# Analytical Comparison of the SIPP and CPS-ASEC Key Longitudinal Estimates

Smanchai Sae-Ung, C. Dennis Sissel, Tracy L. Mattingly
U.S. Census Bureau
Contact Author Smanchai.Sae.Ung@census.gov

**Key Words:** Longitudinal, Design, Non-response, Attrition, Weighting, Simulation, Modeling, Estimates

**Abstract:** The demographic and socioeconomic longitudinal survey SIPP and cross-sectional survey CPS-ASEC have identical survey universes and sampling frames, and are similar in sample design, demographic and socioeconomic contents, questionnaire design profiles, and data collection and processing systems. In contrast, the SIPP collects monthly data but the CPS-ASEC collects annual data. The SIPP sample households are interviewed every four months over three to four years but the CPS-ASEC ones are interviewed once a year for two consecutive years. As a longitudinal survey, the SIPP follows its movers and identifies its survey universe leavers; but, as a cross-section survey, the CPS-ASEC does not. Inspired by similarity between these two surveys, and higher cross-sectional response rates, fresher sample, and establishment stature of the CPS-ASEC, we developed a methodology to construct a quasi-longitudinal dataset from two consecutive CPS-ASEC cross-sectional datasets based on a set of probabilistic models to simulate the survey universe leavers among the sample people in a mix of not-followed movers and universe leavers that could not be definitively identified or separated. We weighted the CPS-ASEC quasi-longitudinal file to reduce the nonresponse bias using the same SIPP longitudinal weighting procedure. As a result of the simulation of survey universe leavers and the longitudinal weighting discussed above, we could use the data of two consecutive CPS-ASEC samples to properly create longitudinal estimates for comparing with those of the SIPP. With addition of identical weighting process to other identical or similar systematic elements mentioned earlier, we could practically attribute the cause of the differences between the SIPP and CPS-ASEC key longitudinal estimates to just among the following nonsampling error sources: attrition/non-response, sample freshness (aging), recall period, four versus twelve month seam, collection of monthly versus yearly data, and detailed content and structure of the questionnaire. We applied this methodology to the 2002 and 2003 CPS-ASEC Supplements and the corresponding 2001 SIPP Panel. This methodology is also directly applicable to the reengineering of the SIPP.

## 1.0 Background and Objective[1]

The Survey of Income and Program Participation (SIPP) is a longitudinal survey about the demographic and socioeconomic well-being of the civilian non-institutional population of the United States (U.S.) and adults in the U.S. Armed Forces living in an off-or-on post household with at least one civilian adult member. The SIPP principally provides monthly information about the cash and non-cash income and program participation of individuals, families, and households. It also collects data on taxes, assets, liabilities, participation in government transfer programs, and health insurance coverage. Data from the SIPP allow the government to evaluate the effectiveness of federal, state, and local programs. The households in a SIPP sample (panel) are interviewed every four months over three to four years. In each interview, the SIPP collects monthly data of the last four months preceding the interview month. The monthly data can be aggregated into annual data.

The Current Population Survey (CPS) is a cross-sectional survey about the monthly labor force of the civilian non-institutional population of the United States. The CPS Annual Social and Economic Supplement (CPS-ASEC) is an annual cross-sectional survey about the social and economic well-being of the civilian non-institutional population of the United States and adults in the U.S. Armed Forces living in an off-or-on post household with at least one civilian adult member. The CPS-ASEC collects annual supplemental data on work experience, income, cash and non-cash benefits including those from public/government assistance programs, migration, and health insurance coverage. The CPS-ASEC is formerly known as the March Supplement or the Annual Demographic Supplement. The households in the CPS-ASEC sample are interviewed once a year (in February or March or April) for two consecutive years. In each interview, the CPS-ASEC collects annual data from the preceding year. In each CPS-ASEC Supplement sample, about half of the sample households were in sample the previous year. Consequently, two consecutive CPS-ASEC samples provide the two years of annual data for the sample

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

people residing in a sample household that was in sample and interviewed in both of the two consecutive samples under consideration. Because (a) some of the interviewed sample households in the first of the two consecutive samples become a non-interviewed sample household in the second of the two consecutive samples and (b) some of the interviewed sample households in the first of the two consecutive samples lose some of their original household members in the second of the two consecutive samples, only about one third (instead of about half) of the sample people in the first of two consecutive CPS-ASEC samples have two years of data that is necessary for making longitudinal estimates (for examples, gross changes and status transition between two consecutive years). However, as a cross-sectional survey, the CPS-ASEC Supplement neither follows the movers nor records the reason that a sample person leaves the survey universe in the second of two consecutive CPS-ASEC interviews. A CPS-ASEC (or SIPP) sample person becomes a survey universe leaver when he/she dies, moves to live in Armed Forces Barracks, emigrates to another country, or is institutionalized. The demographic and socioeconomic characteristics of the movers and the survey universe leavers are likely to be substantially different from those who continue to stay in the same households. Therefore, we could not just directly use the data and the final cross-sectional weights of the sample people interviewed for two consecutive years to produce any longitudinal estimates without concern that such difference may induce a significant level of bias in the longitudinal estimates. One approach to circumvent this concern is to develop a set of probabilistic models, based on the data of the movers and survey universe leavers collected in the first of two consecutive CPS-ASEC interviews, to simulate a mover status or a survey universe leaver status of each of the sample people who is a member of his/her household in the first of two consecutive CPS-ASEC interviews. Development of this set of models is one of the two main objectives of our research discussed in this paper. The other main objective is to reduce the longitudinal nonresponse bias using the SIPP's longitudinal weighting procedure to transfer the initial weights from the simulated movers to the sample people who are either interviewed in two consecutive CPS-ASEC interviews or simulated universe leavers in the second of two consecutive CPS-ASEC interviews, and to perform the post stratification (second stage) weight adjustment. The SIPP's longitudinal weighting procedure is chosen because (a) it is a well-established procedure, and (b) more importantly, it automatically avoids creating another systematic difference between the SIPP and CPS-ASEC Supplement (due to weighting process). As a result of the simulation of survey universe leavers and the longitudinal weighting discussed above, we could use the data of two consecutive CPS-ASEC samples to properly create longitudinal estimates for comparison with those of the SIPP.

Most of the essential systematic features of the SIPP and CPS-ASEC are either identical or similar. Namely, they have identical survey universes and sampling frames, similar sample designs (as discussed in section 2.0), data collection and processing systems, and survey subject contents. Their systematic contrasts are (a) the SIPP collects monthly data but the CPS-ASEC collects yearly data, (b) the data collection recall period is up to four months but up to sixteen months[2] for the CPS-ASEC, (c) The SIPP sample households are in sample for three to four years and interviewed every four months but the CPS-ASEC sample households are generally in sample for 16 months and interviewed monthly for the first four and last four months, and only two of the eight interviews collect the CPS-ASEC data in addition to the basic CPS data, i.e., the SIPP is deemed more burdensome to its respondents than the CPS-ASEC, and (d) the CPS-ASEC refreshes its sample every month by replacing the oldest of its eight rotations (eight self representative sub-samples) with a new rotation but the SIPP, as a longitudinal survey, does not, i.e., the CPS-ASEC sample is generally more up-to-date (fresher) than the SIPP sample.

On the basis of (a) the essential systematic identities and similarities between the two surveys and the addition of an identical longitudinal weight adjustment procedure discussed earlier, and (b) the following CPS-ASEC's attributes: lower attrition rate, fresher sample, less burdensomeness to its respondents, and, more importantly, its establishment stature; the longitudinal estimates produced from two consecutive CPS-ASEC samples could serve as a practical benchmark for gauging the quality of the SIPP data by analytically comparing the same longitudinal estimates produced by the two surveys. More importantly, based on the essential systematic identities and similarities between the two surveys and the identical longitudinal weighting process discussed earlier, we could practically attribute the cause of any differences between the SIPP and CPS-ASEC key longitudinal estimates to just among the following nonsampling error sources: attrition/non-response, sample freshness (aging), recall period, four versus twelve month seam, collection of monthly versus yearly data, and detailed content and structure of the questionnaire. This amounts to a substantial reduction in our effort to search for the cause of any differences between the SIPP and CPS-ASEC key longitudinal estimates.

None of the existing SIPP monthly cross-sectional weights and calendar year and panel longitudinal weights is identical or adequately comparable to the existing CPS-ASEC cross-sectional weights and the longitudinal weight produced from two consecutive CPS-ASEC samples for the following reasons.

---

[2] Except for the taxable income, the recall period is from one to three months for those who file their income tax returns on time.

- A SIPP monthly cross-sectional weight for March represents the survey universe as of March and the data pertaining to the same month (March) though a CPS-ASEC cross-sectional weight represents the survey universe as of March but the data pertaining to the entire previous year.
- A SIPP calendar year longitudinal weight represents the survey universe as of January and the data pertaining to the year starting with the same month (January); though a CPS-ASEC cross-sectional weight represents the survey universe as of March (just two months apart from January) but the data covering the entire *previous* year (a contrast of one year apart from the same year).
- A 1996 or later SIPP longitudinal panel weight represents the survey universe as of the first January (of the panel length) and the data covering the entire panel length (three or four years long). Based on earlier discussion, a longitudinal weight produced from two consecutive CPS-ASEC samples represents the survey universe in first March (based on the fielding time of the first of the two consecutive samples) but the data covering the year before and the same year as the first March. Suppose we choose two consecutive CPS-ASEC samples with the first March in the same year as the first January (of the entire panel length) of a SIPP panel. Since the SIPP panel data of a sample person cover none or only one to three months of the year before the first January, the corresponding longitudinal SIPP panel and CPS-ASEC weights are not comparable (even though their survey universes are only two months apart). Suppose we choose two consecutive CPS-ASEC samples with the first March coinciding with the second March of a SIPP panel so that the SIPP panel also has data in one year before and one year in the second March. However, the corresponding longitudinal SIPP panel and CPS-ASEC weights are not comparable because the panel survey universe for the SIPP is in the first January of the SIPP panel length but the longitudinal survey universe for these two consecutive CPS-ASEC samples is in the second March of the SIPP panel length, i.e., their survey universes are 14 months apart. Thus, even the SIPP longitudinal panel weight is not comparable with the longitudinal weights produced from two consecutive CPS-ASEC samples.

Therefore, to obtain any SIPP longitudinal estimates corresponding to the longitudinal estimates provided by two consecutive CPS-ASEC samples, we also need to produce a SIPP longitudinal weight for the survey universe coinciding with the first March of two consecutive CPS-ASEC samples and for the data in one year before and the same year as the first March (of two consecutive CPS-ASEC samples). Thus, we need to include a development of a procedure for producing such SIPP longitudinal weights in this study.

Based on the above discussion, we developed a methodology to produce the CPS-ASEC and SIPP longitudinal weights for the analytical comparison between the SIPP and CPS-ASEC longitudinal estimates with full details described in Sae-Ung and Dennis (2006 and 2007). In brief, our methodology consists of the following set of procedures:

- Procedure for creating a quasi-longitudinal dataset (micro) from two consecutive CPS-ASEC cross-sectional datasets (micro).
- Procedure for implementing the probabilistic models for the deceased, barracked, emigrated, and institutionalized survey universe leavers constructed based on the universe leaver rates (probabilities) estimated from the universe leaver data from the SIPP Panel corresponding to two consecutive CPS-ASEC samples under consideration.
- Procedure for simulating the deceased, barracked, emigrated, and institutionalized sample people among the CPS-ASEC sample people in the first of two CPS-ASEC consecutive based on the probabilistic models for the deceased, barracked, emigrated, and institutionalized survey universe leavers.
- Procedure for longitudinal weighting of the CPS-ASEC quasi-longitudinal dataset (made valid with the simulated survey universe leavers) and its corresponding SIPP Panel longitudinal dataset, based on the SIPP longitudinal weight adjustment procedure specified in SIPP Branch (2002).

A full description of all essential features in this methodology is provided in sections 3.0 and 4.0. For simplicity and clarity but without loss of generality, we developed our methodology based directly on the 2002 and 2003 CPS-ASEC samples and the corresponding 2001 SIPP Panel which covered 33 months of full sample monthly data from January 2001 to September 2003. A full description of all essential features in our methodology is provided in sections 3.0 and 4.0. We also developed a set of SAS computer programs (software) to implement our methodology. To apply our methodology to other pairs of two consecutive CPS-ASEC samples, all we have to do is replace (or recode) input variables from the 2002 and 2003 CPS-ASEC micro dataset and the 2001 SIPP Panel longitudinal dataset by the corresponding variables from other pairs of two consecutive CPS-ASEC micro dataset and their corresponding SIPP Panel longitudinal micro dataset.

Using the longitudinal weights produced for the 2002 and 2003 CPS-ASEC and its corresponding 2001 SIPP Panel by our software (described above), we could properly carry out comparison for the SIPP and CPS-ASEC keys longitudinal estimates in the period of 2001 and 2002. Since the 2001 SIPP Panel sample was only nationally representative by sample design, our analytical comparison between the SIPP and CPS-ASEC longitudinal estimates in section 5.0 is limited to national estimates. In this report, we provided some analytical comparison of the SIPP and CPS-ASEC key longitudinal estimates in section 5.0.

## 2.0 Similarity and Difference in CPS-ASEC and SIPP Sample Design

The sample design of the CPS-ASEC and SIPP are based on two-stage probability sample design methodology. In brief, each (of the two surveys) forms its primary sample units (PSUs) of housing units based principally on contiguous and practical sized geography and demographic heterogeneity. The size of a PSU is measured as the number of housing units that it contains. In the 1990 Sample Redesign, the two surveys formed they PSUs somewhat differently. Namely, each CPS-ASEC PSU was not allowed to cross any state boundaries because each CPS-ASEC sample was designed to be state representative. However, the some SIPP PSUs were allowed to cross state boundaries (for better overall design) because each SIPP sample was designed to be only nationally representative. However, for 2000 Sample Redesign, the SIPP sample was designed to be state representative like the CPS-ASEC one. Thus, the PSUs for the CPS-ASEC and SIPP are exactly the same (and no PSUs were allowed to cross any state boundaries). [Note that the samples for the 2002 and 2003 CPS-ASEC Supplements directly under consideration in this study are from the 1990 Sample Redesign. The sample for the 2004 CPS-ASEC Supplement is also from the 1990 sample redesign. The samples for the 2005 CPS-ASEC Supplement are from both the 1990 and 2000 Sample Redesigns. The samples for the 2006 CPS-ASEC Supplement and beyond (until the 2010 Sample Redesign) are from the 2000 Sample Redesign.]

Each of the two surveys classifies its PSUs as either a self-representative PSU (SRPSU) or a non-self representative PSU (NSRPSU). An SRPSU is selected with certainty (probability of selection = 1). Each of the two surveys groups its NSRPSUs into many strata based on their demographic similarity and its design requirements or constraints (such as national or state representative design, its most important survey characteristics, field workload constraint, etc.). For a state representative design, NSRPSUs are stratified within each state. For national representative design, NSR PSUs are stratified within each Census Region. The CPS-ASEC sample selects one NSRPSU per stratum with the probability of selection proportional to size (PPS) but the SIPP sample selects two PSUs per stratum with PPS. The size of an SRPSU is generally larger that those of NSRPSUs. The two surveys classify their SRPSU/NSRPSU independently to satisfy their own sample design constraints (for both the 1990 and 2000 Sample Redesigns).

From each sample PSU, the CPS-ASEC makes a systematic selection of sample housing units by sorting the housing units based on the Census demographic data and geography so as to minimize the within-PSU variances of its estimates. For the SIPP, each sample PSU is further stratified in two strata, i.e., a poverty stratum and a non-poverty stratum classified using the Census data. Within each sample PSU, its poverty stratum has significantly higher estimate of low-income households than that of its non-poverty stratum one. The housing units in the poverty strata are selected with a higher rate (selection probability) than those in the non-poverty strata (over sampling from the poverty strata) so as to achieve a higher reliability for poverty rate estimate for a given budgeted sample size. From each of the two strata within each sample PSU, the SIPP makes a systematic selection of sample housing units by sorting the housing units based on the Census demographic data and geography so as to minimize the within-PSU standard error estimates.

Based on the above discussed, the CPS-ASEC and SIPP sample design have both common and different essential features. However, for the analytical comparison between their key longitudinal estimates, the differences between their essential sample design features are adequately accounted for through the standard error of their longitudinal estimates. Namely, a valid comparison between two estimates needs their standard errors (not just the estimates per se) to statistically determine the significance of their difference. Therefore, the essential difference in the CPS-ASEC and SIPP sample design does not invalidate the analytical comparison between the CPS-ASEC and SIPP longitudinal estimates by virtue of including their standard error estimates to account for their essential sample design difference.

## 3.0 Survey Universe Leaver Probabilistic Models

Based on definition of the CPS-ASEC/SIPP survey universe described in section 1.0, there are four types of survey universe leavers in these two surveys as classified below.

- An individual leaves the survey universe because of his/her death; hereinafter, he/she will be simply referred to as *a deceased universe leaver*.
- An individual leaves the survey universe because he/she moves to live in U.S. Armed Forces barracked; hereinafter, he/she will be simply referred to as *a barracked universe leaver*.
- An individual leaves the survey universe because he/she moves to live in other country; hereinafter, he/she will be simply referred to as *an emigrated universe leaver*. Note that term *an emigrated universe leaver* is the same as the term *an expatriated universe leaver* used in Sae-Ung and Sissel (2006 and 2007).
- An individual leaves the survey universe because he/she is institutionalized; hereinafter, he/she will be simply referred to as *an institutionalized universe leaver*.

In this section, we describe the essence of our deceased, barracked, emigrated, and institutionalized universe leaver probabilistic models in sections 3.1 to 3.4 respectively. The fully detailed specifications of these four models are provided in Sae-Ung and Sissel (2007).

### 3.1 Deceased Universe Leaver Probabilistic Model

We model the probability of death (becoming a deceased universe leaver) among the people in the CPS-ASEC/SIPP universe in a given time period (e.g., 12 months) based on the assumption that their probability (rate) of death (POD) can be expressed as a function of a set of demographic and/or socioeconomic variables/characteristics (covariates) such that (a) the realizations of these covariates and their response (POD function) can be directly obtained or derived from the CPS-ASEC's and SIPP's data, and (b) the covariates of their POD function can be partitioned into a finite set of mutually exclusive rectangular domains (cells) such that the POD function within each of these cells is practically constant. We determine these covariates principally based on the report prepared by Hoyert, Heron, et al (2006) for the National Center for Health Statistics (NCHS). Namely, this NCHS report estimates the 2003 annual deceased rate (probability) among the people in the U.S. using administrative records, and classifies the people in the U.S. into a set of cells by sex, race and age such that the annual deceased rate among the people within an individual cell is about the same (approximately constant). Namely, the NCHS deceased rate cell classification by sex, race and age is as shown in tables 1A to 1D. Thus, our POD model is constructed using the same cells classified by the NCHS deceased rate cell classification as shown in column 2 of tables 1A to 1D with the following additional enhancements. We judiciously add health insurance coverage status and family income level variables in the classification of our POD model cells, as shown in columns 3 and 4 of tables 1A to 1D, to further increase the degree of similarity of the deceased rate among the people within the same cell. The addition of health insurance coverage status and family income level variables results in a total of 114 cells for each of these four tables (tables 1A to 1D). We arrange the 114 cells in each of these four tables based on the degree of similarity of their cells' 2003 deceased rates published in this NCHS report by Hoyert, Heron, et al (2006), namely, the closer the cell numbers the more similar the deceased rates of the cells. We also assign scale values for the cells (in each of these four tables) to reflect the relative degrees of similarity in their deceased rates (as shown in column 5 of tables 1A to 1D). In section 4.0, the scale values in each of these four tables are used for collapsing (merging) adjacent cells *if needed*. This will lead to achieving reliable estimates of deceased rates based on the SIPP data because the SIPP sample people available for an individual cell may be too small to provide a reliable estimate of its deceased rate on its own.

### 3.2 Barracked Universe Leaver Probabilistic Model

Similar to the deceased universe leaver model, we model the probability of becoming a barracked universe leaver among the people in the CPS-ASEC/SIPP universe in a given time period (e.g., 12 months) based on the assumption that their probability (rate) of becoming a barracked universe leaver (POB) can be expressed as a function of a set of demographic and/or socioeconomic covariates such that (a) the realizations of these covariates and their response (POB function) can be directly obtained or derived from the CPS-ASEC's and SIPP's data, and (b) the covariates of their POB function can be partitioned into a finite set of mutually exclusive rectangular domains (cells) so the POB function within each of these cells is practically constant. Due to limited available information, we judiciously construct the cells for the model based on (a) the Department of Defense's manpower classified by races in the Statistical Abstract of the Unites States complied by U.S. Census Bureau (2002) and (b) the age requirement for enlistment into the U.S. Armed Forces provided by Usmilitary.about.com (2006). Namely, in the same manner as our POD model specified in section 3.1, our POB model is represented by a set of cells classified by sex, race, and age, and their cell scale values as shown in tables 2A and 2B under the assumptions that (a) the rate of becoming a barracked universe leaver among the people in the same cell is practically constant, and (b) the scale values of the cells (in each of these two tables) reflect the degree of similarity in their rates of becoming a barracked universe leaver. In section 4.0, the scale values in each of these two tables are used (in section 4.0) for collapsing (merging) adjacent cells *if needed*. This will lead to achieving reliable estimates of rates (of becoming a barracked universe leaver) based on the SIPP data because the SIPP sample people available for an individual cell may be too small to provide a reliable estimate of its rate (of becoming a barracked universe leaver) on its own.

### 3.3 Emigrated Universe Leaver Probabilistic Model

Similar to the deceased universe leaver model, we model the probability of becoming an emigrated universe leaver among the people in the CPS-ASEC/SIPP universe in a given time period (e.g., 12 months) in based on the assumption that their probability (rate) of becoming an emigrated universe leaver (POE) can be expressed a function of a set of demographic and/or socioeconomic covariates such that (a) the realizations of these covariates and their response (POE function) can be directly obtained or derived from the CPS-ASEC's and SIPP's data, and (b) the covariates of their POE function can be partitioned into a finite set of mutually exclusive rectangular domains (cells) such the POE function within each of these cells is practically constant. Based on the information about emigration rates among the U.S. born and foreign born people available in Ahmed and Robinson (1994) and Fernandez (1995), we judiciously construct the cells for the model based on the nativity statuses (U.S. born or foreign born) and ages of the people. Namely, in the same manner as our POD model specified in section 3.1, our POE model is represented by a set of cells classified by nativity and age, and their cell scale values as shown in tables 3A and 3B under the assumptions that (a) the rate of becoming an expatriated universe leaver among the people in the same cell is practically constant, and (b) the scale values of the cells (in each of these two tables) reflect the degree of similarity in their rates of becoming an expatriated universe leaver. As mentioned in section 3.1, the scale values in each of these two tables are used (in section 4.0) for collapsing (merging) adjacent cells *if needed*. This will lead to achieving reliable estimates of rates (of becoming an expatriated universe leaver) based on the SIPP data because the SIPP sample people available for an individual cell may be too small to provide a reliable estimate of its rate (of becoming an expatriated universe leaver) on its own.

### 3.4 Institutionalized Universe Leaver Probabilistic Model

Similar to the deceased universe leaver model, we model the probability of becoming an institutionalized universe leaver among the people in the CPS-ASEC/SIPP universe in a given time period (e.g., 12 months) in based on the assumption that their probability (rate) of becoming an institutionalized universe leaver (POI) can be expressed as a function of a set of demographic and/or socioeconomic covariates such that (a) the realizations of these covariates and their response (POI function) can be directly obtained or derived from the CPS-ASEC's and SIPP's data, and (b) the covariates of their POI function can be partitioned into a finite set of mutually exclusive rectangular domains (cells) so the POI function within each of these cells is practically constant. Based on the information about imprisonment rates reported by Bonczar (2003) for the U.S. Bureau of Justice Statistics, and our own judgment (due to lacking information about other forms of institutionalization), we judiciously construct the cells for the model based on sexes, races, and ages of the people. Namely, in the same manner as our POD model specified in section 3.1, our POI model is represented by a set of cells classified by sex, race, and age, and their cell scale values as shown in tables 4A to 4D. The cells in table 4A to 4D are created under the assumptions that (a) the rate of becoming an institutionalized universe leaver among the people in the same cell is practically constant, and (b) the scale values of the cells (in each of these four tables) reflect the degree of similarity in their rates of becoming institutionalized universe leavers. As mentioned in section 3.1, the scale values in each of these four tables are used (in section 4.0) for collapsing (merging) adjacent cells *if needed*. This will lead to achieving reliable estimates of rates (of becoming an institutionalized universe leaver) based on the SIPP data because the SIPP sample people available for an individual cell may be too small to provide a reliable estimate of its rate (of becoming an institutionalized universe leaver) on its own.

## 4.0 Survey Universe Leaver Simulation and Longitudinal Weighting Methodology

As discussed in section 1.0, for simplicity and clarity yet without loss of generality, we developed our methodology for producing the CPS-ASEC and SIPP longitudinal weights for analytical comparison between the SIPP and CPS-ASEC longitudinal estimates based directly on (a) the 2002 and 2003 CPS-ASEC samples and the corresponding 2001 SIPP panel, and (b) the four survey universe leaver probabilistic models described in section 3.0 (for simulating the deceased, barracked, emigrated, and institutionalized universe leavers among the March 2002 CPS-ASEC sample people using the corresponding survey universe leaver data from the SIPP panel 2001). The full detailed description and procedural specifications of our longitudinal weighting methodology are provided in Sae-Ung and Dennis (2006 and 2007). In this section, we describe only the essential features of the procedures that sequentially make up our survey universe leaver simulation and longitudinal weighting methodology in steps 1 to 15 below.

**Step 1 – Creating of the 2002 and 2003 CPS-ASEC Quasi-longitudinal Dataset**

In this step, we create a 2002 and 2003 CPS-ASEC quasi-longitudinal dataset (micro-data file) that (a) is a SAS dataset, (b) combine all household, family, and person level cross-sectional micro data from the 2002 and 2003 CPS-ASECs into a rectangular dataset for easy utilization. Namely, the 2002 and 2003 CPS-ASEC quasi-longitudinal dataset includes (a) all non-interviewed sample households for attrition/nonresponse analysis and weighting, and (b) all interviewed sample households and their household members (sample people) and family units (primary family and sub-families). As a cross-sectional survey, the CPS-ASEC Supplement does not record the reason that a sample person, belonging to sample household interviewed for two consecutive CPS-ASEC Supplements, was interviewed in the first but not the second of the two consecutive CPS-ASECs. All we know is that it could be of one of the followings: because of his/her death or being barracked or emigration or being institutionalized or simply moving to live elsewhere within the country (becoming a mover). As a cross-sectional survey, the CPS-ASEC does not follow sample people who move. As a cross-sectional survey, the CPS-ASEC does not provide a means to uniquely identify a sample person who was interviewed for two consecutive CPS-ASECs even though it does provide a variable to uniquely identify a sample household across time (e.g., two consecutive CPS-ASEC Supplements). Consequently, we could not directly construct a *complete* longitudinal data from two consecutive CPS-ASEC samples, namely, we could only directly construct a *quasi* longitudinal data set like the 2002 and 2003 CPS-ASEC quasi-longitudinal dataset, because of the following two missing elements: element A - the reason a sample person is not interviewed in the second of the two consecutive CPS-ASEC Supplements, and element B - means to uniquely identify a sample person who was interviewed in both of two consecutive CPS-ASEC Supplements. Therefore, we develop the deceased, barracked, emigrated, and institutionalized universe leaver probabilistic models (described in section 3.0) to practically supply element A's information by simulation. Based on the practical guide provided in U.S. Bureau of Labor Statistics and U.S. Census Bureau (2002), one can identify, with high reliability of correctness, a sample person who was interviewed in two consecutive CPS-ASEC Supplements by matching the following information in the two consecutive CPS-ASEC Supplements: sample household unique identification, sample household member line number, sample person sex, sample person race, and sample person age. This is the approach we take to practically supply element B's information. [We demonstrate later in section 5.0 that the overall error or imperfection of this matching approach for the 2002 and 2003 CPS-ASEC Supplements substantially lower than 1.67%.]

**Step 2 – Defining the CPS-ASEC/SIPP Longitudinal Universe**

Based on the data collection months (February, March, April) and the data years of the 2002 and 2003 CPS-ASEC Supplements, we define the CPS-ASEC/SIPP longitudinal universe for the 2002 and 2003 CPS-ASEC Supplements and its corresponding 2001 SIPP Panel as *the civilian non-institutional population of the U.S. including adults (aged 17 and over)*[3] *in the U.S. Armed Forces but living in an off-or-on post household with at least one civilian adult (aged 15 and over) in <u>March 2002</u>*. The data time period (i.e., data years or data months) under consideration for this longitudinal universe is from 2001 through 2002 (January 2001, …, December 2001, January 2002, …, December 2002). The data time period is used for determining the longitudinal interview status of each sample person in the longitudinal weight adjustment process (for both the 2002 and 2003 CPS-ASEC quasi-longitudinal datasets and the 2001 SIPP Panel longitudinal dataset). Hereinafter, this longitudinal universe is referred to as *the March 2002 CPS-ASEC/SIPP longitudinal universe*.

**Step 3 – Defining the Longitudinal Interview Status of the CPS-ASEC Sample People**

Among all sample persons in the 2002 and 2003 CPS-ASEC quasi-longitudinal dataset, we define their longitudinal interview statuses in the following manner:

- A sample person who was interviewed in the 2002 CPS-ASEC Supplement and <u>either</u> was also interviewed in <u>or</u> was classified as a simulated deceased or barracked or emigrated or institutionalized universe leaver in the 2003 CPS-ASEC Supplement in step 11 is a longitudinal interviewed person for the March 2002 CPS-ASEC/SIPP longitudinal universe. Hereinafter, he/she is referred to as a *March 2002 CPS-ASEC longitudinal interviewed sample person*.
- A sample person who was interviewed in the 2002 CPS-ASEC Supplement but was not interviewed in the 2003 CPS-ASEC Supplement for any reasons other than being classified as a simulated deceased or barracked or

[3] By Armed Forces personnel requirement, an individual must be at least 17 years old to enter the Armed Forces Services.

emigrated or institutionalized universe leaver in the 2003 CPS-ASEC Supplement in step 11 is a longitudinal non-interviewed person for the March 2002 CPS-ASEC longitudinal universe. Hereinafter, he/she is referred to as a *March 2002 CPS-ASEC longitudinal non-interviewed sample person*.

- A sample person who is not a March 2002 CPS-ASEC longitudinal interviewed sample person or a March 2002 CPS-ASEC longitudinal non-interviewed sample person is irrelevant with the longitudinal weight process (for the 2002 and 2003 CPS-ASEC quasi-longitudinal dataset), thus he/she is excluded from any further consideration (in step 4 and beyond). Hereinafter, he/she is referred to as a *March 2002 CPS-ASEC longitudinal excluded sample person*.

**Step 4 – Defining the Longitudinal Interview Status of the SIPP Sample People**

Among all sample persons in the 2001 SIPP Panel longitudinal dataset, we define their longitudinal interview statuses consistently with the CPS-ASEC sample people's longitudinal interview statuses (defined in step 3) in the following manner:

- A sample person who was interviewed for March 2002 and <u>either</u> was also interviewed for all other 11 months in 2001 and all 12 months in 2002 <u>or</u> became a deceased or barracked or emigrated or institutionalized universe leaver in any month between April 2002 and March 2003. Hereinafter, he/she is referred to as a *March 2002 SIPP longitudinal interviewed sample person*.
- A sample person who was interviewed for March 2002 but was not interviewed in all other 11 months in 2001 and all 12 months in 2002 for any reasons other than becoming a deceased or barracked or emigrated or institutionalized universe leaver in any month between April 2002 and March 2003. Hereinafter, he/she is referred to as a *March 2002 SIPP longitudinal non-interviewed sample person*.
- A sample person who is not a March 2002 SIPP longitudinal interviewed sample person or a March 2002 SIPP longitudinal non-interviewed sample person is irrelevant with the longitudinal weight process (for the 2001 SIPP Panel longitudinal dataset), thus he/she is excluded from any further consideration (in step 5 and beyond). Hereinafter, he/she is referred to as a *March 2002 SIPP longitudinal excluded sample person*.

We remark that, by definition, only an original sample person (wave 1 interviewed person) has a chance to be qualified as a March 2002 SIPP longitudinal interviewed sample person. This is because only the original sample people were interviewed for the month of January 2001. All non-original sample persons were not interviewed for January 2001, i.e., they certainly had a least one missing interview month between January 2001 through December 2002; thus they could not be qualified as a March 2002 SIPP longitudinal interviewed sample person.

**Step 5 – Determining the initial weights of the SIPP Sample People**

Based on the longitudinal weighting procedure specified in SIPP Branch (2002), the initial weight for each of the March 2002 SIPP longitudinal interviewed and non-interviewed sample people is his/her weight obtained right before the second stage weight adjustment (post-stratification weight adjustment). Thus, by definition, his/her initial weight is his/her cross-sectional non-interviewed adjusted weight for the month of March 2002. Hereinafter, his/her initial weight is referred to as *his/her March 2002 SIPP longitudinal initial weight*.

**Step 6 – Determining the initial weights of the CPS-ASEC Sample People**

As discussed in section 1.0, we also use the SIPP longitudinal weighting procedure specified in SIPP Branch (2002) for the longitudinal weight adjustment for the 2002 and 2003 CPS-ASEC quasi-longitudinal dataset. Based on the SIPP longitudinal weighting procedure, the initial weight for each of the March 2002 CPS-ASEC longitudinal interviewed and non-interviewed sample people is his/her weight obtained right before the second stage weight adjustment (post-stratification weight adjustment). Thus, by definition, his/her initial weight is his/her March 2002 CHIP weight as indicated in the weighting process diagram in attachment A of HSS Branch (2005). Hereinafter, his/her initial weight is referred to as *his/her March 2002 CPS-ASEC longitudinal initial weight*.

**Step 7 – Calculating the March 2002 SIPP Longitudinal Non-Interview Adjusted and Final Longitudinal Weights**

Having determined the March 2002 longitudinal non-interview statuses and initial longitudinal weights of the SIPP sample people, we perform the longitudinal weighting using the SIPP longitudinal weighting procedure specified in SIPP Branch (2002) to produce the longitudinal non-interview adjusted weight and the final longitudinal weight. In brief, the longitudinal non-interview adjusted weights are obtained by transferring the initial longitudinal weights from the longitudinal non-interviewed sample people to the longitudinal interviewed sample people. The weight transferring is based on their similarity in their survey design, demographic, and socioeconomic variables (characteristics) highly related to more important key characteristics of the SIPP (i.e., income, poverty, and program participation) as classified by the longitudinal non-interview weight adjustment cells established for the SIPP in SIPP Branch (2002). The final longitudinal weight is obtained from a post-stratification (second stage) weight adjustment procedure that adjusts the longitudinal non-interview adjusted weights to match a set of demographic controls (benchmark estimates) derived from the basic CPS and the cross-section CPS-ASEC data for March 2002. As indicated in SIPP Branch (2002), the demographic characteristics used to form the controls include race, ethnicity, age, and family type (e.g., female householder with no spouse present but with own children age ≤ 18) to highly reflect more important key characteristics of the SIPP (as well as the CPS-ASEC). As a result of the longitudinal weighting, only the longitudinally interviewed sample people have a positive weight, and all other sample people have a zero weight. This implies that the March 2002 SIPP longitudinal interviewed sample people together with their final longitudinal weights provide a full representation of the March 2002 CPS-ASEC/SIPP longitudinal universe.

Hereinafter, the longitudinal non-interviewed adjusted weight produced in this step is referred to as *the March 2002 SIPP longitudinal non-interview adjusted weight*, and the post-stratification/second-stage longitudinally adjusted weight is referred to as *the March 2002 SIPP final longitudinal weight*. The March 2002 SIPP longitudinally non-interview adjusted weight is used to determine the benchmark rates/probabilities of the deceased, barracked, emigrated, and institutionalized universe leaver for the March 2002 CPS-ASEC/SIPP longitudinal universe as described in step 8. These benchmark rates are used (in step 11) to simulate the universe leavers among the 2002 CPS-ASEC sample people using the universe leaver probabilistic models defined in step 3.

**Step 8 – Determining the Benchmark Rates for Universe Leavers Based on SIPP Data**

To be able to use the four universe leaver probabilistic models (defined in step 3) to simulate the deceased, barracked, emigrated, and institutionalized universe leavers among the 2002 CPS-ASEC sample people, we need to have estimates of the deceased, barracked, emigrated, and institutionalized universe leaver rates (probabilities) for each cell or similar cell groups in tables 1A to 1D, 2A and 2B, 3A and 3B, and 4A to 4D, respectively. To meet this need, we use the data from the March 2002 SIPP longitudinal non-interview adjusted weight (produced in step 7) for the 2001 SIPP Panel to make these estimates in the manner as briefly described below.

Due to limited sample size, some of the cells in tables 1A to 1D (deceased), 2A and 2B (barracked), 3A and 3B (emigrated), and 4A to 4D (institutionalized), have no or only a few deceased or barracked or emigrated or institutionalized universe leaver sample people in them. Consequently, we could not make a reasonable estimate of the universe leaver rate for each of these cells individually. Therefore, we need to collapse (merge) each of these cells with other similar cells based on the closeness of their scale values into a cell group with sufficient universe leavers to render a reasonable universe leaver rate, and yet not to excessively collapse too many cells together so that the degrees of similarity among the collapsed cells are not compromised. Based on our study, we judiciously establish the following criterion for cell collapsing: For each type of universe leaver, if an individual cell of its universe leaver probabilistic model (e.g., tables 1A to 1D if deceased universe leavers is under consideration) has less than 15 sample people being a universe leaver of the type under consideration, then it must be successively collapsed with other similar cell or cells based on the closeness of their scale values until the collapsed cell group have 15 or more sample people being a universe leaver of the type under consideration for the first time. Based on this criterion, we developed a set of cells/collapsed-cell-groups for each of tables 1A to 1D, 2A and 2B, 3A and 3B, 4A to 4D, and calculated the estimates of their corresponding universe leaver rates using the March 2002 SIPP longitudinal non-interview adjusted weights (to be used for simulating each of the four types of universe leavers among the 2002 CPS-ASEC Supplement sample people in step 11). Based on the discussion in section 3.0, we could assume that, for a given type of a universe leaver, the probability of each individual belonging to a given cell/collapsed-cell-group is practically constant (identical). Thus, our estimate of the universe leaver rate in a given cell/collapsed-cell-group was taken as the ratio of {the sum of the March 2002 SIPP longitudinal non-interview adjusted weights of its universe leaver sample people} to {the sum of the March 2002 longitudinal non-interview adjusted weights of all of its the sample people}.

Hereinafter, a set of cells/collapsed-cell-groups and their universe leaver rate estimates developed for each of tables 1A to 1D, 2A and 2B, 3A and 3B, 4A to 4D are referred to in the following manner: the set of table per se is referred to as *a benchmark table*, each of the cell/collapsed-cell-groups in a universe leaver rate benchmark table is referred to as *a universe leaver benchmark cell/collapsed-cell-group*, and the universe leaver rate estimate of each benchmark cell/collapsed-cell-group is referred to as *a universe leaver benchmark rate*. For example, in these terminologies, for the deceased universe leaver probabilistic model for white or Asian or Pacific Islander males, we have a deceased universe leaver benchmark table consisting of a set of deceased universe leaver benchmark cells/collapsed-cell-groups (derived from table 1A) with their deceased universe leaver rates, for the white or Asian or Pacific Islander males.

Analytically, the universe leaver benchmark tables, their cells/collapsed-cell-group, and the universe leaver rate estimates of their cells/collapsed-cell-groups derived for tables 1A to 1D, 2A and 2B, 3A and 3B, 4A to 4D collectively constitute a set of complete specifications for these four universe leaver probabilistic models ready to be used for simulating each of the four type of universe leavers among the 2002 CPS-ASEC Supplement sample people (in step 11). Namely, for example, a deceased universe leaver benchmark table and its deceased universe leaver benchmark cells/collapse-cell-groups with their deceased universe leaver rate benchmark estimates derived from table 1A (in this step) are the complete model specifications for the deceased universe leaver probabilistic model for white or Asian or Pacific Islander males ready to be used for simulating the deceased universe leavers among the 2002 CPS-ASEC Supplement sample males whose races are white or Asian or Pacific Islander.

**Step 9 – Creating Pre Universe-leaver-simulation Longitudinal Interview Status of the CPS-ASEC Sample People**

Consider each of the 2002 CPS-ASEC Supplement sample people whose sample household was also interviewed but he/she was not interviewed for the 2003 CPS-ASEC Supplement because he/she no longer was a member of the sample household by then (and as a cross-sectional survey by design, the CPS-ASEC supplement did not record why he/she was no longer a member of the household). By definition of a March 2002 CPS-ASEC longitudinal interview status provided in step 3, his/her March 2002 CPS-ASEC longitudinal non-interview status is indeterminate, i.e., he/she can be <u>either</u> a March 2002 CPS-ASEC longitudinal interviewed sample person (if he/she were a deceased or barracked or emigrated or institutionalized universe leaver) <u>or</u> a March 2002 CPS-ASEC longitudinal non-interviewed sample person (if he/she were a mover). As discussed in section 1.0, in this study, we determine his/her universe leaver status by simulation in step 11 using the complete specifications for four universe leaver probabilistic models obtained in step 8. As one of the elements needed in the process for simulating his/her universe leaver status, we need to create a variable (denoted by LIS_02A03A) to represent the pre universe-leaver-simulation longitudinal interview status for all the sample people in the 2002 and 2003 CPS-ASEC quasi-longitudinal dataset in the following manner. We assign

- LIS_02A03A = YS for each sample person interviewed in both the 2002 and 2003 CPS-ASEC Supplements to imply that he/she is definitively a March 2002 CPS-ASEC longitudinal interviewed sample person.
- LIS_02A03A = YN for each sample person belonging to a sample household that was interviewed in both the 2002 and 2003 CPS-ASEC Supplements but he/she was interviewed only in the 2002 CPS-ASEC Supplement because he/she no longer was a member of the sample household by the time that the 2003 CPS-ASEC Supplement was in the field. Namely, we assign his/her LIS_02A03A to imply that his/her March 2002 CPS-ASEC longitudinal interview status is indeterminate for the reason described earlier in this step. Hereinafter, we refer to he/she as *a March 2002 CPS-ASEC indeterminate longitudinal-interview-status sample person*.
- LIS_02A03A = NO for each sample person interviewed in the 2002 CPS-ASEC Supplement but not interviewed in the 2003 CPS-ASEC Supplement for any reasons other than being classified as a sample person with LIS_02A03A = YN. Namely, we assign his/her LIS_02A03A = NO to imply that he/she is a March 2002 CPS-ASEC longitudinally non-interviewed sample person.
- LIS_02A03A = EX for each sample person who is not a March 2002 CPS-ASEC longitudinal interviewed sample person or a March 2002 CPS-ASEC longitudinal non-interviewed sample person or a March 2002 CPS-ASEC indeterminate longitudinal-interview-status sample person. Namely, we assign his/her LIS_02A03A = EX to imply that he/she is a March 2002 CPS-ASEC longitudinal excluded sample person (as defined in step 3).

The variable LIS_02A03A created as described above is added to the 2002 and 2003 CPS-ASEC quasi-longitudinal dataset.

**Step 10 – Calculating the Pre-universe-leaver-simulation March 2002 CPS-ASEC Longitudinal Non-interview Adjusted Weight**

As a starting March 2002 CPS-ASEC longitudinal non-interview adjusted weight needed in the process for simulating the universe leaver status of each of the March 2002 CPS-ASEC indeterminate longitudinal-interview-status sample people (LIS_02A03A = YS as defined in step 9), we derive this weight by treating each of the March 2002 CPS-ASEC indeterminate longitudinal-interview-status sample people as though he/she were a March 2002 CPS-ASEC longitudinal interviewed sample person. Hereinafter, *we refer to this weight as the pre-universe-leaver-simulation March 2002 CPS-*

*ASEC longitudinal non-interviewed adjusted weight*. For simple reference, we denote the pre-universe-leaver-simulation March 2002 CPS-ASEC longitudinal non-interview adjusted weight by PRE_LNIAW_02A03A.

In the same manner as step 7, we use the SIPP longitudinal non-interview weight adjustment procedure specified in SIPP Branch (2002) to produce the PRE_LNIAW_02A03A. In brief, the PRE_LNIAW_02A03A weights are obtained by transferring the March 2002 CPS-ASEC longitudinal initial weights (obtained in step 6) from the March 2002 CPS-ASEC longitudinally non-interviewed sample people LIS_02A03A = NO) to the March 2002 CPS-ASEC longitudinal interviewed sample people (LIS_02A03A = YS) and the March 2002 CPS-ASEC indeterminate longitudinal-interview-status sample people (LIS_02A03A = YN) based on their similarity in their survey design, demographic, and socioeconomic variables (characteristics) highly related to income, poverty, and program/public-assistance participation (which are more important key characteristics of the SIPP) as classified by the longitudinal non-interview weight adjustment cells (established for the SIPP) in SIPP Branch (2002). As a result of this longitudinal non-interview weight adjustment, only the March 2002 CPS-ASEC longitudinal interviewed sample people (LIS_02A03A = YS) and the March 2002 CPS-ASEC indeterminate longitudinal-interview-status sample people (LIS_02A03A = YN) have a positive PRE_LNIAW_02A03A, and all other sample people (LIS_02A03A = NO or EX) have zero PRE_LNIAW_02A03A (PRE_LNIAW_02A03A = 0). Namely, at this point, we let the sample people with positive PRE_LINAW_02A03A temporarily provide a full representation of the March 2002 CPS-ASEC/SIPP longitudinal universe.

**Step 11 – Simulating the Universe Leaver Status among the 2002 CPS-ASEC Supplement Sample People**

Based on (a) the complete specifications for the deceased, barracked, emigrated, and institutionalized universe leaver probabilistic models obtained in step 8, (b) the pre universe-leaver-simulation longitudinal interview status (represented by the variable LIS_02A03A) obtained in step 9 and (c) the pre-simulation March 2002 CPS-ASEC longitudinal non-interviewed adjusted weight PRE_LNIAW_02A03A obtained in step 10, we simulate the universe leaver among the sample people with positive PRE_LNIAW_02A03 (temporarily providing a full representation of the March 2002 CPS-ASEC/SIPP longitudinal universe) using the procedure briefly described below.

- For each of the benchmark cells/collapsed-cell-groups derived for each of tables 1A to 1D (for deceased universe leavers), 2A and 2B (for barracked universe leavers), 3A and 3B (for emigrated universe leavers), and 4A to 4D (for institutionalized universe leavers) derived in step 8, we assign the sample people with positive PRE_LNIAW_02A03 to it (the benchmark cell/collapsed-cell-group under consideration).
- Based on the discussion in section 3.0, we could assume that, for a given type of a universe leaver, the probability of each individual belonging to a given cell/collapsed-cell-group is practically constant (identical). Therefore, for each sample person in each cell/collapsed-cell-group assign a unique sub-interval in the zero-to-one interval with a length equal to the ratio of {his/her PRE_LNIAW_02A03A} to {the sum of the PRE_LNIAW_02A03A of all sample people including him/her in his/her cell/collapsed-cell-group}.
- Next, we start with simulating the deceased universe leaver in each of the benchmark cells/collapsed-cell-groups derived from tables 1A to 1D in the following manner. We assign (a) the longitudinally interviewed (LIS_02A03A = YS) CPS-ASEC sample people and (b) the CPS-ASEC sample people with LIS_02A03A = YN (their universe leaver status to be simulated) in each of the deceased universe leaver benchmark cell/collapsed-cell-group established in step 8. We then select a number of people from the CPS-ASEC sample people with LIS_02A03A = YN with probability of selection proportional to their weights PRE_LNIAW_02A03A such that the selected people produce a weighted estimate of the deceased rate matches the deceased rate calculated from SIPP data. This can be algorithmically implemented in the following manner: We generate a sequence of random numbers (zero-to-one uniform distribution realizations) for each of these benchmark cells/collapsed-cell-groups such that the size of the sequence is substantially larger than the number of the sample people (assigned) in the cell/collapsed-cell-group that the sequence is generated for. For each of these cells/collapsed-cell-groups, start from the top (first element) of its random number sequence sequential find the first random number that falls into a sub-interval-in-the-zero-to-one-interval that belongs to a sample person with LIS_02A03A = YN then assign his/her simulated universe leaver status as a deceased universe leaver. Proceed in this manner to select the second, third, … simulated deceased universe leavers from the sample people with LIS_02A03A = YN until the ratio of {the sum of the weights PRE_LNIAW_02A03A of the simulated deceased universe leavers in the cell/collapsed-cell-group under consideration} to {the sum of the weights PRE_LNIAW_02A03A of all sample people in the cell/collapsed-cell-group under consideration} is equal to or larger than the deceased universe leaver rate benchmark estimate for the cell/collapsed-cell-group under consideration, for the first time.

- After we have finished with simulating the deceased universe leavers, we proceed to simulate the barracked universe leaver in the same manner as simulating the deceased universe leavers with one modification that a sample person (with LIS_02A03A = YN) who has already been selected as a simulated deceased universe leaver cannot be selected as a simulated barracked universe leaver.
- After the completion of simulating the barracked universe leavers, we proceed to simulate the emigrated universe leaver in the same manner as simulating the deceased universe leavers with one modification that a sample person (with LIS_02A03A = YN) who has already been selected as a simulated deceased or barracked universe leaver cannot be selected as a simulated emigrated universe leaver.
- After the completion of simulating the emigrated universe leavers, we proceed to simulate the institutionalized universe leaver in the same manner as simulating the deceased universe leavers with one modification that a sample person (with LIS_02A03A = YN) who has already been selected as a simulated deceased or barracked or emigrated universe leaver cannot be selected as a simulated institutionalized universe leaver.
- After the completion of simulating the institutionalized universe leavers, all sample people with LIS_02A03A = YN but not selected as a simulated deceased or barracked or emigrated or institutionalized universe leaver are, by definition, simulated movers.

From the universe leaver simulation results obtained above, we create a variable denoted by SUL_TYPE_02A03A to identify the simulated universe leaver statuses of the sample people in the 2002 and 2003 CPS-ASEC quasi-longitudinal dataset in the following manner. We assign

- SUL_TYPE_02A03A = SD to imply that a sample person is a simulated deceased universe leaver.
- SUL_TYPE_02A03A = SB to imply that a sample person is a simulated barracked universe leaver.
- SUL_TYPE_02A03A = SE to imply that a sample person is a simulated emigrated universe leaver.
- SUL_TYPE_02A03A = SI to imply that a sample person is a simulated institutionalized universe leaver.
- SUL_TYPE_02A03A = SM to imply that a sample person is a simulated mover.
- SUL_TYPE_02A03A = EX to imply that a sample person is not a simulated deceased or barracked or emigrated or institutionalized universe leaver or a simulated mover.

We add this variable (SUL_TYPE_02A03A) to the 2002 and 2003 CPS-ASEC quasi-longitudinal dataset for use in the next two steps.

**Step 12 – Calculating Post-universe-leaver-simulation March 2002 CPS-ASEC Longitudinal Non-interview Adjusted Weight**

By definition, (a) a sample person who is a simulated deceased or barracked or emigrated or institutionalized universe leaver (his/her SUL_TYPE_02A03A = SD or SB or SE or SI) is a March 2002 CPS-ASEC longitudinal interviewed person, and (b) a sample person who is a simulated mover (his/her SUL_TYPE_02A03A = SM) is a March 2002 CPS-ASEC longitudinal non-interviewed person. However, in calculating the pre-universe-leaver-simulation March 2002 CPS-ASEC longitudinal non-interview adjusted weight PRE_LNIAW_02A03A in step 10, each of the sample people with SUL_TYPE_02A03A was treated as though he/she were a March 2002 CPS-ASEC longitudinal interviewed person. To account for the change in the longitudinal interview status of these people after the universe leaver simulation, we need to perform another longitudinal non-interview weight adjustment to transfer the weights PRE_LNIAW_02A03A from the simulated movers to the actual March 2002 CPS-ASEC longitudinal interviewed sample people (LIS_02A03A = YS) and the simulated universe leavers (SUL_TYPE_02A03A = SM or SB or SE or SI). We perform this longitudinal non-interview weight adjustment using the same procedure in step 10 with modification that analytically the pre-universe-leaver-simulation March 2002 CPS-ASEC longitudinal non-interview adjusted weight PRE_LNIAW_02A03A becomes the initial longitudinal weight. Hereinafter, we refer to the longitudinal non-interview adjusted weight produced in this step as *the post-universe-leaver-simulation March 2002 CPS-ASEC longitudinal non-interview adjusted weight*, and denote it by PST_LNIAW_02A03A.

As a result of this longitudinal non-interview weight adjustment, only the actual March 2002 CPS-ASEC longitudinal interviewed sample people (LIS_02A03A = YS) and the simulated universe leaver sample people (SUL_TYPE_02A03A = SD or SB or SE or SI) have a positive PST_LNIAW_02A03A, and all other sample people have zero PST_LNIAW_02A03A (PST_LNIAW_02A03A = 0). Namely, at this point, we let the sample people with positive PRE_LNIAW_02A03A temporarily provide a full representation of the March 2002 CPS-ASEC/SIPP longitudinal universe.

Because of the difference between the weights PST_LNIAW_02A03A and PRE_LNIAW_02A03A, the universe leaver rate estimate for each of the universe leaver benchmark cells/collapsed-cell-groups (derived from tables 1A to 1D, 2A and 2B, 3A and 3B, and 4A to 4D) produced by the weight PST_LNIAW_02A03A is also different from its universe leaver benchmark rate (established in step 8). Therefore, we need to modify (adjust) the weight PST_LNIAW_02A03A such that the modified PST_LNIAW_02A03A produces a universe leaver rate estimate for each of the universe leaver benchmark cells/collapsed-cell-groups that is practically equal its universe leaver benchmark rate. In addition, the modified PST_LNIAW_02A03A must be able to practically preserve the original degree of representation (coverage) provided by the PST_LNIAW_02A03A for each of (a) the universe leaver benchmark cells/collapsed-cell-groups (derived from tables 1A to 1D, 2A and 2B, 3A and 3B, and 4A to 4D in step 8) and the longitudinal non-interview weight adjustment cell/collapsed-cell-groups derived from the longitudinal non-interview weight adjustment table provided in SIPP Branch (2002) in this step. This additional constraint on the modified PST_LNIAW_02A03A implies that (a) the sum of the PST_LNIAW_02A03A and the sum of the modified PST_LNIAW_02A03A of the sample people in each of the universe leaver benchmark cells/collapsed-cell-groups are practically equal, and (b) the sum of the PST_LNIAW_02A03A and the sum of the modified PST_LNIAW_02A03A of the sample people in each of the longitudinal non-interview weight adjustment cell/collapsed-cell-groups derived from the longitudinal non-interview weight adjustment table provided in SIPP Branch (2002) are practically equal. We judiciously determine that, herein, “practically equal” implies “within ±10% difference”. This criterion of being practically equal is used in calculating the modified PST_LNIAW_02A03A in next step. Hereinafter, for simplicity, we make the following terminologies for use.

- We refer to the sum of the weights PST_LNIAW_02A03A in each universe leaver benchmark cells/collapsed-cell-groups (derived from tables 1A to 1D, 2A and 2B, 3A and 3B, and 4A to 4D) as *the benchmark weight sum of a universe leaver benchmark cell/collapsed-cell-group*.
- We refer to each of the longitudinal non-interview weight adjustment cells/collapsed-cell-groups (derived from the longitudinal non-interview weight adjustment table provided in SIPP Branch (2002)) as *a longitudinal non-interview weight adjustment benchmark cell/collapsed-cell-group*. Accordingly, we refer to a table containing all longitudinal non-interview weight adjustment benchmark cells/collapsed-cell-groups as *a longitudinal non-interview weight adjustment benchmark table*.
- We refer to the sum of the weight PST_LNIAW_02A03A in each of longitudinal non-interview weight adjustment benchmark cells/collapsed-cell-groups as *the benchmark weight sum of a longitudinal non-interview weight adjustment benchmark cell/collapsed-cell-group*.

Based on the above discussion, in this step, we retain (keep) the longitudinal non-interview weight adjustment benchmark table and the benchmark weight sum for each of its longitudinal non-interview weight adjustment benchmark cells/collapsed-cell-groups (obtained in the longitudinal non-interview weight adjustment process in this step) for use in the next step. In addition, we also calculate, in this step, the benchmark weight sum for each of the universe leaver benchmark cells/collapsed-cell-groups (established in step 8) for use in next step.

**Step 13 – Modifying the Post-Universe-Leaver-Simulation March 2002 CPS-ASEC Longitudinal Non-interviewed Adjusted Weight**

Based on the discussion in step 12, we need to modify (adjust) the post-universe-leaver-simulation March 2002 CPS-ASEC longitudinal non-interview adjusted weight PST_LNIAW_02A03A such that it satisfied the following three conditions.

- Condition 1 – For each of the universe leaver benchmark cells/collapsed-cell-groups (established in step 8), the sum of the modified PST_LNIAW_02A03A of its sample people is within ±10% of its benchmark sum (calculated in step 12).
- Condition 2 – For each of the universe leaver benchmark cells/collapsed-cell-groups (established in step 8), the ratio of {the sum of the modified PST_LNIAW_02A03A of its simulated universe leaver sample people} to {the sum of the modified PST_LNIAW_02A03A of its sample people} is within ±10% of its universe leaver benchmark rate (calculated in step 8).
- Condition 3 – For each of the longitudinal non-interview weight adjustment benchmark cells/collapsed-cell-groups (established in step 12), the sum of the modified PST_LNIAW_02A03A of its sample people is within ±10% of its benchmark sum (calculated in step 12).

We remark that (a) conditions 1 and 2 simultaneously imposed on each of the universe leaver benchmark cells/collapsed-cell-groups, and (b) for a given type of universe leaver (e.g., deceased universe leavers), the sample people in each of its benchmark cells/collapsed-cell-groups (derived in step 8) are mutually exclusive from one another. Therefore, for each of the universe leaver benchmark cells/collapsed-cell-groups of a given type of universe leavers, we can derive two weight adjustment factors to produce a set of adjusted to-be-modified weights for its sample people that simultaneously satisfy conditions 1 and 2 in an exact manner (note that in exact manner implies 0% difference). Let X and Y denote these two factors. Based on the derivation in Sae-Ung and Sissel (2007), these factors (X and Y) can be expressed as shown in equations 1 and 2 below.

$$X = (UL\_BMR \times UL\_BMS) \div A \qquad \text{Equation 1}$$

$$Y = \{UL\_BMS \times (1 - UL\_BMR)\} \div B \qquad \text{Equation 2}$$

In equations 1 and 2, (a) UL_BMR denotes the universe leaver benchmark rate of the universe leaver benchmark cell/collapsed-cell-group for the universe leaver type under consideration, (b) UL_BMS denotes the benchmark weight sum of the universe leaver benchmark cell/collapsed-cell-group for the universe leaver type under consideration, (c) A denotes the sum of the to-be-modified weights of the universe-leaver-of-the-type-under-consideration sample people in the universe leaver benchmark cell/collapsed-cell-group for the universe leaver type under consideration, and (d) B denotes the sum of the to-be-modified weights of the *not* universe-leaver-of-the-type-under-consideration sample people in the universe leaver benchmark cell/collapsed-cell-group for the universe leaver type under consideration. For each of the universe-leaver-of-the-type-under-consideration sample people in the universe leaver benchmark cell/collapsed-cell-group for the universe leaver type under consideration, let X_UL_TBMW denote his/her to-be-modified weight, and A_X_UL_TBMW denote his/her adjusted to-be-modified weight produced from this weight adjustment. By definition, his/her A_X_UL_TBMW can be expressed as shown in equation 3 below.

$$A\_X\_UL\_TBMW = X \times X\_UL\_TBMW \qquad \text{Equation 3}$$

For each of the *not* universe-leaver-of-the-type-under-consideration sample people in the universe leaver benchmark cell/collapsed-cell-group for the universe leaver type under consideration, let Y_UL_TBMW denote his/her to-be-modified weight, and A_Y_UL_TBMW denote his/her adjusted to-be-modified weight produced from this weight adjustment. By definition, his/her A_Y_UL_TBMW can be expressed as shown in equation 4 below.

$$A\_Y\_UL\_TBMW = Y \times Y\_UL\_TBMW \qquad \text{Equation 4}$$

Similarly, the sample people in each of its longitudinal non-interview weight adjustment benchmark cells/collapsed-cell-groups (established in step 12) are mutually exclusive from one another. Therefore, for each of the longitudinal non-interview weight adjustment benchmark cells/collapsed-cell-groups, we can derive a weight adjustment factor to produce a set of adjusted to-be-modified weights for its sample people that satisfy condition 3 in an exact manner (note that in exact manner implies 0% difference) using equation 5 below.

$$Z = LNI\_BMS \div C \qquad \text{Equation 5}$$

In equation 5, (a) Z denotes the weight adjustment factor that produces a set of adjusted to-be-modified weights for the longitudinal non-interview weight adjustment benchmark cell/collapsed-cell-group under consideration such that it satisfies condition 3 in exact manner, (b) LNI_BMS denotes the benchmark weight sum of the longitudinal non-interview weight adjustment benchmark cell/collapsed-cell-group under consideration, and (c) C denotes the sum of to-be-modified weights of the sample people in the longitudinal non-interview weight adjustment benchmark cell/collapsed-cell-group under consideration.

For each of the sample people in the longitudinal non-interview weight adjustment benchmark cell/collapsed-cell-group under consideration, let Z_LNI_TBMW denote his/her to-be-modified weight, and A_Z_LNI_TBMW denote his/her adjusted to-be-modified weight produced from this weight adjustment. By definition, his/her A_Z_LNI_TBMW can be expressed as shown in equation 6 below.

$$A\_Z\_LNI\_TBMW = Z \times Z\_UL\_TBMW \qquad \text{Equation 6}$$

We remark that the sample people in a universe leaver benchmark cell/collapsed-cell-group for a given type of universe leavers are not necessarily mutually exclusive with <u>either</u> the sample people in any universe leaver benchmark cells/collapsed-cell-groups for any other types of universe leavers <u>or</u> the sample people in any longitudinal non-interview weight adjustment benchmark cells. Therefore, we need to use iterative approach to modify (adjust) the PST_LNIAW_02A03A to produce a set of modified PST_LNIAW_02A03A that satisfy conditions 1 to 3 (defined earlier in this section). We develop an iteration scheme (algorithm) in this regard as briefly described below.

- In the 1st iteration, we start with adjusting (modifying) the PST_LNIAW_02A03A for each of the deceased universe leaver benchmark cells/collapsed-cell-groups (derived from tables 1A to 1D in step 8) using equations 1 to 4 (by taking the PST_LNIAW as the to-be-modified weight). We denote the weight obtained from this adjustment as *the 1st deceased adjustment modified PST_LNIAW_02A03A*.
- We continue the 1st iteration by taking the 1st deceased adjustment modified PST_LNIAW_02A03A as the to-be-modified weight for each of the barracked universe leaver benchmark cells/collapsed-cell-groups (derived from tables 2A and 2B in step 8). We then adjust the1st deceased adjustment modified PST_LNIAW_02A03A for each of the barracked universe leaver benchmark cells/collapsed-cell-groups using equations 1 to 4. We denote the weight obtained from this adjustment as *the 1st barracked adjustment modified PST_LNIAW_02A03A*.
- We continue the 1st iteration by taking the 1st barracked adjustment modified PST_LNIAW_02A03A as the to-be-modified weight for each of the emigrated universe leaver benchmark cells/collapsed-cell-groups (derived from tables 3A and 3B in step 8). We then adjust the 1st barracked adjustment modified PST_LNIAW_02A03A for each of the emigrated universe leaver benchmark cells/collapsed-cell-groups using equations 1 to 4. We denote the weight obtained from this adjustment as *the 1st emigrated adjustment modified PST_LNIAW_02A03A*.
- We continue the 1st iteration by taking the 1st emigrated adjustment modified PST_LNIAW_02A03A as the to-be-modified weight for each of the institutionalized universe leaver benchmark cells/collapsed-cell-groups (derived from tables 4A to 4D in step 8). We then adjust the1st emigrated adjustment modified PST_LNIAW_02A03A for each of the institutionalized universe leaver benchmark cells/collapsed-cell-groups using equations 1 to 4. We denote the weight obtained from this adjustment as *the 1st institutionalized adjustment modified PST_LNIAW_02A03A*.
- We continue the 1st iteration by taking the 1st institutionalized adjustment modified PST_LNIAW_02A03A as the to-be-modified weight for each of the longitudinal non-interview weight adjustment benchmark cells/collapsed-cell-groups (established in step 12). We then adjust the 1st institutionalized adjustment modified PST_LNIAW_02A03A for each of the longitudinal non-interview weight adjustment benchmark cells/collapsed-cell-groups using equations 1 and 2. We denote the weight obtained from this adjustment as *the 1st iteration modified PST_LNIAW_02A03A*.
- Check <u>if</u> the 1st iteration modified PST_LNIAW_02A03A satisfy conditions 1 and 2 for every universe leaver benchmark cells/collapsed-cell-group established in step 8 (note that, algorithmically, the 1st iteration modified PST_LNIAW_02A03A automatically satisfies condition 3 for every longitudinal non-interview weight adjustment cell/collapsed-cell-group established in step 12). If yes, the 1st iteration modified PST_LNIAW_02A03A is our final modified PST_LNIAW_02A03A, and thus we stop our iteration. If no, we then perform the 2nd iteration using the 1st iteration modified PST_LNIAW_02A03A as its to-be-modified weight, the 3rd iteration using the 2nd iteration modified PST_LNIAW_02A03A as its to-be-modified weight, … until we obtain the nth iteration modified PST_LNIAW_02A03A that satisfy conditions 1 and 2 for every universe leaver benchmark cells/collapsed-cell-group established in step 8 (note that, algorithmically, the nth iteration modified PST_LNIAW_02A03A automatically satisfies condition 3 for every longitudinal non-interview weight adjustment cell/collapsed-cell-group established in step 12). Namely, the nth iteration modified PST_LNIAW_02A03A is our final modified PST_LNIAW_02A03A, and thus we stop our iteration.

Hereinafter, we refer to the final modified PST_LNIAW_02A03A (the final modified post-universe-leaver-simulation March 2002 CPS-ASEC longitudinal non-interview adjusted weight) as *the final March 2002 CPS-ASEC longitudinal non-interview adjusted weight*. We denote the final March 2002 CPS-ASEC longitudinal non-interview adjusted weight by FNL_LNIAW_02A03A. Like the post-universe-leaver-simulation March 2002 CPS-ASEC longitudinal non-interview adjusted weight PST_LNIAW_02A03A (calculated in step 12), only the actual March 2002 CPS-ASEC longitudinal interviewed sample people (LIS_02A03A = YS) and the simulated universe leaver sample people (SUL_TYPE_02A03A = SD or SB or SE or SI) have a positive FNL_LNIAW_02A03A, and all other sample people have zero FNL_LNIAW_02A03A (FNL_LNIAW_02A03A = 0).

**Step 14 – Calculating the March 2002 CPS-ASEC Final Longitudinal Weight**

Having obtained the final modified post-universe-leaver-simulation March 2002 CPS-ASEC longitudinal non-interview adjusted weight FNL_LNIAW_02A03A in step 13, we proceed to perform the post-stratification (second-stage) weight adjustment for the FNL_LNIAW_02A03A to produce the March 2002 CPS-ASEC final longitudinal weight using the same procedure used by the SIPP in step 7. Namely, the March 2002 CPS-ASEC final longitudinal weight is obtained from a procedure that adjusts the final March 2002 CPS-ASEC longitudinal non-interview adjusted weight FNL_LNIAW_02A03A to match a set of demographic controls (benchmark estimates) derived from the basic CPS and the cross-section CPS-ASEC data for March 2002. As indicated in SIPP Branch (2002), the demographic characteristics used to form the controls include races, ethnicities, ages, and family types (e.g., female householder with no spouse present but with own children age ≤ 18) to highly reflect more important key characteristics of the SIPP as well as the CPS-ASEC.

Hereinafter, we denote the March 2002 CPS-ASEC final longitudinal weight (produced in this step as described above) A_FLAW_02A03A. Like the final modified post-universe-leaver-simulation March 2002 CPS-ASEC longitudinal non-interview adjusted weight FNL_LNIAW_02A03A (calculated in step 13), only the actual March 2002 CPS-ASEC longitudinal interviewed sample people (LIS_02A03A = YS) and the simulated universe leaver sample people (SUL_TYPE_02A03A = SD or SB or SE or SI) have a positive A_FLAW_02A03A, and all other sample people have zero A_FLAW_02A03A (A_FLAW_02A03A = 0). This implies that the actual March 2002 CPS-ASEC longitudinal interviewed sample people (LIS_02A03A = YS) and the simulated universe leaver sample people (SUL_TYPE_02A03A = SD or SB or SE or SI) together with their final longitudinal weights A_FLAW_02A03A provide a full representation of the March 2002 CPS-ASEC/SIPP longitudinal universe.

**Step 15 – Calculating the Generalized Variance Function Parameters**

Both the CPS-ASEC and SIPP produce their generalized variance function (GVF) parameters (denoted by a and b) based on the GVF model described by equation 7 below.

$$\sigma_x^2 = ax^2 + bx \qquad \text{Equation 7}$$

In equation 7, x denotes an estimated number of units of interest in a survey universe (e.g., number of people having health insurance coverage in 2001), $\sigma_x^2$ denotes variance of the estimated numbers x, and a and b are the GVF parameters derived from a group of estimated numbers for various characteristics that are similar or highly correlated among themselves and the characteristic of the units for which x provides the estimated number.

It is a common practice among analysts to use the GVF parameters (a and b) provided by the CPS-ASEC and SIPP to calculate the standard error of the estimates of their interest. Therefore, in this study, we produce two sets of GVF parameters. Namely, one set of GVF parameters is for the longitudinal estimates derived using the 2001 SIPP Panel longitudinal dataset and the March 2002 SIPP final longitudinal weight (produced in step 7), and another set of GVF parameters is for the longitudinal estimates derived using the 2002 and 2003 CPS-ASEC quasi-longitudinal dataset and the March 2002 CPS-ASEC final longitudinal weight (produced in step 14). Hereinafter, we refer to the GVF parameters for the longitudinal estimates derived using the 2001 SIPP Panel longitudinal dataset and the March 2002 SIPP final longitudinal weight as *the March 2002 SIPP longitudinal GVF parameters*. Similarly, we refer to the GVF parameters for the longitudinal estimates derived using the 2002 and 2003 CPS-ASEC quasi-longitudinal dataset and the March 2002 CPS-ASEC final longitudinal weight as *the March 2002 CPS-ASEC longitudinal GVF parameters*.

In this study, we derive the March 2002 SIPP longitudinal GVF parameters (a and b) from the GVF cross-sectional parameters for waves 4 to 6 provided in SIPP Branch (2005) by scaling them with the ratio of {the effective cross-sectional sample size for the SIPP cross-sectional universe for March 2002} to {the effective longitudinal sample size for the March 2002 CPS-ASEC/SIPP longitudinal universe}. For person-level characteristics, the effective longitudinal sample size for the March 2002 CPS-ASEC/SIPP longitudinal universe is, by definition, the number of sample people in the 2001 SIPP panel with a positive March 2002 SIPP final longitudinal weight (produced in step 7); and the effective cross-sectional sample size for the SIPP cross-sectional universe for March 2002 is, by definition, the number of sample people in the 2001 SIPP Panel with a positive final cross-section weight for March 2002. For household-level characteristics, the effective longitudinal sample size for the March 2002 CPS-ASEC/SIPP longitudinal universe is, by definition, the number of sample household reference persons in the 2001 SIPP panel with a positive March 2002 SIPP final longitudinal weight (produced in step 7); and the effective cross-sectional sample size for the cross-sectional SIPP universe for March 2002 is, by definition, the number of sample household reference persons in the 2001 SIPP Panel with a positive final cross-section weight for March 2002.

Similarly, we derive the March 2002 CPS-ASEC longitudinal GVF parameters (a and b) from the GVF cross-sectional parameters for the 2002 CPS-ASEC Supplement provided in U.S. Bureau of Labor Statistics and U.S. Census Bureau (2002) by scaling them with the ratio of {the effective cross-sectional sample size for the CPS-ASEC cross-sectional universe for March 2002} to {the effective longitudinal sample size for the March 2002 CPS-ASEC/SIPP longitudinal universe}. For person-level characteristics, the effective longitudinal sample size for the March 2002 CPS-ASEC/SIPP longitudinal universe is, by definition, the number of sample people in the 2002 and 2003 CPS-ASEC quasi-longitudinal dataset with a positive March 2002 SIPP final longitudinal weight (produced in step 7); and the effective cross-sectional sample size for the CPS-ASEC cross-sectional universe for March 2002 is, by definition, the number of sample people in the 2002 CPS-ASEC cross-sectional dataset with a positive final cross-section weight for March 2002. For household-level characteristics, the effective longitudinal sample size for the March 2002 CPS-ASEC/SIPP longitudinal universe is, by definition, the number of sample household reference persons in the 2002 and 2003 CPS-ASEC quasi-longitudinal dataset with a positive March 2002 SIPP final longitudinal weight (produced in step 7); and the effective cross-sectional sample size for the CPS-ASEC cross-sectional universe for March 2002 is, by definition, the number of sample household reference persons in the 2002 CPS-ASEC cross-sectional dataset with a positive final cross-section weight for March 2002.

Due to limitation in resource and time, we do not produce any datasets (e.g., replicate weights) or handy computer programs for calculating direct standard error estimation. If the needed resource and time are available in the future, we will provide such datasets and/or handy computer programs using at least one of the following three methods: the Fay's balanced repeated replication (BRR) method, the combine Fay's BRR and successive difference replication (SDR) method, and the Deville's residual technique of linearization (DRL) method as described in Sae-Ung, Hall, and Sissel (2004).

## 5.0 Results of the Study

We applied our methodology described in sections 3.0 and 4.0 to the 2002 and 2003 CPS-ASEC Supplements and their corresponding 2001 SIPP Panel. The results obtained from this application are presented below (result sets 1 to 4).

### Result Set 1 – Longitudinal Attrition and Nonresponse Analysis

By the definition (step 4 in section 4.0), almost all of the March 2002 SIPP longitudinal interviewed sample people are the original (wave 1 interviewed) sample people of the 2001 SIPP Panel. Therefore, our longitudinal attrition and nonresponse analysis for the 2001 SIPP Panel is based only on the original sample people. Hereinafter, we refer to these people as *the SIPP focal original sample people*. In concert with the interview months (February, March, and April) of the 2003 CPS-ASEC Supplement, our longitudinal attrition and nonresponse analysis for the 2001 SIPP Panel covers the period between February 2001 (wave 2) and March 2003 (waves 7 and 8). Our analysis is based on both the weighted and unweighted counts. Results from the unweighted count analysis are theoretically applicable only to the sample itself. Results from the weighted count analysis are applicable to the survey universe. We used the non-interview adjusted weight of reference month 4 in wave 1 as the weight for our analysis. Hereinafter, this weight is referred to as *the SIPP initial weight*.

Similarly, by definition (step 3 in section 4.0), most of the March 2002 CPS-ASEC longitudinal interviewed sample people are from rotations 1 to 4 of the regular March component of the 2002 CPS-ASEC Supplement sample. Therefore, our longitudinal attrition and nonresponse analysis for the 2000 and 2003 CPS-ASEC Supplements is based only on these sample people. Hereinafter, we refer to these people as *the CPS-ASEC focal original sample people*. Based on the period between the interview months of the 2002 and 2003 CPS-ASEC Supplements, our longitudinal and nonresponse analysis for the 2002 and 2003 CPS-SEC Supplements covers the period from March 2002 to March 2003. Similar to the longitudinal attrition and nonresponse analysis for SIPP, our analysis is based on the unweighted and weighted counts. We used the 2002 CPS-ASEC CHIP weight (the cross-sectional weight obtained just prior to the second-stage/post-stratification weight adjustment) for our weighted count analysis. Hereinafter, this weight is simply referred to as *the CPS-ASEC initial weight*.

The longitudinal attrition and nonresponse analysis and its results are provided in table 5A and table 5B for the SIPP and CPS-ASEC, respectively. In the same manner as the current practice in reporting survey attrition and non-interview rates, we did not attempt to make any statistical significance tests for the comparisons of the longitudinal attrition or non-interview rates of the SIPP and CPS-ASEC focal original sample people, presented below.

The overall longitudinal attrition rates of the SIPP and CPS-ASEC focal original sample people are 28.17% and 19.28% for unweighted case (27.84% and 19.39% for weighted case), respectively. As expected, the SIPP's overall longitudinal attrition rate is considerably higher than the CPS-ASEC's one (by about nine percentage points). The black and Hispanic focal original sample people also have similar differences in the SIPP and CPS-ASEC longitudinal attrition rates as shown in table 5A and table 5B. The overall longitudinal attrition rates of the black and Hispanic SIPP focal original sample people are slightly different (35.14% versus 33.72% for unweighted case, and 35.40% versus 34.77% for weighted case) but are much higher than the rate (28.17% for unweighted case, and 27.84% for weighted case) of all SIPP focal original sample people. Similarly, the overall longitudinal attrition rates of the black and Hispanic CPS-ASEC focal original sample people are slightly different (29.75% versus 28.63% for unweighted case, and 31.09% versus 29.57% for weighted case) but are much higher than the rate (19.28% for unweighted case, and 19.39% for weighted case) of all CPS-ASEC focal original sample people. A classification of the longitudinal attrition rates by causes (reasons) of attrition is provided in details in table 5A and table 5B.

The overall longitudinal interview rates of the SIPP and CPS-ASEC focal original sample persons are 82.57% and 80.72% for unweighted case (82.87% and 80.61% for weighted case), respectively. The SIPP's overall longitudinal interview rate is moderately higher than the CPS-ASEC's rate because the SIPP followed the movers for interview but the CPS-ASEC did not, as indicated by the SIPP mover (type D) longitudinal non-interview rate of 2.3% for unweighted case (2.19% for weighted case) versus the CPS-ASEC simulated-mover (type SM) longitudinal non-interview rate of 8.65% for unweighted case (8.39% for weighted case). The SIPP overall longitudinal interview rates of the black and Hispanic focal original sample people are considerably higher than the CPS-ASEC rates as shown in table A and table B. A classification of the longitudinal non-interview rates by causes (reasons) of attrition is provided in details in table 5A and table 5B. As discussed in note 7 of table 5B, the longitudinal attrition and non-interview rates of the CPS-ASEC focal original sample people are the same for the following reason – as defined step 6 of section 4.0, all CPS-ASEC focal original sample people have a positive March 2002 CPS-ASEC cross-sectional weight, thus all of them must be accounted for in the longitudinal non-interview weight adjustment part of the longitudinal weighting process.

As a result of the longitudinal attrition and nonresponse analysis performed in table 5B, we found an imperfection in the longitudinal matching procedure (specified in step 1 of section 4.0) for longitudinal identification of every sample person interviewed in both the 2002 and 2003 CPS-ASEC Supplements. This imperfection resulted in an overall 1.49% of the CPS-ASEC focal original sample people whose interview/non-interview statuses could not be determined (as discussed in note 5 of table 5B). Since the proportion of these sample people is small (1.49% for unweighted case, and 1.59% for weighted case), we judiciously classified each of them as a March 2002 CPS-ASEC longitudinal non-interviewed sample person.

**Result Set 2 – Comparison of SIPP and CPS-ASEC Universe Leaver Estimates with Independent Estimates**

Among the four types of the CPS-ASEC/SIPP survey universe leavers defined in section 3.0, at present, we only have independent estimates for the deceased universe leavers from the National Center for Health Statistics (NCHS) in a report produced by Hoyert, Heron, et al (2006). Therefore, in this study, we only compared the deceased universe leaver estimates from the SIPP survey data and CPS-ASEC simulated data with the independent estimates from the NCHS. In concert with the time period covered by the 2002 and 2003 CPS-ASEC quasi-longitudinal file, the period for the death estimates of interest is between April 2002 and March 2003. The SIPP death estimates for this period were derived from (a) its own survey data, and (b) its final longitudinal weights (produced per the procedure specified in section 4.0). The CPS-ASEC estimates for this period were derived from (a) the simulated death data using the deceased universe leaver probabilistic model described in section 3.1 and the simulation procedure specified in step 11 in section 4.0, and (b) its final longitudinal weights (produced per the procedure specified in section 4.0). The NCHS death estimates were derived from administrative records of deaths in the country. In concert with the CPS-ASEC/SIPP longitudinal universe, the original NCHS death estimates were adjusted to represent the CPS-ASEC/SIPP longitudinal universe by the ratio of the population size in the CPS-ASEC/SIPP survey universe to the size of all population in the country. To be in concert with the period between April 2002 and March 2003, these adjusted NCHS death estimates were then further adjusted in the following manner. We further adjusted each NCHS death estimate for this period by adding 8/12 of the adjusted 2002 NCHS death estimate and 4/12 of the adjusted 2003 NCHS death estimate to obtain each final NCHS deaths estimate between April 2002 and March 2003. Since the NCHS death estimates were derived from administrative records, we judiciously assume that the variance (standard error) of these estimates was negligible small.

Based on the above discussion, we obtained the SIPP, CPS-ASEC, and NCHS death estimates for both sexes, males, and females, in the period between April 2002 and March 2003, as shown in table 6. The differences between the CPS-ASEC and SIPP estimates are 0.86%, 6.78%, and –5.64% for both sexes, males, and females, respectively, and not statistically significant (at 10% significance level test). This implies that the deceased universe leaver probabilistic model specified in section 3.0 performs well as expected. The differences between the SIPP and NCHS death estimates are 3.20%, -5.57%, and 15.06% for both sexes, males, and females, respectively. The differences for both sexes and males are not statistically significant, but the difference for the females is statistically significant. We speculate that the death data for females collected by the 2001 SIPP Panel between April 2002 and March 2003 may not be accurate enough, and this may need further study in future research. The differences between the CPS-ASEC and NCHS death estimates are 4.08%, 0.76%, and 8.58% for both sexes, males, and females, respectively, and not statistically significant. Thus, the death estimates produced by the CPS simulated death data are comparable with the death estimates produced by the NCHS administrative records. We deem this as a validation of our deceased universe leaver probabilistic model (specified in section 3.1).

**Result Set 3 – Comparison of SIPP and CPS-ASEC Poverty Estimates**

Consider the estimates of poverty rates in 2001 among the people in the CPS-ASEC/SIPP longitudinal universe. The poverty rates under consideration are based on 100%, 150%, and 200% official poverty thresholds for families in 2001. We produced two sets of these poverty rates for 11 cohorts of people (all people, all adults, all children, black adult males, …) as listed in table 7. The first set of estimates of poverty rates was derived using the data and final longitudinal weights (produced per the specification in section 4.0) of the 2002 and 2003 CPS-ASEC quasi-longitudinal, and the second set was derived from the data and final longitudinal weights (produced per the specifications in section 4.0) of the 2001 SIPP Panel longitudinal file. These two sets of poverty rates are tabulated in table 7.

As indicated in table 7, the difference between the SIPP and CPS-ASEC estimates of the poverty rates at 100% threshold for each of the 11 cohorts of people is not statistically significant (at 10% significance level test). For example, the SIPP and CPS-ASEC estimates of the poverty rates at 100% threshold for all people are 12.48% (35,178,696 people) and 12.12% (34,183,628 people), respectively, and their relative difference is 2.91%. Based on these comparisons, the 2001 SIPP Panel and the 2002 and 2003 CPS-ASEC Supplements generally provide comparable poverty estimates at 100% poverty threshold. However, as indicated in the statistical test results in table 7, the 2001 SIPP Panel and the 2002 and 2003 CPS-ASEC Supplements generally do not provide comparable poverty estimates at 150% and 200% poverty thresholds. Since the poverty thresholds are predicated on the total family income level, we speculate that a dominant cause for the significant differences in the SIPP and CPS-ASEC poverty estimates at higher poverty thresholds may be as follow. As a family income increases, the effect of the difference in the ways that the two surveys collect their income data becomes more significant (e.g., the SIPP collects monthly data but the CPS-ASEC collects yearly income data). Another cause may stem from the large difference in longitudinal attrition rates between the two surveys (as described in result set 1 in this section and shown in table 5A and 5B). Therefore, we need further study to determine the causes of the difference in the next phase of this project.

**Result Set 4 – Comparison of SIPP and CPS-ASEC Health Insurance Coverage Estimates**

Consider the estimates of rates of health insurance coverage in 2001 among the people in the CPS-ASEC/SIPP longitudinal universe. We define an individual having health insurance coverage in 2001 if he/she was covered by any health insurance for any duration of time in 2001 (for example, if he/she has just one month of health insurance coverage in 2001 the he/she is classified as having health insurance coverage in 2001). We produced two sets of these health insurance coverage rates for 15 cohorts of people (all people, all males, all female, all blacks, …) as listed in table 8. The first set of estimates of health insurance coverage rates was derived using the data and final longitudinal weights (produced per the specification in section 4.0) of the 2002 and 2003 CPS-ASEC quasi-longitudinal, and the second set was derived from the data and final longitudinal weights (produced per the specifications in section 4.0) of the 2001 SIPP Panel longitudinal file. These two sets of health insurance coverage rates are tabulated in table 8.

Among the 15 cohorts of people under consideration, the difference between the SIPP and CPS-ASEC health insurance coverage rate for each of the 15 cohorts of people is statistically significant (at 10% significance level test) for only five cohorts, as indicated in table 8. These five cohorts are the all male, all non-black, all non-black males, Hispanic males, and non-Hispanic male cohorts (with relative difference of –0.25%, -0.49%, 0.61%, -2.26%, and –0.015%, respectively). The relative difference for each of the other 10 cohorts has a range from –1.54% (non-black female cohort) to –13.11% (the female black cohort). Based on the relative differences and the statistical significance test results of the real differences in table 8, the estimates of the 2001 health insurance coverage rates produced by the SIPP are generally higher than those produced by the CPS-ASEC Supplement. Based on the above discussion, the CPS-ASEC Supplement and SIPP generally do not produce comparable estimates for the 2001 health insurance coverage rates. The degrees of difference in estimates are considerably higher among the blacks and Hispanics. Therefore, we need further study to determine the causes of the difference in the next phase of this study.

## 6.0 Discussion and Conclusion

In this study, we identified as discussed in section 1.0 that, for each pair of two consecutive CPS-ASEC Supplements, about one third (instead of about half) of the sample people in the first of two consecutive samples are also interviewed in the second if the two consecutive samples. Thus, these sample people have two consecutive years of data that is necessary for making longitudinal estimates. On the other hand, some of the sample people interviewed in the first of the two consecutive samples leave their households by time their households are in sample for the second of the two consecutive samples because they become a deceased or barracked or emigrated or institutionalized universe leaver or a mover. As a cross-sectional survey, the CPS-ASEC Supplement does not record the reasons that these sample people leave their sample households. Therefore, to be able to properly create longitudinal estimates from two consecutive CPS-ASEC samples, we need to identify that these sample people leave they sample households as a universe leaver or a mover. To fulfill this need, we developed the deceased, barracked, emigrated, and institutionalized universe leaver models as specified in section 3.0 to simulate the reasons that these sample people leave their sample households. The simulation procedure is fully described in section 4.0. Since the main objective of this research project is to provide analytical comparison of the SIPP and CPS-ASEC key longitudinal estimates, we use the SIPP longitudinal weighting process for the longitudinal weight adjustment of the combined micro data of two consecutive CPS-ASEC samples to reduce the bias due to longitudinal attrition/nonresponse. In this study, we refer to the combined micro data of two consecutive CPS-ASEC samples as *a CPS-ASEC quasi-longitudinal file (dataset)*. As a result of the simulation of survey universe leavers and the longitudinal weighting discussed above, we could use the data from a CPS-ASEC quasi-longitudinal file to properly create longitudinal estimates for comparing with those of the SIPP. Based on the essential systematic identities and similarities between the two surveys discussed in sections 1.0 and 2.0 and their identical longitudinal weighting process discussed in section 2.0, we could practically attribute the cause of any differences between the SIPP and CPS-ASEC key longitudinal estimates to just among the following non-sampling error sources: attrition/non-response, sample freshness (aging), recall period, four versus twelve month seam, collection of monthly versus yearly data, and detailed content and structure of the questionnaire. This amounts to a substantial reduction in our effort to search for the cause of any differences between the SIPP and CPS-ASEC key longitudinal estimates.

For the first installment of our analytical comparison of the SIPP and CPS-ASEC key estimates, we applied our methodology described above to the 2002 and 2003 CPS-ASEC quasi-longitudinal file and its corresponding 2001 SIPP longitudinal file. In this application, we developed a set of SAS computer programs for implementing our methodology (specified in section 4.0) based directly on the 2002 and 2003 CPS-ASEC Supplements and its corresponding 2001 SIPP Panel. However, this set of SAS computer programs can be easily adapted for application to other pairs of two consecutive CPS-ASEC Supplements and their corresponding SIPP Panels in the following manner. The adaptation can be done by simply changing (recoding) the names of the input variables specifically for the 2002 and 2003 CPS-Supplements and 2001 SIPP Panel to the names specifically for the pair of two consecutive CPS-ASEC Supplements under consideration and its corresponding SIPP Panel. The results from this application were full presented in section 5.0, and briefly summarized and discussed as shown below.

- The SIPP longitudinal attrition rates are considerably higher than the CPS-ASEC longitudinal attrition rates. Consequently, to determine the sources of any SIPP and CPS-ASEC estimates that are significantly different, we need to consider longitudinal attrition as one of the possible sources of the difference.
- We found that our deceased universe leaver probabilistic model for simulating the CPS-ASEC deceased universe leavers produced the death estimates comparable with the NCHS death estimates (derived based on death administrative records). We deem this as a validation of our deceased universe probabilistic model. We will continue to search for other independent estimates to validate our barracked, emigrated, and institutionalized universe leaver probabilistic models.
- Our comparison of the SIPP and CPS-ASEC poverty estimates for 2001 indicates that both surveys generally produced comparable poverty rates at 100% official poverty threshold but not at 150% and 200% official poverty thresholds. We need further study to determine the causes of the significant difference at higher official poverty thresholds.
- We compared the SIPP and CPS-ASEC health insurance coverage estimates for 2001. We found that the health insurance coverage rates from both surveys generally are not comparable. The degrees of difference in estimates are considerably higher among the blacks and Hispanics. We need further study to determine the causes of the differences.

In the next phase of our study, we plan to fully carry out our analytical comparison of the SIPP and CPS-ASEC key longitudinal estimates produced from the 2002 and 2003 CPS-ASEC quasi-longitudinal file and the 2001 SIPP longitudinal file. The key longitudinal estimates to be considered include: poverty, individual earnings, family income, public (government) assistance program participation, and health insurance coverage. The types of estimates include 2001 calendar year estimates, estimates of gross change between 2001 and 2002, estimates of status transition (exit/entry) between 2001 and 2002 (e.g., estimate of the proportion of people in poverty in 2001 but not in poverty in 2002). We will also continue to search for means and opportunities to refine and further validate our deceased, barracked, emigrated, and institutionalized probabilistic models.

**Source of Estimates and Statistical Accuracy**

The data in this report are from the 2002 and 2003 Annual Social and Economic Supplements (ASEC) to the Current Population Survey (CPS) and the 2001 Survey of Income and Program Participation (SIPP) Panel conducted by the Census Bureau. The population represented in these two surveys is the civilian non-institutionalized population living in the United States, and members of the armed forces living off or on post are included if at least on civilian adult lives in the household. Statistics from surveys are subject to sampling and non-sampling error.

All comparisons presented in this report have taken sampling error into account and meet the U.S. Census Bureau's standards for statistical significance. Nonsampling errors in surveys may be attributed to a variety of sources, such as how the survey was designed, how respondents interpret questions, how able and willing respondents are to provide correct answers, and how accurately the answers are coded and classified. The Census Bureau employs quality control procedures throughout the production process – including the overall design of surveys, the wording of questions, review of the work of interviewers and coders, and statistical review of reports.

For further information on statistical standards and the computation and use of standard errors for this report, contact Sam Sae-Ung, Demographic Statistical Methods Division, at 301-763-4221 or Smanchai.Sae.Ung@census.gov.

**Acknowledgement**

We gratefully thank Patrick E. Flanagan and David V. Hornick for their support and review of this research.

**References**

Ahmed, B., and Robinson, J. (1994). *Estimates of Emigration of Foreign-born Population: 1980-1990*. Working Paper No. 9. Population Division, U.S. Census Bureau. Washington, DC.

Fernandez, E.W. (1995). *Estimation of the Annual Emigration of U.S. Born Persons by Using Foreign Censuses and Selected Administrative Data: Circa 1980*. Working Paper No. 10. Population Division, U.S. Census Bureau. Washington, DC.

U.S. Census Bureau (1998). *Survey of income and Program Participation (SIPP) Quality Profile*. Washington, DC.

U.S. Bureau of Labor Statistics and U.S. Census Bureau (2002). *2002 Current Population Survey Annual Social and Economic Supplement (March) Supplement Documentation*. Washington, DC.

U.S. Bureau of Labor Statistics and U.S. Census Bureau (2002). *Source and Accuracy of the Data for the March 2002 Current Population Survey Microdata File*. http://www.bls.census.gov/cps/ads/2002/S&A_02.pdf. Washington, DC.

SIPP Branch (2002). *SIPP 96: Specifications for Longitudinal Weighting of Sample People (Draft)*. Memorandum from A. R. Tupek, Chief, Demographic Statistical Methods Division (DSMD), for C. E. Bowie, Chief, Demographic Surveys Division, U.S. Census Bureau. Washington, DC.

U.S. Census Bureau (2002). *Statistical Abstract of the United States: Section 10 of National Defense and Veteran Affairs*. Washington, D.C.

Bonczar, T.P. (2003). *Prevalence of Imprisonment in the U.S. Population, 1974-2001*. Special Report, U.S. Bureau of Justice Statistics. Washington, DC.

U.S. Census Bureau (2003). *2003 Current Population Survey Annual Social and Economic Supplement (March) Supplement Documentation*. Washington, DC.

U.S. Census Bureau (2004). *Survey of income and Program Participation (SIPP) Users' Guide (Revised)*. Washington, DC.

Sae-Ung, S., Hall, D.W., and Sissel, C. D. (2004). *Evaluation of Variance Estimation Methods Based on an SPD Longitudinal File*. Proceedings of the Survey Research Methods. American Statistical Association, 4296-4303.

Hoyert, D.L., Heron, M.P., et al (2006). *Deaths: Final Data for 2003*. National Vital Statistics Report, Vol. 54, No. 13. National Center for Health Statistics (NCHS), Center for Disease Control and Prevention (CDC), U.S. Department of Health and Human Services.

HSS Branch (2005). *CPS Supplements: 2005 Weighting Specifications for the CPS Annual Social and Economic(ASEC)Supplement – Draft 3*. Memorandum from A. R. Tupek, Chief, Demographic Statistical Methods Division (DSMD), for C. E. Bowie, Chief, Demographic Surveys Division, U.S. Census Bureau. Washington, DC.

SIPP Branch (2005). *Survey of Income and Program Participation (SIPP) 2001 Panel: Source and Accuracy Statement for Wave 1 – Wave 9 Public Use Files (S&A-2)*. Memorandum from A. R. Tupek, Chief, Demographic Statistical Methods Division (DSMD), for W. C. Odom, Jr., Chief, Administrative and Customer Services Division, U.S. Census Bureau. Washington, DC.

U.S. Bureau of Labor Statistics and U.S. Census Bureau (2006). *Current Population Survey CPS) Design and Methodology Technical Paper 66*. Washington, DC.

Usmilitary.about.com (2006). United States Military Enlistment Standards. Website: http://usmilitary.about.com.

Sae-Ung, S., and Sissel, C.D. (2006). *Procedure for Creating a Quasi-Longitudinal CPS March/ASEC Rectangular SAS Dataset from Two Consecutive Cross-sectional CPS March/ASEC SAS Dataset*. Memorandum for Documentation, Survey of Income and Program Participation (SIPP) Branch, Demographic Statistical Methods Division (DSMD), U.S. Census Bureau. Washington, DC.

Sae-Ung, S., and Sissel, C.D. (2007) *Longitudinal Weighting Procedure for a Quasi-longitudinal CPS March/ASEC Dataset and Its Corresponding SIPP Longitudinal Dataset*. Memorandum for Documentation, Survey of Income and Program Participation (SIPP) Branch, Demographic Statistical Methods Division (DSMD), U.S. Census Bureau. Washington, DC.

Table 1A – The cells and their scale values of the deceased universe leaver probabilistic model for white or Asian or Pacific Islander males, associated with the March 2002 CPS-ASEC/SIPP Longitudinal Survey Universe. The complete table is provided in Table 17A of Sae-Ung and Sissel (2007).

| Cell Number (1) | Age Interval in March 2002 (2) | Annual Family Income in 2001 as Percent of Poverty Threshold (3) | With Health Insurance in 2001 (4) | Scale Value for Gauging Similarity among Cells (5) |
|---|---|---|---|---|
| 1 | 5 to 9 | Less Than 176 | No | 10 |
| 2 | 5 to 9 | Less Than 176 | Yes | 11 |
| 3 | 5 to 9 | 176 to 450 | No | 15 |
| 4 | 5 to 9 | 176 to 450 | Yes | 16 |
| 5 | 5 to 9 | Larger than or Equal to 450 | No | 24 |
| 6 | 5 to 9 | Larger than or Equal to 450 | Yes | 25 |
| 7 | 10-14 | Less Than 176 | No | 50 |
| 8 | 10-14 | Less Than 176 | Yes | 51 |
| 9 | 10-14 | 176 to 450 | No | 55 |
| 10 | 10-14 | 176 to 450 | Yes | 56 |
| 11 | 10-14 | Larger than or Equal to 450 | No | 64 |
| 12 | 10-14 | Larger than or Equal to 450 | Yes | 65 |
| . | . | . | . | . |
| . | . | . | . | . |
| . | . | . | . | . |
| 109 | 85 and Over | Less Than 176 | No | 2640 |
| 110 | 85 and Over | Less Than 176 | Yes | 2665 |
| 111 | 85 and Over | 176 to 450 | No | 2700 |
| 112 | 85 and Over | 176 to 450 | Yes | 2725 |
| 113 | 85 and Over | Larger than or Equal to 450 | No | 2770 |
| 114 | 85 and Over | Larger than or Equal to 450 | Yes | 2795 |

Table 1B – The cells and their scale values of the deceased universe leaver probabilistic model for white or Asian or Pacific Islander females, associated with the March 2002 CPS-ASEC/SIPP Longitudinal Survey Universe. The complete table is provided in Table 17A of Sae-Ung and Sissel (2007).

| Cell Number (1) | Age Interval in March 2002 (2) | Annual Family Income in 2001 as Percent of Poverty Threshold (3) | With Health Insurance in 2001 (4) | Scale Value for Gauging Similarity among Cells (5) |
|---|---|---|---|---|
| 1 | 5 to 9 | Less Than 176 | No | 10 |
| 2 | 5 to 9 | Less Than 176 | Yes | 11 |
| 3 | 5 to 9 | 176 to 450 | No | 15 |
| 4 | 5 to 9 | 176 to 450 | Yes | 16 |
| 5 | 5 to 9 | Larger than or Equal to 450 | No | 24 |
| 6 | 5 to 9 | Larger than or Equal to 450 | Yes | 25 |
| 7 | 10-14 | Less Than 176 | No | 50 |
| 8 | 10-14 | Less Than 176 | Yes | 51 |
| 9 | 10-14 | 176 to 450 | No | 55 |
| 10 | 10-14 | 176 to 450 | Yes | 56 |
| 11 | 10-14 | Larger than or Equal to 450 | No | 64 |
| 12 | 10-14 | Larger than or Equal to 450 | Yes | 65 |
| . | . | . | . | . |
| . | . | . | . | . |
| . | . | . | . | . |
| 109 | 85 and Over | Less Than 176 | No | 3100 |
| 110 | 85 and Over | Less Than 176 | Yes | 3125 |
| 111 | 85 and Over | 176 to 450 | No | 3160 |
| 112 | 85 and Over | 176 to 450 | Yes | 3185 |
| 113 | 85 and Over | Larger than or Equal to 450 | No | 3230 |
| 114 | 85 and Over | Larger than or Equal to 450 | Yes | 3255 |

Table 1C – The cells and their scale values of the deceased universe leaver probabilistic model for black or American Indian or Aleut or Eskimo males, associated with the March 2002 CPS-ASEC/SIPP Longitudinal Survey Universe. The complete table is provided in Table 17A of Sae-Ung and Sissel (2007).

| Cell Number (1) | Age Interval in March 2002 (2) | Annual Family Income in 2001 as Percent of Poverty Threshold (3) | With Health Insurance in 2001 (4) | Scale Value for Gauging Similarity among Cells (5) |
|---|---|---|---|---|
| 1 | 5 to 9 | Less Than 176 | No | 10 |
| 2 | 5 to 9 | Less Than 176 | Yes | 11 |
| 3 | 5 to 9 | 176 to 450 | No | 15 |
| 4 | 5 to 9 | 176 to 450 | Yes | 16 |
| 5 | 5 to 9 | Larger than or Equal to 450 | No | 24 |
| 6 | 5 to 9 | Larger than or Equal to 450 | Yes | 25 |
| 7 | 10-14 | Less Than 176 | No | 50 |
| 8 | 10-14 | Less Than 176 | Yes | 51 |
| 9 | 10-14 | 176 to 450 | No | 55 |
| 10 | 10-14 | 176 to 450 | Yes | 56 |
| 11 | 10-14 | Larger than or Equal to 450 | No | 64 |
| 12 | 10-14 | Larger than or Equal to 450 | Yes | 65 |
| . | . | . | . | . |
| . | . | . | . | . |
| . | . | . | . | . |
| 109 | 85 and Over | Less Than 176 | No | 4010 |
| 110 | 85 and Over | Less Than 176 | Yes | 4035 |
| 111 | 85 and Over | 176 to 450 | No | 4070 |
| 112 | 85 and Over | 176 to 450 | Yes | 4095 |
| 113 | 85 and Over | Larger than or Equal to 450 | No | 4140 |
| 114 | 85 and Over | Larger than or Equal to 450 | Yes | 4165 |

Table 1D – The cells and their scale values of the deceased universe leaver probabilistic model for black or American Indian or Aleut or Eskimo females, associated with the March 2002 CPS-ASEC/SIPP Longitudinal Survey Universe. The complete table is provided in Table 17A of Sae-Ung and Sissel (2007).

| Cell Number (1) | Age Interval in March 2002 (2) | Annual Family Income in 2001 as Percent of Poverty Threshold (3) | With Health Insurance in 2001 (4) | Scale Value for Gauging Similarity among Cells (5) |
|---|---|---|---|---|
| 1 | 5 to 9 | Less Than 176 | No | 10 |
| 2 | 5 to 9 | Less Than 176 | Yes | 11 |
| 3 | 5 to 9 | 176 to 450 | No | 15 |
| 4 | 5 to 9 | 176 to 450 | Yes | 16 |
| 5 | 5 to 9 | Larger than or Equal to 450 | No | 24 |
| 6 | 5 to 9 | Larger than or Equal to 450 | Yes | 25 |
| 7 | 10-14 | Less Than 176 | No | 50 |
| 8 | 10-14 | Less Than 176 | Yes | 51 |
| 9 | 10-14 | 176 to 450 | No | 55 |
| 10 | 10-14 | 176 to 450 | Yes | 56 |
| 11 | 10-14 | Larger than or Equal to 450 | No | 64 |
| 12 | 10-14 | Larger than or Equal to 450 | Yes | 65 |
| . | . | . | . | . |
| . | . | . | . | . |
| . | . | . | . | . |
| 109 | 85 and Over | Less Than 176 | No | 3220 |
| 110 | 85 and Over | Less Than 176 | Yes | 3245 |
| 111 | 85 and Over | 176 to 450 | No | 3280 |
| 112 | 85 and Over | 176 to 450 | Yes | 3305 |
| 113 | 85 and Over | Larger than or Equal to 450 | No | 3350 |
| 114 | 85 and Over | Larger than or Equal to 450 | Yes | 3375 |

Table 2A – The cells and their scale values of the barracked universe leaver probabilistic model for black and non-black males, associated with the March 2002 CPS-ASEC/SIPP Longitudinal Survey Universe.

| Cell Number | Race | Age Interval in March 2002 | Scale Factor |
|---|---|---|---|
| 1 | Non-black | 36-59* | 10 |
| 2 | Non-black | 16*-35 | 15 |
| 3 | Black | 16*-35 | 100 |
| 4 | Black | 36-59* | 120 |

*Based on Usmilitary.about.com (2006), the minimum age and maximum age for enlistment is 17-59, respectively. However, the minimum age limit of 16 in March 2002 is used to accounted for a 16 year old individual in March 2002 may become 17 years old between March 2002 and March 2003 that is the period for determination of a individual becomes a survey universe leaver.

Table 2B – The cells and their scale values of the barracked universe leaver probabilistic model for black and non-black females, associated with the March 2002 CPS-ASEC/SIPP Longitudinal Survey Universe.

| Cell Number | Race | Age Interval in March 2002 | Scale Factor |
|---|---|---|---|
| 1 | Non-black | 36-59* | 10 |
| 2 | Non-black | 16*-35 | 15 |
| 3 | Black | 16*-35 | 100 |
| 4 | Black | 36-59* | 120 |

*Based on Usmilitary.about.com (2006), the minimum age and maximum age for enlistment is 17-59, respectively. However, the minimum age limit of 16 in March 2002 is used to accounted for a 16 year old individual in March 2002 may become 17 years old between March 2002 and March 2003 that is the period for determination of a individual becomes a survey universe leaver.

Table 3A – The cells and their scale values of the emigrated universe leaver probabilistic model for U.S. born residents, associated with the March 2002 CPS-ASEC/SIPP Longitudinal Survey Universe.

| Cell Number | Age Interval in March 2002 | Scale Factor |
|---|---|---|
| 1 | 0 – 14 | 10 |
| 2 | 15 – 29 | 20 |
| 3 | 40 and Over | 30 |

Table 3B – The cells and their scale values of the emigrated universe leaver probabilistic model for foreign born residents, associated with the March 2002 CPS-ASEC/SIPP Longitudinal Survey Universe.

| Cell Number | Age Interval in March 2002 | Scale Factor |
|---|---|---|
| 1 | 0 – 14 | 10 |
| 2 | 15 – 29 | 20 |
| 3 | 40 and Over | 30 |

Table 4A – The cells and their scale values of the institutionalized universe leaver probabilistic model for non-black males, associated with the March 2002 CPS-ASEC/SIPP Longitudinal Survey Universe.

| Cell Number | Age Interval in March 2002 | Scale Factor |
|---|---|---|
| 1 | 0 – 14 | 10 |
| 2 | 15 – 25 | 20 |
| 3 | 26 – 54 | 40 |
| 4 | 55 – 64 | 60 |
| 5 | 65 and Over | 70 |

Table 4B – The cells and their scale values of the institutionalized universe leaver probabilistic model for non-black females, associated with the March 2002 CPS-ASEC/SIPP Longitudinal Survey Universe.

| Cell Number | Age Interval in March 2002 | Scale Factor |
|---|---|---|
| 1 | 0 – 14 | 10 |
| 2 | 15 – 25 | 20 |
| 3 | 26 – 54 | 40 |
| 4 | 55 – 64 | 60 |
| 5 | 65 and Over | 70 |

Table 4C – The cells and their scale values of the institutionalized universe leaver probabilistic model for black males, associated with the March 2002 CPS-ASEC/SIPP Longitudinal Survey Universe.

| Cell Number | Age Interval in March 2002 | Scale Factor |
|---|---|---|
| 1 | 0 – 14 | 10 |
| 2 | 15 – 25 | 20 |
| 3 | 26 – 54 | 40 |
| 4 | 55 – 64 | 60 |
| 5 | 65 and Over | 70 |

Table 4D – The cells and their scale values of the institutionalized universe leaver probabilistic model for black females, associated with the March 2002 CPS-ASEC/SIPP Longitudinal Survey Universe.

| Cell Number | Age Interval in March 2002 | Scale Factor |
|---|---|---|
| 1 | 0 – 14 | 10 |
| 2 | 15 – 25 | 20 |
| 3 | 26 – 54 | 40 |
| 4 | 55 – 64 | 60 |
| 5 | 65 and Over | 70 |

Table 5A – Longitudinal attrition and nonresponse analysis for the SIPP focal original sample people (defined in section 5.0) in the 2001 SIPP Panel during the period between January 2001 (wave 1) and March 2003 (waves 7 and 8). The weighted count and rate are based on the SIPP initial weight (defined in section 5.0).

| Items under Consideration | All People | | | All Blacks | | | All Hispanics | | |
|---|---|---|---|---|---|---|---|---|---|
| | Both Sexes | Male | Female | Both Sexes | Male | Female | Both Sexes | Male | Female |
| *Focal Original Sample People* | | | | | | | | | |
| Unweighted Count | 75,060 | 35,964 | 39,096 | 10,656 | 4,743 | 5,913 | 9,572 | 4,654 | 4,918 |
| Weighted Count | 207,723,181 | 99,635,041 | 108,088,140 | 25,563,139 | 11,373,559 | 14,189,581 | 23,751,138 | 11,519,275 | 12,231,861 |
| *Longitudinal Interview (defined in section 4.0)* | | | | | | | | | |
| Unweighted Count | 53,022 | 25,238 | 27,784 | 6,801 | 2,975 | 3,826 | 6,153 | 2,960 | 3,193 |
| Weighted Count | 147,409,623 | 70,244,242 | 77,165,381 | 16,254,511 | 7,098,710 | 9,155,802 | 15,331,958 | 7,345,326 | 7,986,631 |
| *Longitudinal Interview until Becoming an Universe Leaver (defined in section 4.0)* | | | | | | | | | |
| Unweighted Count | 893 | 514 | 379 | 110 | 67 | 43 | 191 | 114 | 77 |
| Weighted Count | 2,487,941 | 1,422,995 | 1,064,946 | 258,871 | 158,215 | 100,656 | 468,557 | 279,901 | 188,655 |
| *Longitudinal Non-interview Type A*(1) | | | | | | | | | |
| Unweighted Count | 8,990 | 4,256 | 4,734 | 1,522 | 656 | 866 | 1,282 | 603 | 679 |
| Weighted Count | 24,774,470 | 11,788,067 | 12,986,394 | 3,642,603 | 1,573,178 | 2,069,425 | 3,187,858 | 1,512,658 | 1,675,201 |
| *Longitudinal Non-interview Type D*(2) | | | | | | | | | |
| Unweighted Count | 1,545 | 769 | 776 | 375 | 169 | 206 | 307 | 162 | 145 |
| Weighted Count | 3,970,084 | 1,996,435 | 1,973,649 | 887,768 | 413,259 | 474,509 | 693,184 | 365,477 | 327,707 |
| *Longitudinal Non-interview Type O*(3) | | | | | | | | | |
| Unweighted Count | 848 | 469 | 379 | 162 | 94 | 68 | 195 | 119 | 76 |
| Weighted Count | 2,246,913 | 1,227,418 | 1,019,495 | 379,599 | 219,787 | 159,812 | 468,103 | 286,893 | 181,210 |
| *Excluded from Longitudinal Weighting Process (defined in section 4.0) Type A*(4) | | | | | | | | | |
| Unweighted Count | 7,142 | 3,384 | 3,758 | 1,077 | 475 | 602 | 801 | 345 | 456 |
| Weighted Count | 19,994,316 | 9,486,386 | 10,507,930 | 2,683,549 | 1,175,708 | 1,507,841 | 1,996,043 | 860,952 | 1,135,090 |
| *Excluded from Longitudinal Weighting Process Type D*(5) | | | | | | | | | |
| Unweighted Count | 1,656 | 843 | 813 | 435 | 209 | 226 | 437 | 229 | 208 |
| Weighted Count | 4,276,298 | 2,174,065 | 2,102,233 | 1,028,503 | 492,627 | 535,876 | 1,102,995 | 575,113 | 527,882 |
| *Excluded from Longitudinal Weighting Process Type O*(6) | | | | | | | | | |
| Unweighted Count | 964 | 491 | 473 | 174 | 98 | 76 | 206 | 122 | 84 |
| Weighted Count | 2,563,536 | 1,295,424 | 1,268,112 | 427,735 | 242,075 | 185,660 | 502,440 | 292,955 | 209,485 |
| *Overall Longitudinal Attrition Rate*(7) | | | | | | | | | |
| Unweighted Count | 28.17% | 28.40% | 27.96% | 35.14% | 35.85% | 34.57% | 33.72% | 33.95% | 33.51% |
| Weighted Count | 27.84% | 28.07% | 27.62% | 35.40% | 36.19% | 34.77% | 33.47% | 33.80% | 33.16% |

| Items under Consideration | All People | | | All Blacks | | | All Hispanics | | |
|---|---|---|---|---|---|---|---|---|---|
| | Both Sexes | Male | Female | Both Sexes | Male | Female | Both Sexes | Male | Female |
| *Type A Longitudinal Attrition Rate* | | | | | | | | | |
| Unweighted Count | 21.49% | 21.24% | 21.72% | 24.39% | 23.85% | 24.83% | 21.76% | 20.37% | 23.08% |
| Weighted Count | 21.55% | 21.35% | 21.74% | 24.75% | 24.17% | 25.21% | 21.83% | 20.61% | 22.98% |
| *Type D Longitudinal Attrition Rate* | | | | | | | | | |
| Unweighted Count | 4.26% | 4.48% | 4.06% | 7.60% | 7.97% | 7.31% | 7.77% | 8.40% | 7.18% |
| Weighted Count | 3.97% | 4.19% | 3.77% | 7.50% | 7.96% | 7.12% | 7.56% | 8.17% | 6.99% |
| *Type O Longitudinal Attrition Rate* | | | | | | | | | |
| Unweighted Count | 2.41% | 2.67% | 2.18% | 3.15% | 4.05% | 2.44% | 4.19% | 5.18% | 3.25% |
| Weighted Count | 2.32% | 2.53% | 2.12% | 3.16% | 4.06 | 2.43% | 4.09% | 5.03% | 3.19% |
| *Overall Longitudinal Interview Rate*[(8)] | | | | | | | | | |
| Unweighted Count | 82.57% | 82.42% | 82.71% | 77.05% | 76.80% | 77.24% | 78.05% | 77.67% | 78.42% |
| Weighted Count | 82.87% | 82.68% | 83.04% | 77.08% | 76.69% | 77.39% | 78.42% | 77.89% | 78.92% |
| *Overall Longitudinal Non-interview Rate*[(9)] | | | | | | | | | |
| Unweighted Count | 17.43% | 17.58% | 17.29% | 22.95% | 23.20% | 22.76% | 21.95% | 22.33% | 21.58% |
| Weighted Count | 17.13% | 17.32% | 16.96% | 22.92% | 23.31% | 22.61% | 21.58% | 22.11% | 21.08% |
| *Type A Longitudinal Non-interview Rate* | | | | | | | | | |
| Unweighted Count | 13.77% | 13.62% | 13.90% | 16.97% | 16.56% | 17.29% | 15.77% | 15.23% | 16.28% |
| Weighted Count | 13.70% | 13.60% | 13.78% | 17.00% | 16.62% | 17.30% | 15.82% | 15.45% | 16.71% |
| *Type D Longitudinal Non-interview Rate* | | | | | | | | | |
| Unweighted Count | 2.37% | 2.46% | 2.28% | 4.18% | 4.27% | 4.11% | 3.78% | 4.09% | 3.48% |
| Weighted Count | 2.19% | 2.30% | 2.09% | 4.14% | 4.37% | 3.97% | 3.44% | 3.73% | 3.16% |
| *Type O Longitudinal Non-interview Rate* | | | | | | | | | |
| Unweighted Count | 1.30% | 1.50% | 1.11% | 1.81% | 2.37% | 1.36% | 2.40% | 3.01% | 1.82% |
| Weighted Count | 1.24% | 1.42% | 1.08% | 1.77% | 2.32% | 1.34% | 2.32% | 2.93% | 1.75% |

(1) *Type A March 2002 SIPP longitudinal non-interviewed sample person* is an individual whose household directly or indirectly refused to be interviewed starting at any point during February 2001 (wave 2) and March 2003 (wave 7 or 8) and continued to do so until March 2003.

(2) *Type D March 2002 SIPP longitudinal non-interviewed sample person* is an individual whose household moved to an unknown address starting at any point during February 2001 and March 2003 and continued to remained so until March 2003.

(3) *Type O March 2002 SIPP longitudinal non-interviewed sample person* is an individual who is a March 2002 longitudinal non-interviewed sample person for any reasons other than being a type A or type D March 2002 SIPP longitudinal non-interviewed person.

(4) *Type A March 2002 SIPP longitudinal excluded sample person* is an individual whose household directly or indirectly refused to be interviewed starting at any point during February 2001 and March 2003 and continued to do so until March 2003.

(5) *Type D March 2002 SIPP longitudinal excluded sample person* is an individual whose household moved to an unknown address starting at any point during February 2001 and March 2003 and remained to do so until March 2003.

(6) *Type O March 2002 SIPP longitudinal excluded sample person* is an individual who is a March 2002 longitudinal excluded sample person for any reasons other than being a type A or type D March 2002 SIPP longitudinal excluded person.

(7) *Overall longitudinal attrition rate* = {(type A + type D + type O March 2002 SIPP longitudinal non-interviewed sample people) + (type A + type D + type O March 2002 SIPP longitudinal excluded sample people)} ÷ {(March 2002 SIPP longitudinal interviewed sample people) + (sample people who were March 2002 SIPP longitudinal interviewed until they became a universe leaver) + (type A + type D + type O March 2002 SIPP longitudinal non-interviewed sample people) + (type A + type D + type O March 2002 SIPP longitudinal excluded sample people)}.

(8) *Overall longitudinal interview rate* = {(March 2002 SIPP longitudinal interviewed sample people) + (sample people who were March 2002 SIPP longitudinal interviewed until they became a universe leaver)} ÷ {(March 2002 SIPP longitudinal interviewed sample people) + (sample people who were March 2002 SIPP longitudinal interviewed until they became a universe leaver) + (type A + type D + type O March 2002 SIPP longitudinal non-interviewed sample people)}.

(9) *Overall longitudinal non-interview rate* = {(type A + type D + type O March 2002 SIPP longitudinal non-interviewed sample people)} ÷ {(March 2002 SIPP longitudinal interviewed sample people) + (sample people who were March 2002 SIPP longitudinal interviewed until they became a universe leaver) + (type A + type D + type O March 2002 SIPP longitudinal non-interviewed sample people)}.

Table 5B – Longitudinal attrition and nonresponse analysis for the CPS-ASEC focal original sample people (defined in section 5.0) in the 2001 and 2003 CPS-ASEC quasi-longitudinal file (dataset). The weighted count and rate are based on the CPS-ASEC initial weight (defined in section 5.0).

| Items under Consideration | All People | | | All Blacks | | | All Hispanics | | |
|---|---|---|---|---|---|---|---|---|---|
| | Both Sexes | Male | Female | Both Sexes | Male | Female | Both Sexes | Male | Female |
| *Focal Original Sample People* | | | | | | | | | |
| Unweighted Count | 63,752 | 30,889 | 32,863 | 6,403 | 2,941 | 3.462 | 6,059 | 3,005 | 3,054 |
| Weighted Count | 80,965,806 | 39,039,646 | 41,926,161 | 7,646,541 | 3,520,069 | 4,126,472 | 6,768,053 | 3,351,907 | 3,416,147 |
| *Longitudinal Interview (defined in section 4.0)* | | | | | | | | | |
| Unweighted Count | 50,866 | 24,515 | 26,351 | 4,419 | 1,974 | 2,445 | 4,270 | 2,081 | 2,189 |
| Weighted Count | 64,499,745 | 30,899,262 | 33,600,483 | 5,186,188 | 2,307,164 | 2,879,024 | 4,717,340 | 2,294,619 | 2,422,721 |
| *Longitudinal Interview until Becoming an Simulated Universe Leaver (defined in section 4.0)* | | | | | | | | | |
| Unweighted Count | 592 | 318 | 274 | 79 | 53 | 26 | 73 | 35 | 38 |
| Weighted Count | 768,681 | 411,636 | 357,045 | 82,999 | 55,909 | 27,090 | 81,104 | 40,981 | 40,123 |
| *Longitudinal Non-interview Type A*[(1)] | | | | | | | | | |
| Unweighted Count | 2,670 | 1,286 | 1,384 | 356 | 163 | 193 | 289 | 146 | 143 |
| Weighted Count | 3,553,458 | 1,701,864 | 1,851,594 | 436,553 | 206,575 | 229,978 | 358,814 | 178,635 | 180,179 |
| *Longitudinal Non-interview Type B*[(2)] | | | | | | | | | |
| Unweighted Count | 3,144 | 1,561 | 1,583 | 491 | 218 | 273 | 377 | 204 | 173 |
| Weighted Count | 4,040,441 | 2,027,093 | 2,013,348 | 678,122 | 308,112 | 370,010 | 432,277 | 236,694 | 195,583 |
| *Longitudinal Non-interview Type C*[(3)] | | | | | | | | | |
| Unweighted Count | 16 | 6 | 10 | 5 | 2 | 3 | 2 | 0 | 2 |
| Weighted Count | 22,932 | 6,501 | 16,431 | 8,287 | 2,416 | 5,871 | 1,159 | 0 | 1,159 |
| *Longitudinal Non-interview Type SM*[(4)] | | | | | | | | | |
| Unweighted Count | 5,516 | 2,723 | 2,793 | 911 | 455 | 456 | 938 | 471 | 467 |
| Weighted Count | 6,795,279 | 3,339,647 | 3,455,632 | 1,081,337 | 543,935 | 537,402 | 1,078,102 | 538,895 | 539,208 |
| *Longitudinal Non-interview Type ILM*[(5)] | | | | | | | | | |
| Unweighted Count | 948 | 480 | 468 | 142 | 76 | 66 | 110 | 68 | 42 |
| Weighted Count | 1,285,271 | 653,643 | 631,628 | 173,055 | 95,958 | 77,097 | 99,258 | 62,083 | 37,174 |
| *Overall Longitudinal Interview Rate*[(6)] | | | | | | | | | |
| Unweighted Count | 80.72% | 80.39% | 81.02% | 70.25% | 68.92% | 71.37% | 71.68% | 70.42% | 72.92% |
| Weighted Count | 80.61% | 80.20% | 80.99% | 68.91% | 67.13% | 70.43% | 70.90% | 69.68% | 72.09% |
| *Overall Longitudinal Attrition or Non-interview Rate*[(7)] | | | | | | | | | |
| Unweighted Count | 19.28% | 19.61% | 18.98% | 29.75% | 31.08% | 28.63% | 28.32% | 29.58% | 27.08% |
| Weighted Count | 19.39% | 19.80% | 19.01% | 31.09% | 32.87% | 29.57% | 29.10% | 30.32% | 27.91% |
| *Type A Longitudinal Attrition or Non-interview Rate* | | | | | | | | | |
| Unweighted Count | 4.19% | 4.16% | 4.21% | 5.56% | 5.54% | 5.57& | 4.77% | 4.86% | 4.68% |
| Weighted Count | 4.39% | 4.36% | 4.42% | 5.71% | 5.87% | 5.57% | 5.30% | 5.33% | 5.27% |

| Items under Consideration | All People | | | All Blacks | | | All Hispanics | | |
|---|---|---|---|---|---|---|---|---|---|
| | Both Sexes | Male | Female | Both Sexes | Male | Female | Both Sexes | Male | Female |
| *Type B Longitudinal Attrition or Non-interview Rate* | | | | | | | | | |
| Unweighted Count | 4.93% | 5.05% | 4.82% | 7.67% | 7.41% | 7.89% | 6.22% | 6.79% | 5.66% |
| Weighted Count | 4.99% | 5.19% | 4.80% | 8.87% | 8.75% | 8.97% | 6.39% | 7.06% | 5.73% |
| *Type C Longitudinal Attrition or Non-interview Rate* | | | | | | | | | |
| Unweighted Count | 0.0251% | 0.0194% | 0.0304% | 0.0781% | 0.0680% | 0.0867% | 0.0330% | 0% | 0.0655% |
| Weighted Count | 0.0283% | 0.0167% | 0.0392% | 0.1084% | 0.0686% | 0.1423% | 0.0171% | 0% | 0.0339% |
| *Type SM Attrition or Non-interview Rate* | | | | | | | | | |
| Unweighted Count | 8.65% | 8.82% | 8.50% | 14.23% | 15.47% | 13.17% | 15.48% | 15.67% | 15.29% |
| Weighted Count | 8.39% | 8.55% | 8.24% | 14.14% | 15.45% | 13.02% | 15.93% | 16.08% | 15.78% |
| *Type ILM Longitudinal Attrition or Non-interview Rate* | | | | | | | | | |
| Unweighted Count | 1.49% | 1.55% | 1.42% | 2.22% | 2.58% | 1.91% | 1.82% | 2.26% | 1.38% |
| Weighted Count | 1.59% | 1.67% | 1.51% | 2.26% | 2.73% | 1.87% | 1.47% | 1.85% | 1.09% |

*(1)* *Type A March 2002 CPS-ASEC longitudinal non-interviewed sample person* is an individual whose household directly or indirectly refused to be interviewed for the 2003 CPS-ASEC Supplement.

*(2)* *Type B March 2002 CPS-ASEC longitudinal non-interviewed sample person* is an individual whose household was not interviewed for the 2003 CPS-ASEC Supplement because it was temporarily ineligible (e.g., vacant, under construction, entire household in Armed Forces, etc.)

*(3)* *Type C March 2002 CPS-ASEC longitudinal non-interviewed sample person* is an individual whose household was not interviewed for the 2003 CPS-ASEC Supplement because it was permanently ineligible (e.g., demolished, condemned, permanently converted to permanent business or storage facility, etc.)

*(4)* *Type SM March 2002 CPS-ASEC longitudinal non-interviewed sample person* is an individual who was simulated as a mover within the CPS-March/ASEC survey universe.

*(5)* *Type ILM March 2002 CPS-ASEC longitudinal non-interviewed sample person* is an individual who was interviewed in the 2002 CPS-ASEC Supplement but could not be identified as either longitudinally interviewed or type A or B or C or SM non-interviewed. This situation happens because of imperfection in the longitudinal matching procedure (specified in step 1 of section 4.0) for longitudinally identifying every sample person interviewed in both the2002 and 2003 CPS-ASEC Supplements. In reality, a sample person of this type could be either longitudinally interviewed or non-interviewed. Since this type of sample people are proportionally small (1.49%) among all the CPS-ASEC focal original sample people, we judiciously classified them as longitudinal non-interviewed sample people.

*(6)* *Overall longitudinal interview rate* = {(March 2002 CPS-ASEC longitudinal interviewed sample people) + (sample people who were March 2002 CPS-ASEC longitudinal interviewed until they became a universe leaver in the 2003 CPS-ASEC Supplement)} ÷ {(March 2002 CPS-ASEC longitudinal interviewed sample people) + (sample people who were March 2002 CPS-ASEC longitudinal interviewed until they became a universe leaver in the 2003 CPS-ASEC Supplement) + (type A + type B + type C + type SM + type ILM March 2002 SIPP longitudinal non-interviewed sample people)}.

*(7)* *Overall longitudinal attrition or non-interview rate* = {type A + type B + type C + type SM + type ILM March 2002 SIPP longitudinal non-interviewed sample people} ÷ {(March 2002 CPS-ASEC longitudinal interviewed sample people) + (sample people who were March 2002 CPS-ASEC longitudinal interviewed until they became a universe leaver in the 2003 CPS-ASEC Supplement) + (type A + type B + type C + type SM + type ILM March 2002 SIPP longitudinal non-interviewed sample people)}. *Note that*, in this case, the longitudinal attrition rate and the longitudinal non-interviewed rate are the same for the following reason. As defined step 6 of section 4.0, all CPS-ASEC focal original sample people have a positive March 2002 CPS-ASEC longitudinal weight, thus all of them must be accounted for in the longitudinal non-interview weight adjustment part of the longitudinal weighting process.

Table 6 – Comparison of the SIPP and CPS-ASEC deceased universe leaver estimates with the NCHS death estimates in the period between April 2002 and March 2003.

| Cohort of People | Death Estimates in the CPS-ASEC/SIPP Longitudinal Universe between April 2002 and March 2003 | | |
|---|---|---|---|
| | SIPP Estimate ± Standard Error* | CPS-ASEC Estimate ± Standard Error* | NCHS Estimate ± Standard Error* |
| | | | |
| Both Sexes | 2,500,613 ± 137,648[C1, C4] | 2,522,005 ± 142,935[C1, C7] | 2,423,074 ± 0[C4, C7] |
| Males | 1,315,870 ± 99,798[C2, C5] | 1,404,051 ± 106,915[C2, C8] | 1,393,438 ± 0[C5, C8] |
| Females | 1,184,743 ± 94,792[C3, C6] | 1,117,954 ± 95,463[C3, C9] | 1,029,636 ± 0[C6, C9] |
| | | | |

* The statistical test of significant difference is at 10% significance level, and the statistical test is only for comparing the SIPP, CPS-ASEC, and NCHS estimates in the same row of the table as indicated below.

[C1] Comparison C1 - This pair of SIPP and CPS-ASEC estimates are not statistically different at 10% significance level test.
[C2] Comparison C2 - This pair of SIPP and CPS-ASEC estimates are not statistically different at 10% significance level test.
[C3] Comparison C3 - This pair of SIPP and CPS-ASEC estimates are not statistically different at 10% significance level test.

[C4] Comparison C4 - This pair of NCHS and SIPP estimates are not statistically different at 10% significance level test.
[C5] Comparison C5 - This pair of NCHS and SIPP estimates are not statistically different at 10% significance level test.
[C6] Comparison C6 - This pair of NCHS and SIPP estimates are statistically different at 10% significance level test.

[C7] Comparison C7 - This pair of NCHS and CPS-ASEC estimates are not statistically different at 10% significance level test.
[C8] Comparison C8 - This pair of NCHS and CPS-ASEC estimates are not statistically different at 10% significance level test.
[C9] Comparison C9 - This pair of NCHS and CPS-ASEC estimates are not statistically different at 10% significance level test.

Table 7 – Comparison of the CPS-ASEC and SIPP estimates of poverty rates at 100%, 150%, and 200% official family poverty thresholds among the people in the CPS-ASEC/SIPP longitudinal universe in 2001.

| Cohort of People | Comparison between SIPP and CPS-ASEC Estimates of Poverty Rates in 2001 | | | |
|---|---|---|---|---|
| | Poverty Threshold | SIPP Estimate ± Standard Error | CPS-ASEC Estimate ± Standard Error | Statistical Significance* of the Difference between SIPP and CPS-ASEC Estimates |
| | | | | |
| All People | 100% | 12.48% ± 0.1944% | 12.12% ± 0.2560% | No |
| All Adults (15 and Older) | 100% | 10.60% ± 0.2043% | 10.28% ± 0.2688% | No |
| All Children (14 and Younger) | 100% | 19.45% ± 0.5609% | 18.87% ± 0.5817% | No |
| | | | | |
| Black Adult Males | 100% | 16.35% ± 1.057% | 16.74% ± 1.438% | No |
| Black Adult Females | 100% | 26.08% ± 1.129% | 24.15% ± 1.483% | No |
| Non-black Adult Males | 100% | 7.355% ± 0.2643% | 7.320% ± 0.3517% | No |
| Non-black Adult Females | 100% | 10.78% ± 0.3061% | 10.35% ± 0.4008% | No |
| | | | | |
| Hispanic Adult Males | 100% | 14.80% ± 1.137% | 14.48% ± 1.267% | No |
| Hispanic Adult Females | 100% | 22.08% ± 1.353% | 22.20% ± 1.524% | No |
| Non-Hispanic Adult Males | 100% | 7.416% ± 0.2677% | 7.473% ± 0.3583% | No |
| Non-Hispanic Adult Females | 100% | 11.52% ± 0.3127% | 10.80% ± 0.4054% | No |
| | | | | |
| All People | 150% | 22.46% ± 0.2176% | 21.17% ± 0.2202% | Yes |
| All Adults (15 and Older) | 150% | 19.85% ± 0.2347% | 18.73% ± 0.2373% | Yes |
| All Children (14 and Younger) | 150% | 32.12% ± 0.6617% | 30.10% ± 0.5338% | Yes |
| | | | | |
| Black Adult Males | 150% | 28.93% ± 1.296% | 26.65% ± 1.252% | No |
| Black Adult Females | 150% | 41.35% ± 1.266% | 36.86% ± 1.229% | Yes |
| Non-black Adult Males | 150% | 14.97% ± 0.3202% | 14.36% ± 0.3254% | No |
| Non-black Adult Females | 150% | 20.33% ± 0.3520% | 19.36% ± 0.3572% | Yes |
| | | | | |
| Hispanic Adult Males | 150% | 32.24% ± 0.1497% | 31.41% ± 1.229% | No |
| Hispanic Adult Females | 150% | 41.35% ± 1.606% | 39.12% ± 1.316% | No |
| Non-Hispanic Adult Males | 150% | 14.24% ± 0.3165% | 13.45% ± 0.3194% | Yes |
| Non-Hispanic Adult Females | 150% | 20.65% ± 0.3513% | 19.34% ± 0.3544% | Yes |
| | | | | |
| All People | 200% | 32.65% ± 0.2445% | 30.48% ± 0.2481% | Yes |
| All Adults (15 and Older) | 200% | 29.75% ± 0.2689% | 27.78% ± 0.2724% | Yes |
| All Children (14 and Younger) | 200% | 43.37% ± 0.7023% | 40.39% ± 0.5710% | Yes |
| | | | | |
| Black Adult Males | 200% | 42.77% ± 1.414% | 37.52% ± 1.371% | Yes |
| Black Adult Females | 200% | 54.85% ± 1.279% | 47.85% ± 1.273% | Yes |
| Non-black Adult Males | 200% | 23.84% ± 0.3825% | 22.49% ± 0.3874% | Yes |
| Non-black Adult Females | 200% | 30.22% ± 0.4016% | 28.77% ± 0.4093% | Yes |
| | | | | |
| Hispanic Adult Males | 200% | 48.48% ± 1.600% | 45.60% ± 1.319% | No |
| Hispanic Adult Females | 200% | 55.19% ± 1.622% | 52.94% ± 1.346% | No |
| Non-Hispanic Adult Males | 200% | 22.68% ± 0.3792% | 21.04% ± 0.3816% | Yes |
| Non-Hispanic Adult Females | 200% | 30.55% ± 0.3998% | 28.43% ± 0.4048% | Yes |
| | | | | |

* The statistical test of significant difference is at 10% significance level, and the statistical test is only for comparing the SIPP and CPS-ASEC estimates in the same row of the table.

Table 8 – Comparison of the CPS-ASEC and SIPP estimates of health insurance coverage rates among the people in the CPS-ASEC/SIPP longitudinal universe in 2001 (see the definition of health insurance coverage in 2001 in section 5.0)

| SIPP Estimate ± Standard Error | Comparison between SIPP and CPS-ASEC Estimates of Health Insurance Coverage Rates in 2001 | | | |
|---|---|---|---|---|
| | SIPP Estimate ± Standard Error | CPS-ASEC Estimate ± Standard Error | Statistical Significance* of the Difference between SIPP and CPS-ASEC Estimates | Relative Difference between the SIPP and CPS-ASEC Estimates |
| | | | | |
| All People | 83.63% ± 0.2429% | 85.07% ± 0.1284% | Yes | -1.690% |
| All Males | 83.83% ± 0.3457% | 84.05% ± 0.1888% | No | -0.2542% |
| All Females | 83.44% ± 0.3412% | 86.04% ± 0.1747% | Yes | -3.031% |
| | | | | |
| All Blacks | 72.55% ± 0.7268% | 80.92% ± 0.4368% | Yes | -10.35% |
| Black Males | 73.35% ± 1.057% | 78.88% ± 0.6662% | Yes | -7.009% |
| Black Females | 71.85% ± 1.000% | 82.69% ± 0.5744% | Yes | -13.11% |
| Non-black People | 85.25% ± 0.2492% | 85.68% ± 0.1352% | No | -0.4932% |
| Non-black Males | 85.28% ± 0.3549% | 84.76% ± 0.1976% | No | 0.6122% |
| Non-black Females | 85.23% ± 0.3500% | 86.56% ± 0.1847% | Yes | -1.543% |
| | | | | |
| All Hispanics | 62.52% ± 0.9275% | 66.19% ± 0.4429% | Yes | -5.540% |
| Hispanic Males | 63.20% ± 1.294% | 64.66% ± 0.6269% | No | -2.258% |
| Hispanic Females | 61.82% ± 1.329% | 67.78% ± 0.6248% | Yes | -8.796% |
| Non-Hispanic People | 86.86% ± 0.2381% | 87.95% ± 0.1260% | Yes | -1.247% |
| Non-Hispanic Males | 87.14% ± 0.3385% | 87.16% ± 0.1858% | No | -0.0159% |
| Non-Hispanic Females | 86.58% ± 0.3349% | 88.70% ± 0.1708% | Yes | -2.389% |

*The statistical test of significant difference is at 10% significance level, and the statistical test is only for comparing the SIPP and CPS-ASEC estimates in the same row of the table.