error see http://www.census.gov/sipp/source.html.

ending due to particular events. Therefore, we use a proportional hazards model to study the conditional probability that a job is reported ending in calendar year 2009 and 2010 for each set of test data, respectively. We condition on the reported duration of the job and the survey method used in all specifications. In all the tables that follow, we present transformed coefficients. These are best interpreted as the ratio of probabilities between the two surveys. The numbers below in parentheses are p-values.

We find mixed results. In both tables, in columns (3) and (4) a nonlinear time trend was included (though not reported) by using month indicator variables. As reported in Table 2, we also see that the probability of reporting a job ending in SIPP-EHC (relative to SIPP) is lower at the beginning of the year and higher later on in the year, even after accounting for the general time trend. While this is somewhat indicative of a recall problem, the trend is not monotonically increasing as the year goes on. Therefore, it could be that SIPP-EHC respondents tend to group the responses of reported ending dates into certain months. Given the landmark structure of the EHC, this could be due to where landmark events tend to fall. This would indicate that there was possibly a problem with recall present in the 2010 data. While the overall model suggests that conditional job duration is not different between the two surveys, we are finding that jobs are reported as ending more often in particular months in SIPP-EHC relative to SIPP.

Table 2: Job Proportional Hazard Models: 2010 Data[10]

VARIABLES	(1)	(2)	(3)	(4)
SIPP-EHC	1.002	1.085	0.675	0.618**
	(0.980)	(0.278)	(0.109)	(0.0496)
Seam Month		1.466***	1.351**	
		(0.000178)	(0.0199)	
SIPP-EHC*Mar			1.862*	1.858*
			(0.0528)	(0.0540)
SIPP-EHC*Apr			1.237	1.213
			(0.534)	(0.571)
SIPP-EHC*May			1.669	1.664
			(0.131)	(0.134)
SIPP-EHC*June			2.214**	2.236**
			(0.0155)	(0.0145)
SIPP-EHC*July			1.563	1.560
			(0.209)	(0.211)
SIPP-EHC*Aug			1.300	1.282
			(0.445)	(0.471)
SIPP-EHC*Sept			1.436	1.438
			(0.292)	(0.291)
SIPP-EHC*Oct			1.974*	2.005*
			(0.0555)	(0.0503)
SIPP-EHC*Nov			1.383	1.389
			(0.387)	(0.381)
SIPP-EHC*Dec			1.730	2.302**
			(0.131)	(0.0180)
Month Indicator Variables	No	No	Yes	Yes
Observations	63,682	63,682	63,682	63,682

p-value in parentheses*** p<0.01, ** p<0.05, * p<0.1

[10] Source: U.S. Census Bureau, Survey of Income and Program Participation (SIPP), 2008 Panel Waves 2-5; SIPP – Event History Calendar (SIPP-EHC), 2010 Panel.

Reconciling the results from the two years of test data is not straightforward. Several changes were instituted between fielding the 2010 and 2011 tests of SIPP-EHC. The major change in the section reporting on jobs was in the direction of recall; 2010, they had been asked to start recall with January 2009 and move forward. In 2011, respondents were asked about their current employment status and then were to go "backwards" in time describing their status throughout the year. This was done to improve recall. Table 3 shows that the data from the 2011 test were different than that of the 2010 test. There were also changes in training the field representatives between 2010 and 2011, which could have resulted in better data collection for the second field test. We now see an overall model where jobs reported in SIPP-EHC are 0.84 times as likely to have a reported end versus jobs reported in SIPP, conditional on the duration of the job up until that point. However, we no longer see a pattern where jobs are more likely to be reported ending near the end of the calendar year in the 2011 test data, or in any particular month in SIPP-EHC relative to SIPP; none of the coefficients on the interaction variables are significant.

Table 3: Job Proportional Hazard Models, 2011 Data[11]

VARIABLES	(1)	(2)	(3)	(4)
SIPP-EHC	0.840**	0.897	0.736	0.691
	(0.0255)	(0.184)	(0.254)	(0.162)
Seam Month		1.381***	1.259**	
		(0.00127)	(0.0392)	
SIPP-EHC*Mar			1.004	0.999
			(0.991)	(0.998)
SIPP-EHC*Apr			1.381	1.373
			(0.379)	(0.387)
SIPP-EHC*May			0.919	0.915
			(0.821)	(0.811)
SIPP-EHC*June			1.196	1.198
			(0.609)	(0.604)
SIPP-EHC*July			1.296	1.292
			(0.509)	(0.513)
SIPP-EHC*Aug			0.895	0.891
			(0.765)	(0.757)
SIPP-EHC*Sept			1.535	1.529
			(0.282)	(0.286)
SIPP-EHC*Oct			1.399	1.402
			(0.409)	(0.405)
SIPP-EHC*Nov			1.372	1.370
			(0.423)	(0.424)
SIPP-EHC*Dec			1.244	1.563
			(0.573)	(0.222)
Month Indicator Variables	No	No	Yes	Yes
Observations	49,370	49,370	49,370	49,370

p-value in parentheses
*** p<0.01, ** p<0.05, * p<0.1

[11] Source: U.S. Census Bureau, Survey of Income and Program Participation (SIPP), 2008 Panel Waves 5-8; SIPP-EHC 2011 Panel.

In both years of test data, we find that the inclusion of the indicator variable for a seam month is significant. Jobs are more likely to be reported as ending in the last month of the reference period. However, this concept may be more relevant for the SIPP respondents than for the SIPP-EHC respondents. While the reference period for SIPP-EHC was calendar year 2009 or 2010 for the respective years of data collection, most interviews took place between mid- January and early March of the following year. So while the seam month is December for both years of data collection, jobs could have ended later than that and could have been reported as such in SIPP-EHC. Therefore, what we may be seeing in this data is a tendency of respondents to report jobs ending on seam months, and the seam months are more spread out in SIPP-EHC relative to SIPP. This is something that would be worth exploring, once the wave 2 version of SIPP-EHC is fielded in early 2012.[12]

The other difference between SIPP and SIPP-EHC worth mentioning is that the respondents are coming from different waves of the panel. There is a large amount of attrition between the first and second waves of a SIPP panel. Therefore, the composition of respondents is different because we are only including non-attriters in SIPP while we have a wave 1 sample in SIPP-EHC. These respondents may be more stable in general, because we are able to find them for follow-up surveys and they are more willing to participate. Non-attriters also tend to have a higher attachment to the labor market, so average job duration might be longer for those completing later wave interviews (see Vaughn and Scheuren or Zabel, 1998). This could lead to the SIPP respondents being less likely to report a job end overall than SIPP-EHC respondents, which would bias the coefficient on the SIPP-EHC variable upwards. Therefore, the results reported here could be underestimated. However, we may find a more comparable sample of nonattriters with the re-interviews in 2012, which we could use to test this explanation. The individuals who are reporting job dates in the SIPP can also be different in another critical way. The start date of the job is asked in the first interview where the respondent mentions a job. For SIPP respondents, this could be as early as 2008. Therefore, for all jobs that began before 2008 that are in the analysis sample, the question about the start date was asked approximately two years earlier for respondents in SIPP than in SIPP-EHC.

As mentioned in Section 3, we have also considered a few demographic variables which may influence the performance of the two survey instruments. We allow here for the possibility that recall is different between men and women. Again, we find mixed results. The 2010 data seem to show that there is an increased probability that women report job endings in the SIPP-EHC relative to SIPP (see column (2) in Table 4); however, this is not reinforced in the 2011 data. It is also worth noting that with full interactions between sex indicators, SIPP-EHC indicators and month indicators we have 40 control variables in this specification. With only 732 job endings reported in calendar year 2009, these results may be spurious due to sample size.

Similarly, for marital status we find that married individuals are less likely to report job ends in the 2010 SIPP-EHC data but this is also not reinforced in 2011. This could be due to a decrease in respondent burden; between 2010 and 2011 the interview time decreased by an average of 17 minutes.

[12] As a check for the robustness of these results, we ran the models separately for the self employed and those employed for an employer. Although the sample size was small for the self employed, we generally find the same patterns as presented here and so those results are not reported.

Table 4: Proportional Hazards Models by Demographics, 2010 Data[13]

VARIABLES	(1)	(2)	(3)	(4)
SIPP-EHC	0.940	0.394***	1.150	0.689
	(0.533)	(0.00500)	(0.118)	(0.202)
Female	0.927	0.749		
	(0.523)	(0.353)		
SIPP-EHC*Female	1.143	2.374**		
	(0.330)	(0.0358)		
Married			1.089	1.047
			(0.486)	(0.881)
SIPP-EHC*Married			0.681***	0.721
			(0.00684)	(0.433)
Month Dummies Included	No	Yes	No	Yes
Interactions with Month Included	No	Yes	No	Yes
SIPP-EHC#Month Interactions Included	No	Yes	No	Yes
Observations	63,682	63,682	63,649	63,649

p-value in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 5: Proportional Hazards Models with Demographics, 2011 Data[14]

VARIABLES	(1)	(2)	(3)	(4)
SIPP-EHC	0.862	0.774	0.813**	0.544*
	(0.181)	(0.464)	(0.0286)	(0.0662)
Female	1.137	0.982		
	(0.221)	(0.952)		
SIPP-EHC*Female	0.950	0.796		
	(0.735)	(0.664)		
Married			0.718***	0.469**
			(0.00350)	(0.0347)
SIPP-EHC*Married			1.111	2.243
			(0.523)	(0.150)
Month Dummies Included	No	Yes	No	Yes
Interactions with Month Included	No	Yes	No	Yes
SIPP-EHC*Month Interactions Included	No	Yes	No	Yes
Observations	49,370	49,370	49,370	49,370

p-value in parentheses
*** p<0.01, ** p<0.05, * p<0.1

[13] Sources: U.S. Census Bureau, Survey of Income and Program Participation (SIPP), 2008 Panel Waves 2-5; SIPP – Event History Calendar (SIPP-EHC), 2010 Panel.

[14] Sources: U.S. Census Bureau, Survey of Income and Program Participation (SIPP), 2008 Panel Waves 5-8; SIPP – Event History Calendar (SIPP-EHC), 2011 Panel.

4.B Results for Program Participation

In calendar year 2009, 35 respondents ended a spell of participation in SNAP or WIC. Table 1 shows the average durations of participation in programs that ended for each year of observation; this does not account for censoring or any of the controls discussed above. Average months of program participation was 4.1 for SIPP and 3.9 for SIPP-EHC. This is encouraging. However, as in the models for jobs, the proportional hazard models for programs show the probability of respondents reporting a program ending is lower in SIPP-EHC (see table 6). This is true across all four models.

As in the models for jobs, the seam month indicator is significant. However, the direction is opposite, this indicates that it is more likely that SIPP-EHC respondents reported ending programs the last month of the calendar year than SIPP respondents reported ending during an interview month. There is no significant effect in the interactions with sex or marital status.

Table 6: Food Stamp and WIC Duration Models: 2010 Data[15]

VARIABLES	(1)	(2)	(3)	(4)
SIPP-EHC	0.344***	0.369***	0.241**	0.442*
	(0.00146)	(0.00290)	(0.0377)	(0.0601)
Seam Month		0.229*	0.233*	0.239
		(0.0941)	(0.0996)	(0.110)
Female			0.749	
			(0.283)	
SIPP-EHC*Female			1.844	
			(0.435)	
Married				1.805**
				(0.0204)
SIPP-EHC*Married				0.721
				(0.646)
Observations	756	756	756	756

p-value in parentheses *** p<0.01, ** p<0.05, * p<0.1

5. Conclusions

We find that in the two surveys with different recall periods, the probability of reported job and program participation endings is lower in the SIPP-EHC with a longer recall period than in SIPP during the same calendar years. This could be due to the longer recall period. However, further testing is warranted to clearly determine the cause of lower reports of jobs and program participation ending in SIPP-EHC. An upcoming third test of the SIPP-EHC survey will be a re-interview of respondents that were questioned in the 2011 test. This will be more similar to the SIPP data that we are using as a comparison. Additionally, changes to the 2012 survey and training may help improve data collection.

There are some differences between the 2010 and 2011 SIPP-EHC comparisons of jobs; in the 2010 data (referring to calendar year 2009) we found that there was generally an increase in the probability of a reported job ending in SIPP-EHC in particular months, whereas in the 2011 data we do not see such a pattern. This could indicate that the changes in the way that data were collected (going forwards in time in 2010 vs. backwards in 2011) improved recall

[15] Sources: U.S. Census Bureau, Survey of Income and Program Participation (SIPP), 2008 Panel Waves 5-8; SIPP – Event History Calendar (SIPP-EHC), 2011 Panel.

in the earlier part of the year. However, the differences could also be related to the different geographies included in the tests between the two years or the location of landmark events. The 2012 test will use the same geography as the 2011 test and could help eliminate geography as an explanation for this difference.

The two demographic characteristics included in our models found little evidence that the sexes respond differently to the two surveys or that married and unmarried individuals respond differently. There is some indication in the 2010 test data for jobs that women may have been more likely to report job ends in SIPP-EHC relative to SIPP, but that is not supported in the second iteration of testing nor is it the case in program participation. Similarly, the 2010 test data for jobs showed that married individuals were less likely to report job endings conditional on duration, but this was also not true in the next field test or in program participation. Future analysis will include other demographic variables such as household size and age of respondent that may influence recall or respondent burden.

References

Brown, James and Light, Audrey. 1992. "Interpreting Panel Data on Job Tenure," *Journal of Labor Economics,* Vol. 10 (3): 219-257.

Hu, Luojia and Taber, Christopher. 2011. "Asymmetric Information and Heterogeneous Human Capital," *Journal of Labor Economics*, Vol. 29 (1): 113-152.

Kuhn, Peter and Sweetman, Arthur. 1999. "Vulnerable Seniors: Unions, Tenure, and Wages following Permanent Job Loss," *Journal of Labor Economics*, Vol 17 (4): 671-693.

Pearce, James. 1990. "Tenure, Unions, and the Relationship between Employer Size and Wages," *Journal of Labor Economics,*Vol 8 (2): 251-269.

Pierret, Charles. 2001. "Event History Data and Survey Recall," *The Journal of Human Resources,* Vol. 36 (3): 439-466.

Ruhm, Christopher. 1986. "Job Tenure and Cyclical Changes in the Labor Market," *The Review of Economics and Statistics*, Vol. 69 (2): 372-378.

SIPP-EHC Data Evaluation Workgroup, U.S. Census Bureau. 2011. "An Initial Evaluation of the 2010 Field Test of Re-Engineered SIPP." *SEHSD Working Papers.*

Vaughn, Denton and Scheuren, Fritz. "Longitudinal Attrition in the Survey of Income and Program Participation." Census Working Paper 242. http://www.census.gov/sipp/workpapr/wp242.pdf.

Zabel, Jeffrey. "An Analysis of Attrition in the Panel Study if Income Dynamics and the Survey of Income and Program Participation with an Application to a Model of Labor Market Behavior." *The Journal of Human Resources,* Vol. 33 (2): 479-506.