Using the Annual Social and Economic Supplement as Part of a Current Population Survey Panel

Sarah M. Flood[*] and **José Pacas**[*]

Minnesota Population Center, University of Minnesota, 50 Willey Hall, 225 19th Ave S, Minneapolis, MN 55455

Abstract

The Annual Social and Economic Supplement (ASEC) is the most widely used type of Current Population Survey (CPS) data, but it is cumbersome to use the ASEC as part of a longitudinal CPS panel, especially linking to non-March months. In this paper, we detail the challenges associated with linking the ASEC to monthly CPS data, outline the creation of an identifier that links the ASEC and the March Basic Monthly data from 1989 through 2017, and provide substantive examples that illustrate the value of combining the ASEC with monthly data. The variable, MARBASECID, which we created to link ASEC and March monthly CPS data, represents a significant contribution to social and economic data infrastructure, saving individual researchers from having to duplicate the effort required to create linkages between ASEC and monthly CPS data.

Keywords

Current Population Survey; data infrastructure; data integration; linking; panel data

Section 1 – Introduction

The Current Population Survey (CPS) is the primary source of information about the United States' labor force [1]. The data are collected monthly, contain information on social and economic indicators, and are the source from which the monthly national unemployment rate is calculated. In addition to collecting labor force data on a monthly basis, referred to henceforth as the Basic Monthly CPS, additional topical data are also often collected via supplements to the monthly survey. The Annual Social and Economic Supplement (ASEC) is the most commonly used file from the CPS with its rich information about employment, union membership, health insurance, and taxes [1 p11-2]. These data are used to calculate the official poverty rate of the United States and have been used to measure health insurance coverage rates in the post-Affordable Care Act period. In addition to analyzing rich, monthly data along with information on supplemental topics, researchers can link CPS data across time to create short panels. Elsewhere we have described the underuse of CPS for longitudinal research and our creation of CPSID, which facilitates linking CPS Basic

---

[*]Corresponding authors. floo0017@umn.edu, pacas002@umn.edu.

Monthly files across time [2]. Unfortunately, however, it is challenging to include the popular ASEC file as part of the CPS panel.

In this paper, we describe how IPUMS (www.ipums.org), a leader in the production of easily accessible, population-level data, is facilitating the use of the ASEC as part of CPS panel studies in social and economic research. We begin with a brief overview of the CPS and a description of the ASEC, detailing key differences between the ASEC and other months of CPS data and the implications of oversample changes for researchers who want to link ASEC files across time and/or to Basic Monthly CPS files. We then describe our process for creating MARBASECID, which facilitates easy linkages between the ASEC and March Basic Monthly CPS data and drastically simplifies combining ASEC data with other Basic Monthly CPS data to create a panel, and delineate the choices we made when we encountered various obstacles. Finally, we provide substantive examples to illustrate the value of combining the ASEC with other Basic Monthly CPS data.

Section 2 – Brief Overview of the Current Population Survey

Understanding the purpose and design of the CPS is necessary for linking respondents from the ASEC to Basic Monthly CPS data. The primary function of the CPS is to be "the source of the official Government statistics on employment and unemployment" in the United States [3]. These data have been collected on a monthly basis since 1940 when record levels of unemployment during the Great Depression heightened the need for reliable unemployment statistics. To that point there had been little effort to count the number of jobless persons in the country, much less to develop precise definitions and concepts of employment. During the late 1930s, these concepts were developed and adopted for a national survey of households, named the Sample Survey of Unemployment, which was implemented by the Works Progress Administration in 1940 [4,5,6,7]. In 1942, the Census Bureau took over the survey, and in 1948, the survey was renamed to the Current Population Survey, "to reflect the survey's expanding role as a source for data on a wide variety of demographic, social, and economic characteristics of the population" [1]. In short, the CPS has historically been, and continues to be, a monthly labor force survey [1].

The Basic Monthly CPS is a sample representative of the civilian, household-based population of the United States. The CPS samples households[1] (addresses) and surveys their occupants. Since 1953, occupants of households selected for participation in the CPS have been surveyed in four consecutive months, left out of the sample for the following eight months, and then re-interviewed in each of the following four months [5,7]; the rotation pattern is illustrated in Table 1. CPS refers to each interview month as a Month-in-Sample (MIS), and thus there are at most eight MIS observations for a particular household (MIS 1– MIS 8). For example, consider a household that is first interviewed in January of 2001. The

---

[1]Household: "A household consists of all the people who occupy a housing unit. A house, an apartment or other group of rooms, or a single room, is regarded as a housing unit when it is occupied or intended for occupancy as separate living quarters; that is, when the occupants do not live with any other persons in the structure and there is direct access from the outside or through a common hall. A household includes the related family members and all the unrelated people, if any, such as lodgers, foster children, wards, or employees who share the housing unit. A person living alone in a housing unit, or a group of unrelated people sharing a housing unit such as partners or roomers, is also counted as a household. The count of households excludes group quarters." (https://www.census.gov/programs-surveys/cps/technical-documentation/subject-definitions.html#household)

individuals in the household will also be interviewed in February 2001, March 2001, and April 2001. For the following eight months (May 2001–December 2001), they will not be interviewed. The individuals in the household will then be interviewed four more times: January 2002, February 2002, March 2002, and April 2002.

Despite its panel component, researchers typically use these data as repeated cross sections in part due to the effort required to correctly link the data across years. Researchers have documented the difficulties of and strategies for linking CPS monthly data as well as how to link adjacent years of ASEC data [2,8,9,10,11,12,13,14]. Absent, however, is thorough documentation about how to link ASEC and Basic Monthly CPS data. Making these linkages possible will facilitate new research using the panel aspect of the data as well as combining data collected as part of topical supplements fielded in different moths.

Section 2a – CPS Supplements

In addition to the Basic Monthly CPS, supplements to the CPS are frequently fielded. CPS supplements vary widely in scope and type (see Table 11-1 [1] for a complete list of CPS supplements) and usually contain only individuals who also complete the Basic Monthly Survey in the month the supplement is fielded.[2] For example, the Voting and Registration supplement is fielded biennially and administered only to respondents from the November Basic Monthly CPS. Eligibility for participation in the supplement varies, however, meaning that some respondents to a Basic Monthly CPS will not receive the supplemental questionnaire fielded in that month. The Displaced Worker supplement is asked of workers 20 years of age and older who were displaced from their jobs and who were interviewed in the January Basic Monthly CPS. The ASEC is an exception to the rule about supplements being fielded only for respondents who participate in a specific month; this supplement is administered during the March Basic Monthly CPS, but also includes CPS participants from other months who are not scheduled to receive the March Basic Monthly CPS. This unique aspect of the ASEC requires special handling in the process of linking it to other CPS data files.

Section 2b – The "March Supplement"

The most popular CPS supplement is the "March Supplement." Introduced in 1947 as the Annual Demographic File (ADF), the technical name for this supplement since 2003 has been the Annual Social and Economic Supplement (ASEC). Between 1947 and 1955, the ASEC was administered in April and included households from the April Basic Monthly CPS (see Table 2). After 1955, the ASEC was implemented in March and began being commonly referred to as the "March Supplement." Between 1956 and 1975, the ASEC consisted only of respondents from the March Basic Monthly CPS. Over time, the ASEC has expanded to improve the reliability of information about certain subpopulations (i.e., persons of Spanish [Hispanic] origin and low-income children who do not have health

---

[2]There are three exceptions to the Basic Monthly CPS serving as the sampling scheme for CPS supplements. The Housing Vacancy Supplement (HVS) and the American Time Use Survey (ATUS) are surveys that base their samples from the Basic Monthly CPS but do not happen at the time of the Basic Monthly CPS. The HVS collects information on housing units that were vacant at the time of the Basic Monthly CPS. The ATUS collects information on how respondents spend their time and is conducted a few months after a respondent's final Basic Monthly CPS survey. The third is the ASEC described in detail here.

insurance coverage). Currently, the ASEC contains basic monthly demographic and labor force data as well as supplementary data on work experience, income, noncash benefits, and migration [3]. Given the expansion and implementation of the ASEC—drawing the ASEC sample from the March Basic Monthly CPS and households from non-March Basic Monthly CPS samples—there are complications for longitudinal linking of the ASEC with Basic Monthly CPS files [1].

Section 2c – ASEC Oversamples

The ASEC oversampling scheme has important ramifications for researchers who want to link CPS respondents across time. Though all ASEC respondents participate in the Basic Monthly CPS, only ASEC households that were administered the March Basic Monthly CPS can be easily matched. Linking ASEC oversample respondents to their Basic Monthly CPS observations is extremely tedious and labor intensive at best, and, in some cases, impossible. We therefore focus our efforts on matching the March Basic Monthly CPS to the ASEC. Figure 1 graphs the size of the ASEC oversample from 1989 to 2017 and shows that in each year the ASEC is larger than the March Basic Monthly CPS with larger differences when the SCHIP (State Children's Health Insurance Program) oversample is introduced (details on this below). MARBASECID has been created for the 1989 to 2017 period and will be added annually for the most recent ASEC and backward in time (to the 1976 ASEC) so that CPSID can be made available via IPUMS CPS.

From March 1976 through 2000 the Census Bureau increased the reliability of estimates for people of "Spanish origin" by conducting additional interviews with November households (from the previous year) that contained one or more persons of Spanish origin [15]. The ASEC oversample of people of Spanish origin is commonly referred to as the "November Hispanic oversample." The November Hispanic oversample increased the size of the ASEC by about 2,500 households (see the 1989–2000 range in Figure 1). Because of the CPS rotation pattern (4-8-4), all of the Spanish households identified in November are out of the CPS sample when the March Basic Monthly CPS is conducted; no households interviewed in November are eligible for the March interview based on the 4-8-4 CPS rotation pattern. For example, a household in MIS 1 in November will be MIS 4 in February and thus out of the CPS in March (similarly, MIS 5 households in November will be MIS 8 in February). Because the oversample households would not have otherwise been in the ASEC, we refer to these extra visits as MIS 9 and 10, respectively; in the data, however, the Hispanic oversample cases are assigned MIS values between one and eight, making the oversample cases more complicated to identify.[3] During the extra interviews, the November Hispanic oversample receives both the March Basic Monthly CPS and the ASEC, though the responses to the Basic Monthly are never released [1]. Thus, the November Hispanic oversample results in additional cases in the ASEC from other months of the CPS that would not have otherwise been in the March Basic Monthly CPS.

The second CPS sample expansion in 2001 was funded by a Congressional allocation of $10 million annually to the Census Bureau, which included both a general expansion and an

[3]In the ASEC files, all households receive an MIS value of 1–8. In Census documentation, extra interviews are occasionally referred to as MIS 9. For convenience, we refer to the second additional interview for oversample cases as MIS 10.

additional oversample to the ASEC. The expansion was motivated by an interest in producing reliable state-level estimates on low-income children without health insurance and to measure the effects of the State Children's Health Insurance Program (SCHIP) established by Congress in 1997 [7]. The general expansion added 12,000 units to the sample monthly [7]. The Basic Monthly CPS sample size increases were completed between September 2000 and July 2001 as is evident by the increasing sample sizes during this period in Figure 1. While the sample increases were completed in July 2001, the expansion is not evident in the March Basic Monthly CPS until 2002 [7].

In contrast to the November Hispanic oversample, the SCHIP oversample is drawn using two strategies: "split-path" assignment and month-in-sample 9 (MIS 9) assignment. The "split-path" strategy selects respondents from the February Basic Monthly CPS and April Basic Monthly CPS (adjacent months to the March Basic Monthly CPS). February households with MIS 4 and 8 that contain children (18 or younger) or non-White household members complete the ASEC at the time of the February Basic Monthly CPS interview. Similarly, April households in MIS 1 and 5 that include children (18 or younger) or non-White household members receive the ASEC during the April CPS Basic interview. Neither the "split-path" eligible respondents from February or April would have otherwise participated in the ASEC because of the survey's rotation pattern. The term "split-path" thus refers to these February and April cases that would normally have received the supplements assigned for February and April but instead are "split" into the ASEC.

The second set of households in the SCHIP oversample—the MIS 9 households—is administered an extra interview[4] [7]. These cases are contacted for a ninth interview in either February or April. From 2001 to 2003, these households were drawn from the November Basic Monthly CPS of the previous year if they were in MIS 6, 7, or 8 in November *and* they were not part of the November Hispanic oversample *and* they were not Hispanic *and* either had at least one child 18 years or younger or a non-White member. Note that these households would have completed all eight interviews of the CPS rotation pattern by January at the latest (for MIS 6). Starting in 2004, the MIS 9 oversample was chosen from August (MIS 8), September (MIS 8) and October (MIS 8); the same condition applies as before of either having at least one child 18 years or younger or a non-White member and being non-Hispanic [7].

Section 2d – 2014 ASEC Redesign

In 2014, the ASEC included a series of redesigned income and health insurance questions. Three-eighths of the 2014 ASEC sample received the redesigned income questions while the other five-eighths of the sample received the "traditional" income questions [16]. The redesigned health insurance questions were asked of the entire 2014 ASEC sample. To later evaluate the effects of the redesigned health insurance questions, the Census Bureau used a

---

[4]Notice that the November Hispanic oversample respondents are also contacted an extra time outside of their 8 scheduled CPS Basic interviews. In fact, since the November Hispanic oversample includes households from all MIS's, it is possible that a household from the November Hispanic oversample is contacted two extra times. Thus, though the common understanding of the CPS rotation pattern implies that households are interviewed at most 8 times is technically incorrect. As part of the ASEC oversampling, it is possible for Hispanic oversample households to be contacted 10 separate times and MIS-9 oversample households to be contacted 9 times. Unfortunately, MIS values in the public use data do not exceed 8.

split-path assignment to randomly select about 6,000 households from the 2016 and 2017 March Basic Monthly sample to answer the complete pre-2014 ASEC questionnaire [17]. As a result, we cannot locate a subset of March Basic respondents in the ASEC for 2016 and 2017 (see Table 3).

Section 3 – IPUMS CPS Constructed Identifiers for Linking CPS Data

IPUMS CPS (https://cps.ipums.org) is eliminating barriers for researchers who want to use linked CPS data. We are simplifying access to Basic Monthly CPS data and facilitating the linking of CPS observations over time via the creation of a new unique identifier, CPSID [3], available only via IPUMS. Until now the potential of using CPSID to facilitate analyses of CPS panel data has been limited by the absence of the ASEC due to the incompatible aspects of the ASEC compared to the other Basic Monthly CPS data. The creation of MARBASECID makes the linkage between the ASEC and the March Basic Monthly CPS data straightforward, thereby simplifying the analysis of ASEC data as part of a panel of CPS observations. This effort is of enormous value given the widespread use of the ASEC and underutilization of linked CPS data and promises to save the research community countless hours of duplicated effort, to eliminate a huge potential source of error, and to increase replicability of research results.

Section 3a – CPSID

The ASEC is unique among CPS data, as described above, and that has implications for linking, which we describe in the next section. To lay the foundation for our work on MARBASECID, we first outline the procedures for linking CPS Basic Monthly data. Using linking keys available on all public use Basic Monthly CPS files, researchers can link observations over time to create short sixteen-month panels with up to eight observations per person. This work, however, is cumbersome and expensive for each individual researcher to perform independently. The several obstacles researchers face in linking Basic Monthly CPS observations, including recycled identifiers, changing linking keys, and the household rather than the person as a sampling unit, are detailed elsewhere [2].

As a service to users, IPUMS CPS delivers a single unique identifier, CPSID, that lowers the barrier to using repeated observations of individuals from Basic Monthly CPS files as a panel. CPSID uses the original linking keys provided by CPS to match records over time, accounts for the complex CPS rotation pattern, and assigns a new unique identifier to each record in the Basic Monthly CPS. However, CPSID was not initially created for ASEC files because the ASEC lacks all of the linking keys required for matching records to other Basic Monthly CPS files.

Section 3b – MARBASECID

To make CPSID available on the ASEC, we create MARBASECID, a variable we use and deliver to users to match individuals in the ASEC to the March Basic Monthly CPS. MARBASECID is a 10-digit variable on both the March Basic Monthly CPS and ASEC files. MARBASECID consists of two digits (either 00 or 11), a two-digit year, and a six-digit sequence number. Individuals who appear in both the March Basic Monthly CPS and

the ASEC in the same year are assigned 11 as the first two digits in MARBASECID; the six-digit sequence number begins at 000001 and increments by one for each additional person who is in both files. For a matched observation in the 1989 March Basic Monthly CPS, MARBASECID is 11 + 89 + six-digit sequence number. The two-digit number for unlinked observations in both the March Basic Monthly CPS and the ASEC is 00. For unlinked March Basic Monthly CPS individuals, MARBASECID takes the form: 00 + two-digit year + six-digit sequence number starting at 000001 and incrementing by one for each unlinked March Basic Monthly CPS person. For unlinked ASEC individuals, MARBASECID is a concatenation of 00, two-digit year, and a six-digit number starting at 500,000 and incrementing by one for each unlinked ASEC observation. This method ensures that MARBASECID is unique within and across years. For example, a MARBASECID value of 1100012345 is decoded as follows: "11" refers to the individual who is in both the March Basic Monthly CPS and ASEC; "00" refers to the ASEC survey year of 2000; "012345" refers to that household being given the randomly sequenced order number of 12345. Similarly, a MARBASECID of 0098000012 refers to an unlinked person from the March Basic of 1998 while 0098500012 refers to an unlinked person from the 1998 ASEC.

We created MARBASECID using Stata 13 for the 1989 to 2013 period. Given the ease of matching in later years as described below, we matched ASEC and Basic Monthly CPS files and constructed MARBASECID for 2014 forward using a program written in Perl. All of our code for the construction of MARBASECID from 1989 to 2017 is available on GitHub (https://github.com/mnpopcenter/cps-march-asec-linking).

Section 3c – Creating MARBASECID

The creation of MARBASECID is a critical step in the process of attaching CPSID, a unique IPUMS-created identifier, to the ASEC. As discussed previously, CPSID allows researchers to easily and reliably link data across CPS months, including the ASEC. The creation of MARBASECID eliminates the need for individual researchers to perform the tedious and cumbersome process of linking the March Basic Monthly Survey and the ASEC, which is complicated for two primary reasons. First, the variables required to link the ASEC to CPS monthly files are not available for all years on the ASEC. As a result of omitted linking keys and the ASEC oversample, duplicate and false matches are problematic. Second, we speculate, despite the absence of technical documentation to be certain, that the Census Bureau transition to a computer-based interview resulted in more prominent data quality issues for linking across months even if they did not compromise the integrity of each individual month of data.

The algorithm matching the March Basic Monthly CPS to ASEC overcomes these difficulties and allows us to put CPSID on the ASEC for easy linkages to other CPS monthly data. With CPSID on the ASEC, opportunities for using the CPS as a panel multiply since the ASEC is the premier CPS supplement. Theoretically, Census-provided household and person identifiers should be sufficient to link the March Basic Monthly CPS and ASEC files. If that were the case, researchers would have many demographic variables with which to check the validity of matches. Practically, however, the linking keys that should uniquely identify records do not always [9]. Table 3 details by year (1989–2017) the variables used to

link the March Basic Monthly CPS and ASEC, the number of persons in the March Basic Monthly CPS, the number of matches/non-matches to the ASEC, and the number of invalidated matches.[5] From 1994 forward, we validate matches based on AGE, SEX, and RACE (we follow the evaluation of validity using age, sex and race in line with [8]), and we find high validation rates for links made between the March Basic Monthly CPS and ASEC files.

Our strategy for creating MARBASECID depended on the types of problems we encountered linking the March Basic Monthly CPS and ASEC files. From 2005 to 2017, the matching algorithm is very simple. Using the variables listed in Table 3, we can uniquely identify all March Basic Monthly CPS respondents and ASEC respondents. March Basic Monthly CPS observations are easily located in the ASEC except, as described above, for the split-path households in the 2016 and 2017 ASEC.

From 1996 to 2004, the matching algorithm is more complicated because of duplicate records caused by non-unique linking keys. Our strategy for handling the duplicate records is as follows. During the first stage, duplicate records based on the first stage linking keys in Table 3 are identified and flagged in both the March Basic Monthly CPS and the ASEC. Then, March Basic Monthly CPS records that are not uniquely identified are dropped from the file. Within a pair of ASEC duplicates, we keep the duplicate with the lowest H_SEQ (the household identifier created by Census Bureau that is unique within a given survey month) value since these records are part of the March Basic Monthly CPS rather than an ASEC oversample [18]. We then merge the pruned March Basic Monthly CPS and ASEC files using the first stage linking keys in Table 3. The second stage of work uses the observations from the duplicate record file (i.e., the "pruned" observations) and the non-matches from the first round of matching. We link records using as few variables as possible. Even then, the data sometimes require a close analysis of a few observations in order to find the correct match.

The period 2001 to 2004 was especially problematic because of the SCHIP expansion of the ASEC oversample. Though CPS documentation details the variables researchers should use for linking, these variables do not uniquely identify records, thus complicating the process [10]. The 1996 to 2000 period was also problematic for unknown reasons. Complications we encountered in both periods are detailed in Appendix A.

The greatest challenges in creating MARBASECID occur prior to 1996 when data quality problems (e.g., duplicate records based on linking keys and missing observations) are more common. In 1995, we employed the two-stage matching approach and were unable to match 951 observations from the March Basic Monthly CPS to the ASEC. It is possible that the observations are missing from the ASEC, though we have yet to find documentation about this specific issue. To link the 1994 March Basic Monthly CPS and ASEC files, we employ first and second stage matching (Table 3) and also make additional adjustments. The most important adjustment is that the ASEC file must contain a corrected version of HRHHID (the originally released 1994 ASEC file contained an error in HRHHID resulting from the

---

[5]For convenience we provide both IPUMS and original variable names. We refer to IPUMS variable names throughout the text.

program that created the variable [19]; the corrected version of the file is available via IPUMS CPS); matching is impossible without the corrected version of HRHHID. Several other minor adjustments must be made, including harmonizing age for a few observations (e.g., an observation may have age of 81 in the March Basic Monthly CPS but 80+ in the ASEC) and handling duplicate observations; our handling of these issues is detailed in Appendix B.[6]

Section 4 – Linking Research Potential

The ability to easily link CPS observations over time to the ASEC creates opportunities for many lines of research that have previously been inaccessible without deep knowledge of the ASEC and the CPS, more generally. The ASEC is especially important for researchers who want to leverage information about taxes, health insurance, and public benefit use, among other things, because these data are not collected in the CPS outside of the ASEC. Below, we provide two substantive examples of research made possible by linking CPS Basic Monthly data to the ASEC that demonstrate the potential of MARBASECID for the research community.

Because of the CPS rotation pattern, in which each household appears in the CPS up to eight times denoted by their MIS value, researchers can link individuals who participate in the March Basic Monthly CPS and the ASEC to up to three months prior (December, January, February) and up to three months after (April, May, June) (see Table 1).[7] In theory 25% of the December/June Basic Monthly CPS will link to March (respondents in MIS 1, 5 for December of the previous year, respondents in MIS 4, 8 for June of the current year); 50% of the January/May Basic Monthly CPS will link to March (respondents in MIS 1, 2, 5, 6 for January of the current year, respondents in MIS 3, 4, 7, 8 for May of the current year); 75% of the February/April Basic Monthly CPS will link to March (respondents in MIS 1, 2, 3, 5, 6, 7 for February of the current year, respondents in MIS 2, 3, 4, 6, 7, 8 for April of the current year). In practice, however, mobility, mortality, births, and non-response are major issues for the CPS, resulting in actual linkage rates that are lower than possible linkage rates.

Section 4a – Substantive Example #1: Child Tax Credit Receipt and Food Security

The link between food insecurity and low income in the United States has been widely documented [20]. Using the CPS, a researcher could examine the relationship between receiving the Child Tax Credit (CTC) and food security. The CTC reduces the amount of taxes families pay (depending on adjusted gross income) by $1,000 dollars per qualifying child. This tax credit increases the disposable annual income of low-income families, potentially reducing their food insecurity.

Investigating this relationship using CPS requires linking ASEC and December Food Security files since tax credit questions are only asked in the ASEC and food security is only

---

[6]Prior to 1994, Census released very few variables for linking surveys. Despite trying to avoid matching March Basic Monthly CPS and ASEC observations on AGE, SEX, and RACE, we used these variables to uniquely identify and match records between 1989 and 1993.
[7]We do not consider oversample cases here because we have been unable to locate documentation for linking oversample members to their respective Basic Monthly CPS observations.

assessed in the Food Security Supplement. For illustration purposes, we link the 2005, 2006, 2007, and 2008 Food Security Supplement respondents (from December) to their ASEC records in the following year (2006–2009). We use CPSID to link the Food Security Supplement to the March Basic Monthly CPS and MARBASECID to link March Basic Monthly CPS and ASEC observations (see Table 4). Only one quarter of December Basic Monthly CPS respondents (those in MIS1 and 5) are eligible to be linked to the March Basic Monthly CPS using CPSID (see Column 3 of Table 4). We match 90% of eligible observations between the December and March Basic Monthly surveys (or 22% of the entire December Basic Monthly CPS sample [see Table 4, Column 5]). The linkage rates are consistent with other observations four months apart [2].

After linking December and March Basic Monthly CPS observations using CPSID, we use MARBASECID to link to the March Basic Monthly CPS to the ASEC. The March Basic Monthly CPS to ASEC linkage is perfect, but the number of records with both food security and tax responses is slightly lower because the CTC variable is only available for persons 15 and older (Table 4, Column 7). The final linked and eligible sample is about 15% of the December Basic Monthly CPS (Table 4, Column 8) and about 62% of the total MIS 1 and 5 observations from the December Basic Monthly CPS (Table 4, Column 9). Nonetheless, sample sizes for examining the relationship between CTC receipt and food security are large in each of the four years (~20,000 respondents).[8]

One may also use CPS data to examine how the CTC has affected population food security over time. This requires extending the December (MIS 1) to ASEC linkage (MIS 4) forward in time to include the subsequent December (MIS 5) and ASEC (MIS 8) observations. For example, individuals from MIS 1 in December 2005 are linked first to the 2006 March Basic Monthly CPS using CPSID (and then to the 2006 ASEC using MARBASECID); CPSID is then used again to link to December 2006 (when respondents are MIS 5) and then to March 2007 (via CPSID and then to the 2007 ASEC via MARBASECID). About 5% of respondents from each of the December Basic Monthly CPS surveys from 2005–2008, with both food security and child tax credit data in both years, may be linked in this way (see Table 4, Column 11). Of those eligible (MIS 1 in December of a given year), about 70% are linked (see Table 4, Column 12), and the resulting samples are sizeable (~7,000 observations for each year).

Section 4b – Substantive Example #2: Outgoing Rotation Groups and the ASEC

A popular set of employment questions have been asked only of the outgoing rotation groups (ORG) from MIS 4 and 8 of each Basic Monthly Survey. This set of variables is commonly known as the ORG questions or the Earner Study questions. Information is collected on topics such as usual hours worked, hourly wage rate, usual weekly earnings,

---

[8]This very process of linking first the March and December Basic Monthly CPS files and then the December Basic Monthly file to the ASEC file was omitted in a recently published paper on poverty and food insecurity [21]. The authors draw on poverty data from the ASEC and food security data from the December Food Security supplement. Rather than make linkages between the ASEC and Food Security supplement, the authors impute poverty, which is available in the ASEC, for the sample of December respondents in their analysis for whom poverty is not available. Imputation allows the authors to retain more cases since they are not linking to the ASEC, though making the linkages to the ASEC would allow the authors to get exact rather than imputed measures of poverty. At the very least, using CPSID and MARBASECID, the authors could compare the imputed and actual poverty values for the linked sample.

union membership, class of worker, and multiple job holdings (hereafter referred to as the "earner study" questions). Earnings data from the earner study refer to a usual week in the last month while the earnings data collected in the ASEC refer to the "past year." Given these differences, researchers may want to use these variables in combination with ASEC variables. For example, any research question looking at unionization alongside taxes paid or poverty will require the use of both the ORG and the ASEC. Using the earner study variables from the March Basic Monthly CPS with the ASEC, researchers encounter a substantial reduction in observations—typically about 20% of the March Basic Monthly CPS (see Table 5). This limitation may be overcome by linking ASEC and ORG data from surrounding non-March months. By linking the ASEC to subsequent ORG responses from the Basic Monthly CPS in April, May, and June, the number of cases is nearly quadrupled, increasing the power for combining earnings from the ASEC with information about union participation from the Basic Monthly CPS.

Because not all earner study variables are part of the ASEC file, using the full set of earner study variables with the ASEC data requires, at the very least, linking the March Basic Monthly CPS with the ASEC; MARBASECID drastically simplifies this effort. About 25% of March Basic Monthly CPS respondents would have this weekly wage data available since only the outgoing rotations (MIS 4 and 8) respond to the earner study questions. Using CPSID to link ORG data from different months to the ASEC, researchers can easily increase their sample sizes. Table 5 shows for the years 2005–2009 the number of cases in the March Basic, the number of linkages to other months to get ORG data, and the number of cases with earnings data collected for individuals[9] in ORGs. Researchers can leverage the power of the short panel aspect of the CPS and use CPSID to link from March to April, May, and June and MARBASECID to link from March to the ASEC; the resulting sample size for 2005 is 100,042 compared to 26,503 if only ORG data from March are used (Table 5, Column 4). Larger sample sizes, for example, allow for detailed subgroup analyses which would otherwise be limited.

Researchers may also make these kinds of linkages across years to get ASEC data combined with earnings data from two points in time. The process just described would be performed for two points in time, year x and year x+1. Year x and year x+1 are then linked together using CPSID. As Table 5 shows, this is possible for just over one quarter of the March Basic Monthly CPS respondents (26% in 2005),[10] which is substantially higher than if we only linked individuals from MIS 4 of the 2005 March Basic Monthly CPS to 2006, which would be 7% (or 9,548 individuals) of the 2005 March Basic Monthly CPS sample (see Table 5, Column 7). Including individuals in the March Basic Monthly whose earnings data come from April, May, or June increases the total sample size to 35,030 in 2005 (see Table 5, Column 6). Patterns are similar for the 2006–2009 period.

The value of MARBASECID and CPSID is evident in recent research on unionization and poverty [22]. Doe et al. (2017) replicate a previous study using the ASEC union status

---

[9]Recall that only civilians age 15 and older who are currently employed as a wage or salaried worker respond to ORG questions.
[10]Only half of the March Basic Monthly CPS respondents are linkable across years because individuals in MIS 5–8 in year x will not be in the CPS in year x+1.

variable. The limitations of the analysis are acknowledged: "the CPS asks the union membership question only for one-fourth of the sample (the two outgoing rotation groups). As a result the CPS samples are much smaller" [22 p886]. The substantial loss of observations the authors lament could easily be overcome with MARBASECID and CPSID. Using MARBASECID to link the March Basic Monthly CPS and the ASEC and then CPSID to link the ASEC to other monthly surveys, the authors could have combined earnings from the ASEC with union membership information from April, May, and June and retained a much larger number of observations (roughly four times as many).

Section 5 – Conclusion

With support from the National Institute for Child Health and Human Development, we are developing integrated data, dissemination software, and associated metadata that will make combining information from the ASEC and other CPS Basic Monthly files dramatically easier. The creation of MARBASECID, which unlocks the vast research potential of longitudinal CPS data by facilitating the inclusion of the ASEC, promises to serve the scientific community. MARBASECID and CPSID will both be freely available exclusively via IPUMS CPS and will be updated as new data become available; these data are fully documented and easily accessible for researchers around the world. These investments in data infrastructure eliminate the need for each individual researcher to perform the tedious task of linking Basic Monthly CPS data to ASEC data, reduce technical errors in linking, simplify replication of existing studies, and encourage researchers to rethink the possibilities of CPS data.

Linking the ASEC, with its oversamples, to the CPS monthly data is dauntingly complex. But the linked data has amazing potential for social science, economic, and health research. These barriers to use are real as evidenced by the limited research linking ASEC to monthly CPS data and previous work documenting how to link ASEC observations one year apart [8]. The opportunities are also rich as indicated by recent work that uses ASEC data with other monthly data through imputation as opposed to direct linkages. The availability of the ASEC as a part of a panel of linked CPS data dramatically magnifies the utility of CPSID— the variable that links CPS monthly observations across time—given the widespread use of the ASEC.

Acknowledgments

References

1. U.S. Bureau of Labor Statistics. Design and Methodology: Current Population Survey. Washington, D.C.: U.S. Department of Labor, Bureau of the Census; 2006. Technical Paper 66

2. Drew JA, Flood S, Warren JR. Making full use of the longitudinal design of the Current Population Survey: Methods for linking records across 16 months. Journal of Economic and Social Measurement. 2014; 39(3):121. [PubMed: 26113770]

3. U.S. Census Bureau. Current Population Survey, 2010 ASEC. Washington, D.C.: U.S. Department of Labor, Bureau of the Census; 2011. Technical Paper

4. Thompson M, Shapiro G. The Current Population Survey: An Overview. Annals of Economic and Social Measurement. 1973 Apr; 2(2):105–129.

5. Wolter KM, Polivka AE, Lubich A. Evolution of the Current Population Survey. Wiley StatsRef: Statistics Reference Online. 2015

6. Frankel LR, Stock JS. On the sample survey of unemployment. Journal of the American Statistical Association. 1942; 37:77–80.

7. U.S. Bureau of Labor Statistics. Design and Methodology: Current Population Survey. Washington, D.C.: U.S. Department of Labor, Bureau of the Census; 2002. Technical Paper 63RV

8. Madrian BC, Lefgren LJ. An approach to longitudinally matching Current Population Survey (CPS) respondents. Journal of Economic and Social Measurement. 2000 Jan 1; 26(1):31–62.

9. Feng S. The longitudinal matching of Current Population Surveys: A proposed algorithm. Journal of Economic and Social Measurement. 2001 Jan 1; 27(1–2):71–91.

10. Feng S. Longitudinal matching of recent Current Population Surveys: Methods, non-matches and mismatches. Journal of Economic and Social Measurement. 2008 Jan 1; 33(4):241–52.

11. Katz A, Tenter K, Sidel P. Comparison of alternative ways of deriving panel data from the annual demographic files of the Current Population Survey. Review of Public Data Use. 1984 Mar 1; 12(1):35–44.

12. Pitts, A. Matching adjacent years of the Current Population Survey [unpublished manuscript]. Los Angeles, CA: Unicon Research Corporation; 1988.

13. Bureau of Labor Statistics. Using the Current Population Survey as a Longitudinal Data Base. US: 1980. Report 608

14. Allen JT. A guide to the 1960–1971 Current Population Survey Files. The Annals of Economic and Social Measurement. 1973 Apr; 2(2):187–197.

15. U.S. Census Bureau. Annual Demographic File (March Supplement of Current Population Survey). Washington, D.C.: U.S. Department of Labor, Bureau of the Census; 1978. Technical Documentation

16. Bureau of Labor Statistics. Redesign of the sample for the Current Population Survey: Current Population Survey. Washington, D.C.: Bureau of the Census; 2014. Technical Documentation

17. Semega, JL., Welniak, E, Jr. The effects of the changes to the Current Population Survey Annual Social and Economic Supplement on estimates of income; Paper presented at: Allied Social Science Association (ASSA) Research Conference; 2015 Jan 3–5; Boston, MA, USA.

18. U.S. Census Bureau. American Time Use Survey User's Guide: Understanding ATUS 2003 to 2013. Washington, D.C.: Bureau of the Census; 2014. Technical Paper

19. Appendix for March 1994. Washington, D.C.: U.S. Census Bureau; 1996. Available from http://www.nber.org/morg/docs/usernote.asc

20. Gundersen C, Ziliak JP. Childhood food insecurity in the US: Trends, causes, and policy options. The Future of Children. 2014; 24(2):1–9.

21. Wight V, Kaushal N, Waldfogel J, Garfinkel I. Understanding the link between poverty and food insecurity among children: Does the definition of poverty matter? Journal of Children and Poverty. 2014 Jan 2; 20(1):1–20. [PubMed: 25045244]

22. Brady D, Baker RS, Finnigan R. When Unionization Disappears State-Level Unionization and Working Poverty in the United States. American Sociological Review. 2013 Oct 1; 78(5):872–96.

23. U.S. Census Bureau. Analysis of Perturbed and Unperturbed Age Estimates. 2008. Available from https://www.census.gov/cps/user_note_age_estimates.html

Appendix A

Match Validity

Match rates based on the algorithms we use are extremely high except in 1995 and 1996. We validate matches from 1994 forward by comparing AGE, SEX, and RACE in the March Basic Monthly CPS to the ASEC. As is evident in Appendix C, nearly all matches are validated. We document the problems we encounter in years where we fail to match or match incorrectly. Due to duplicate IDs in the 1989–1993 period, we are required to use AGE, SEX, and RACE as part of our linking algorithm and thus cannot validate on AGE, SEX, and RACE (but will validate on these variables by construction).

As evident in Table 3, the match rates based on the matching algorithms results in extremely high match rates, with the exception of 1995 and 1996. The validity of these matches is confirmed with a rather higher success rate based on age sex and race. This section provides potential explanations for invalid matches.

1994

115 observations fail to match on sex, 95 on race and 212 on age. We found no documentation to explain any of these failures. Our analysis shows that no observations fail to match on all three variables while only 2 observations fail to match on both sex and race.

1995

951 records in the March Basic Monthly CPS cannot be linked to the ASEC. Despite trying to match using various algorithms, IPUMS-CPS was unable to find links for these records. Furthermore, no Census Bureau documentation is available on this issue. It is possible that these non-links are a result of the CPS redesign that occurred in 1994.

1996

3 person records cannot be matched. No explanation has been uncovered.

2001

Several linked records do not match on age, sex, or race. No Census Bureau documentation on this issue has been located. However, 2001 was a CPS redesign year, which may be an explanation.

2002

The 128 age non-matches are most likely due to age perturbation. In August 2002 "depending on the demographic characteristics of all members of the household, ages of selected household members were adjusted to increase confidentiality protection" [23]. Since the ASEC is typically released in September, it is plausible and likely that these age non-matches are due to this perturbation issue.

2003

The 3,957 cases that do not match on age are a coding issue. That is, the ASEC topcodes at 85 while the Basic topcodes at 80. Thus, in validating the matches, persons ages 81–85 in the ASEC will be assigned the value of their actual age while their age in the Basic will be topcoded.

2004

The 1,832 age non-matches are also top code issues. Similarly, the 154 non-match cases on race are also a coding issue. The Basic Monthly CPS codes "3 or more races" while the ASEC actually lists out the three races.

2009

The 6 observations that do not match on age are most likely an age perturbation issue.

Appendix B

Details of Merging

The merging of the 1994 March Basic Monthly CPS to ASEC file requires three manual corrections in order to match correctly. These changes are apparent upon visual inspection. First, a correction is needed for the HUHHNUM of one three households. In 1994, a unique household in the Basic can be identified by the HRHHID and HRHHID2 while in the ASEC it can be identified by its HSEQ number (note that all variables here refer to IPUMS variable names). As is apparent in the figure below, there are three distinct households with the same HRHHID. The problem when it comes to linking is that only HUHHNUM is common to both files and all are equal to 1 (resulting in duplicate records based on HRHHID and HUHHNUM). Visual inspection makes it clear that the household from the Basic with HRHHID2 63011 is identical to the ASEC household with HSEQ 12307. Similarly, HRHHID2 63001 is identical to HSEQ 12306. For matching across the Basic Monthly CPS and ASEC, we re-assign the HUHHNUM values for these households ("New HUHHNUM" 2 and 3, respectively)

	HRHHID	HRHHID2	HUHHNUM	New HUHHNUM	YEAR	MIS	NUMPREC	LINENO	AGE	SEX	RACE
Basic	880669103209	63021	1	1	1994	2	4	1	34	2	2
Basic	880669103209	63021	1	1	1994	2	4	2	15	1	2
Basic	880669103209	63021	1	1	1994	2	4	3	11	1	2
Basic	880669103209	63021	1	1	1994	2	4	4	10	1	2
Basic	880669103209	63011	1	2	1994	2	1	1	72	2	2
Basic	880669103209	63001	1	3	1994	2	2	1	34	2	1
Basic	880669103209	63001	1	3	1994	2	2	2	21	1	2
	HRHHID	HSEQ	HUHHNUM	New HUHHNUM	YEAR	MIS	NUMPREC	LINENO	AGE	SEX	RACE
ASEC	880669103209	12308	1	1	1994	2	4	1	34	2	2
ASEC	880669103209	12308	1	1	1994	2	4	2	15	1	2
ASEC	880669103209	12308	1	1	1994	2	4	3	11	1	2
ASEC	880669103209	12308	1	1	1994	2	4	4	10	1	2

ASEC	880669103209	12307	1		2	1994	2	1	1	72	2	2
ASEC	880669103209	12306	1		3	1994	2	2	1	33	2	1
ASEC	880669103209	12306	1		3	1994	2	2	2	33	1	2

Another correction that must be made involves the following case. Looking closely at the data, it is evident that the Basic observation with HRHHID2 63001 should be matched to the ASEC observation with HSEQ 258.

	HRHHID	HRHHID2	HUHHNUM	YEAR	MIS	NUMPREC	LINENO	AGE	SEX	EDUC	RACE	OCC
Basic	160999430499	63021	1	1994	1	1	1	48	2	40	1	−1
Basic	160999430499	63001	1	1994	1	1	1	48	2	40	1	20
	HRHHID	HSEQ	HUHHNUM	YEAR	MIS	NUMPREC	LINENO	AGE	SEX	EDUC	RACE	OCC
ASEC	160999430499	258	1	1994	1	1	1	48	2	40	1	20
ASEC	160999430499	260	1	1994	1	1	1	48	2	40	1	0

A more complicated duplicate is below in which the records are identical save for HRHHID2. Leveraging the longitudinal component of the survey to figure out the correct match, we locate the observations in 1995 March Basic Monthly CPS and the 1995 ASEC when both should have been in MIS 6 according to the CPS rotation pattern. Only HRHHID2 63001 and HSEQ 13308 is in the 1995 March Basic Monthly CPS and ASEC files, so we match and retain the records which also appear in 1995.

	HRHHID	HRHHID2	HUHHNUM	YEAR	MIS	NUMPREC	LINENO	AGE	SEX	EDUC	RACE	OCC	MARST
Basic	930479150329	63001	1	1994	2	1	1	71	1	32	1	0	3
Basic	930479150329	63011	1	1994	2	1	1	71	1	32	1	0	3
ASEC	930479150329	13308	1	1994	2	1	1	71	1	32	1	0	3
ASEC	930479150329	13309	1	1994	2	1	1	71	1	32	1	0	3

In 2003, we identify two problematic cases in the March Basic Monthly CPS. Consider the two 68 year olds of the same sex, race, and education level (below). Using HRHHID2 from the March Basic Monthly CPS, we see that the person who is age 48 is in the same household as the first 68 year old. Thus, we match the first 68 year old in the ASEC (HSEQ=62477) to the 68 year old in HRHHID2=76261 in the Basic.

	HRHHID	HRHHID2	HUHHNUM	YEAR	MIS	NUMPREC	PERNUM	AGE	SEX	EDUC	RACE
Basic	130962064655659	76261	1	2003	4	2	1	48	1	40	1
Basic	130962064655659	76001	1	2003	4	2	1	78	1	36	1
Basic	130962064655659	76261	1	2003	4	2	2	68	2	39	1
Basic	130962064655659	76001	1	2003	4	2	2	68	2	39	1
	HRHHID	HSEQ	HUHHNUM	YEAR	MIS	NUMPREC	PERNUM	AGE	SEX	EDUC	RACE
ASEC	130962064655659	62476	1	2003	4	2	1	78	1	36	1

| ASEC	130962064655659	62477	1	2003	4	2	1	48	1	40	1
ASEC	130962064655659	62476	1	2003	4	2	2	68	2	39	1
ASEC	130962064655659	62477	1	2003	4	2	2	68	2	39	1

The next set of observations that is troublesome contains two people whose age does not match between the Basic Monthly CPS and the ASEC. In order to match, we change the age for one of the observations.

Note that original values are maintained in the original files.

	HRHHID	HRHHID2	YEAR	MIS	NUMPREC	LINENO	AGE	SEX	EDUC	RACE
Basic	67843683692593	76261	2003	3	2	2	66	2	34	1
	HRHHID	HUHHNUM	YEAR	MIS	NUMPREC	LINENO	AGE	SEX	EDUC	RACE
ASEC	67843683692593	1	2003	3	2	2	67	2	34	1

In 2004, we need only adjust the age of the following observation. Again, the original age values are retained in the original files.

	HRHHID	HRHHID2	YEAR	MIS	NUMPREC	LINENO	AGE	SEX	EDUC	RACE
Basic	263943067909060	76261	2004	6	1	1	80	1	34	1
	HRHHID	HUHHNUM	YEAR	MIS	NUMPREC	LINENO	AGE	SEX	EDUC	RACE
ASEC	263943067909060	1	2004	6	1	1	85	1	34	1
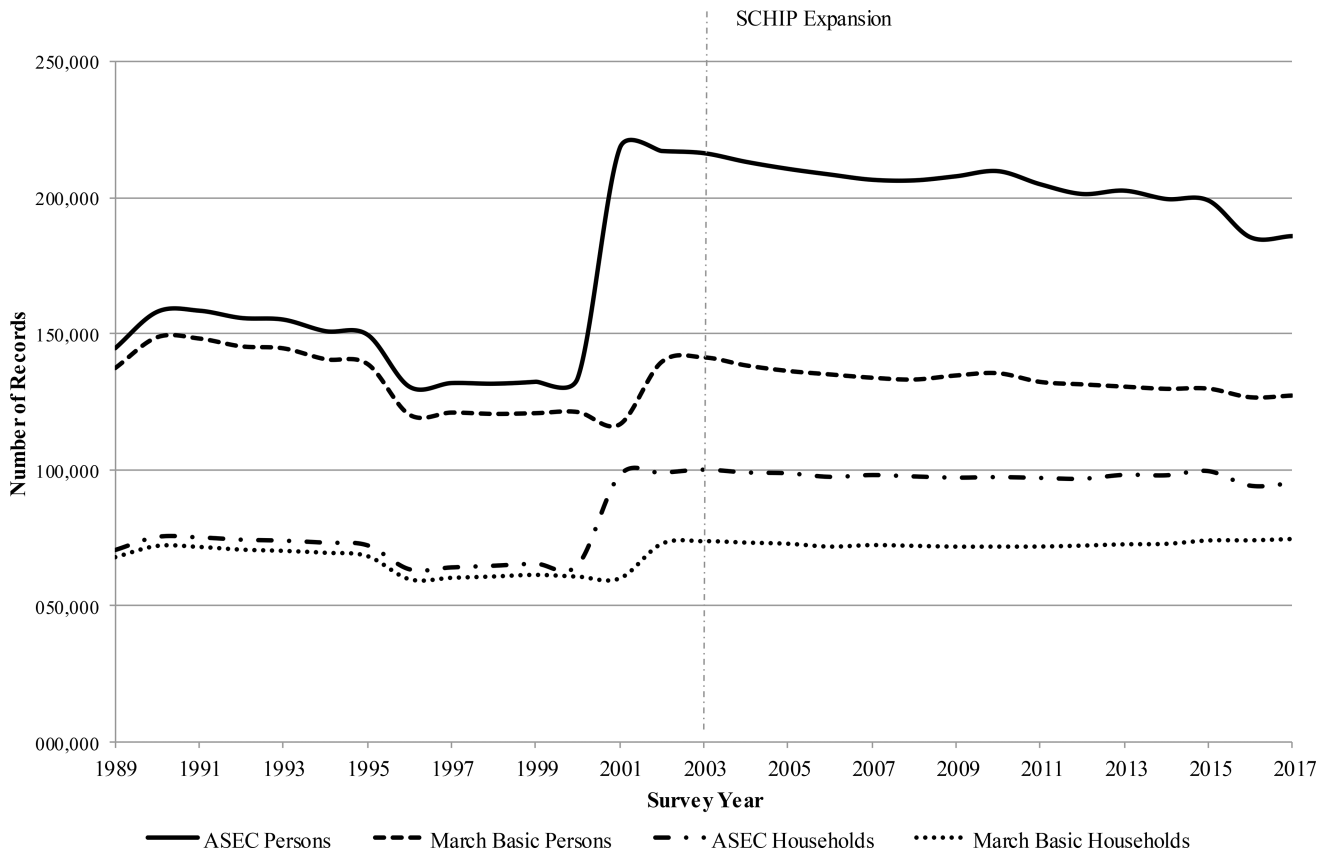
Figure 1.
Comparing the Size of the March Basic Monthly Survey and the ASEC Oversample, 1989–2017

CPS Rotation Pattern

Table 1

| | Year X | | | | | | | | | | | | Year X+1 | | | |
	Jan	**Feb**	**Mar**	**Apr**	**May**	**Jun**	**Jul**	**Aug**	**Sep**	**Oct**	**Nov**	**Dec**	**Jan**	**Feb**	**Mar**	**Apr**
MiS 1	Jan_X	Feb_X	Mar_X	Apr_X	May_X	Jun_X	Jul_X	Aug_X	Sep_X	Oct_X	Nov_X	Dec_X	Jan_{X+1}	Feb_{X+1}	Mar_{X+1}	Apr_{X+1}
MiS 2	Dec_{X-1}	Jan_X	Feb_X	Mar_X	Apr_X	May_X	Jun_X	Jul_X	Aug_X	Sep_X	Oct_X	Nov_X	Dec_X	Jan_{X+1}	Feb_{X+1}	Mar_{X+1}
MiS 3	Nov_{X-1}	Dec_{X-1}	Jan_X	Feb_X	Mar_X	Apr_X	May_X	Jun_X	Jul_X	Aug_X	Sep_X	Oct_X	Nov_X	Dec_X	Jan_{X+1}	Feb_{X+1}
MiS 4	Oct_{X-1}	Nov_{X-1}	Dec_{X-1}	Jan_X	Feb_X	Mar_X	Apr_X	May_X	Jun_X	Jul_X	Aug_X	Sep_X	Oct_X	Nov_X	Dec_X	Jan_{X+1}
MiS 5	Jan_{X-1}	Feb_{X-1}	Mar_{X-1}	Apr_{X-1}	May_{X-1}	Jun_{X-1}	Jul_{X-1}	Aug_{X-1}	Sep_{X-1}	Oct_{X-1}	Nov_{X-1}	Dec_{X-1}	Jan_X	Feb_X	Mar_X	Apr_X
MiS 6	Dec_{X-2}	Jan_{X-1}	Feb_{X-1}	Mar_{X-1}	Apr_{X-1}	May_{X-1}	Jun_{X-1}	Jul_{X-1}	Aug_{X-1}	Sep_{X-1}	Oct_{X-1}	Nov_{X-1}	Dec_{X-1}	Jan_X	Feb_X	Mar_X
MiS 7	Nov_{X-2}	Dec_{X-2}	Jan_{X-1}	Feb_{X-1}	Mar_{X-1}	Apr_{X-1}	May_{X-1}	Jun_{X-1}	Jul_{X-1}	Aug_{X-1}	Sep_{X-1}	Oct_{X-1}	Nov_{X-1}	Dec_{X-1}	Jan_X	Feb_X
MiS 8	Oct_{X-2}	Nov_{X-2}	Dec_{X-2}	Jan_{X-1}	Feb_{X-1}	Mar_{X-1}	Apr_{X-1}	May_{X-1}	Jun_{X-1}	Jul_{X-1}	Aug_{X-1}	Sep_{X-1}	Oct_{X-1}	Nov_{X-1}	Dec_{X-1}	Jan_X

Note: Table reports the month and year in which respondents began the CPS, separately by calendar month and survey month-in-sample. For example, "Oct_{X-2}" in the bottom left cell means that respondents in month-in-sample 8 in January of Year X first entered the CPS in October of Year X−2.

Table 2

Administering the ASEC to Oversamples Over Time.

ASEC Sample Includes:	Month ASEC Administered	1947–1955	1956–1975	1976–2000	2001–2003	2004-present
April Basic	April	x	--	--	--	--
March Basic	March	--	x	x	x	x
"Hispanic Oversample" November						
MIS 1	February[c]	--	--	x	x	x
MIS 2	March[c]	--	--	x	x	x
MIS 3	March[c]	--	--	x	x	x
MIS 4	March[c]	--	--	x	x	x
MIS 5	February[c]	--	--	x	x	x
MIS 6	March[c]	--	--	x	x	x
MIS 7	March[c]	--	--	x	x	x
MIS 8	March[c]	--	--	x	x	x
Non-Hispanic [a]						
November (MIS 6,7,8)	March[d]	--	--	--	x	--
August (MIS 8)	February[d]	--	--	--	--	x
September (MIS 8)	February[d]	--	--	--	--	x
October (MIS 8)	April[d]	--	--	--	--	x
February (MIS 4, 8)[b]	February[e]	--	--	--	x	x
April (MIS 1, 5)	April[e]	--	--	--	x	x

[a] Includes non-Hispanic non-Whites and non-Hispanic Whites with children 18 years or less

[b] These cases are identified in November as MIS 1,5

[c] These dwellings are interviewed a 9th and 10th time which can be considered MIS 9 and MIS 10

[d] These dwellings are interviewed a 9th time which can be considered MIS 9

[e] These dwellings are part of the split-path supplement assignment.

Table 3

ASEC–March Basic Matches & Match Validation by Year, 1989–2017.

Linking Period

1989–1993

First Stage Merge Variables: IPUMS (HRHHID, LINENO, AGE, SEX, RACE); Census (H-IDNUM, A-LINENO, A-AGE, A-SEX, A-RACE)

Additional Variables for Second Stage Merge: N/A

Year	Persons	1st Stage Matches	2nd Stage Matches	Non-Matches	Match Rate (Matches/Persons)	Unvalidated Matches		
						Sex	Race	Age
1989	137,384	137,384	N/A	0	100.00	0	0	0
1990	148,730	148,730	N/A	0	100.00	0	0	0
1991	148,228	148,228	N/A	0	100.00	0	0	0
1992	145,355	145,355	N/A	0	100.00	0	0	0
1993	144,618	144,618	N/A	0	100.00	0	0	0

1994–1995[a]

First Stage Merge Variables: IPUMS (HRHHID, HUHHNUM, STATECENSUS, LINENO); Census (H-IDNUM, H-HHNUM, GESTCEN, A-LINENO)

Additional Variables for Second Stage Merge: IPUMS (AGE, SEX, RACE, NUMPREC); Census (PRTAGE, PESEX, PTDTRACE, HRNUMHOU)

Year	Persons	1st Stage Matches	2nd Stage Matches	Non-Matches	Match Rate (Matches/Persons)	Unvalidated Matches		
						Sex	Race	Age
1994	140,625	140,079	546	0	100.00	115	95	212
1995[b]	138,872	137,473	448	951	99.32	0	0	0

1996–2000[a]

First Stage Merge Variables: IPUMS (HRHHID, HUHHNUM, STATECENSUS, LINENO); Census (H-IDNUM, H-HHNUM, GESTCEN, A-LINENO)

Additional Variables for Second Stage Merge: IPUMS (AGE, SEX, RACE, NUMPREC); Census (PRTAGE, PESEX, PTDTRACE, HRNUMHOU)

Year	Persons	1st Stage Matches	2nd Stage Matches	Non-Matches	Match Rate (Matches/Persons)	Unvalidated Matches		
						Sex	Race	Age
1996[b]	120,186	120,179	4	3	99.9975	3	3	3
1997	120,989	120,981	8	0	100.00	0	0	0
1998	120,507	120,504	3	0	100.00	0	0	0
1999	120,776	120,760	16	0	100.00	0	0	0
2000	121,194	121,149	45	0	100.00	0	0	0

Author Manuscript   Author Manuscript   Author Manuscript   Author Manuscript

2001–2004[a]

First Stage Merge Variables: IPUMS (HRHHID, HUHHNUM, STATECENSUS, LINENO); Census (H-IDNUM, H-HHNUM, GESTCEN, A-LINENO)

Additional Variables for Second Stage Merge: IPUMS (AGE, SEX, RACE, NUMPREC, OCC); Census (PRTAGE, PESEX, PTDTRACE, HRNUMHOU, PEIO1OCD)

Year	Persons	1st Stage Matches	2nd Stage Matches	Non-Matches	Match Rate (Matches/Persons)	Unvalidated Matches		
						Sex	Race	Age
2001[b]	116,663	116,585	78	0	100.00	10	213	573
2002[b]	139,660	139,592	68	0	100.00	0	0	128
2003[b]	141,288	141,220	68	0	100.00	0	0	3957
2004[b]	138,350	138,277	73	0	100.00	0	154	1832

2005–2017

First Stage Merge Variables: IPUMS (HRHHID, HRHHID2, LINENO); Census (H-IDNUM, H-IDNUM2, A-LINENO)

Additional Variables for Second Stage Merge: N/A

Year	Persons	1st Stage Matches	2nd Stage Matches	Non-Matches	Match Rate (Matches/Persons)	Unvalidated Matches		
						Sex	Race	Age
2005	136,315	136,315	N/A	0	100.00	0	0	0
2006	135,028	135,028	N/A	0	100.00	0	0	0
2007	133,817	133,817	N/A	0	100.00	0	0	0
2008	133,155	133,155	N/A	0	100.00	0	0	0
2009[b]	134,650	134,650	N/A	0	100.00	0	0	6
2010	135,478	135,478	N/A	0	100.00	0	0	0
2011	132,275	132,275	N/A	0	100.00	0	0	0
2012	131,372	131,372	N/A	0	100.00	0	0	0
2013	130,534	130,534	N/A	0	100.00	0	0	0
2014	129,727	129,727	N/A	0	100.00	0	0	0
2015	129,811	129,811	N/A	0	100.00	0	0	0
2016[c]	126,628	117,990	N/A	8,638	93.18	0	0	0
2017[c]	127,265	118,650	N/A	8,615	93.23	0	0	0

[a]See Appendix for details on second-stage merging.

[b]See Appendix for notes about non-matches.

[c]CPS-ASEC 2016 and 2017 included a split-panel sample resulting in March Basic Monthly households not being included in the ASEC.

Table 4

Linking the Food Security Supplement (December) to the Child Tax Credit (ASEC), 2005–2008.

Panel A: Linking December to ASEC

	(1)	December Basic		Linked to March Basic[1]			Linked to ASEC who have FSS and CTC[2]		
		(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	# of Persons	# Eligible to Link to March	% Eligible to Link to March (Col 2/Col 1)		% of Persons in December Basic (Col 4/Col 1)	% of Persons Linked of Those Eligible (Col 4/Col 2)	# of Linked Persons	% Linked of December Basic (Col 7/Col 1)	% Linked of Eligible (Col 7/Col 2)
December 2005	138,903	34,024	24.49%	29,807	21.46%	87.61%	21,091	15.18%	61.99%
December 2006	136,174	32,851	24.12%	29,407	21.60%	89.52%	20,324	14.93%	61.87%
December 2007	135,275	32,798	24.25%	29,757	22.00%	90.73%	20,455	15.12%	62.37%
December 2008	133,672	32,606	24.39%	28,873	21.60%	88.55%	19,975	14.94%	61.26%

Panel B: Linking Across Years

	(10)	December to ASEC in year x and December to ASEC in year x+1		(13)	(14)
		(11)	(12)		
	# of Persons in December of year x	# Linked from December of year x to March of year x+2	% Linked from December of year x to March of year x+2 (Col 11/Col 10)	December MIS 1	Linked Obs / December MIS 1 Obs (Col 11/Col 13)
December 2005	138,903	7,213	5.19%	10,755	67.07%
December 2006	136,174	7,143	5.25%	10,398	68.70%
December 2007	135,275	7,081	5.23%	10,350	68.42%
December 2008	133,672	7,068	5.29%	9,948	71.05%

[1] The corresponding linkage between the March Basic and ASEC is perfect. That is, all respondents from the March Basic are identified in the ASEC.

[2] These numbers are conditional on having food security responses in the December Basic Monthly CPS and the Child Tax Credit question in the ASEC.

Table 5

Linking the ASEC to ORG observations from April, May, and June, 2005 – 2009.

March MIS	Month of MIS 4 or 8	(1) Number of Cases in March Basic	Linking Within a Year					Linking Across Years	
			(2) Number Linked to MIS 4 and 8	(3) Percentage Linked to March (Col 2/Col 3)	(4) Number Linked Cases with Earnings Data	(5) Percentage Linked with Earnings Data (Col 4/Col 2)		(6) Number Linked Cases with Earnings Data	(7) Percentage Linked with Earnings Data (Col 6/Col 2)
2005									
MIS 4, 8		34,521	34,521	100.00%	26,503	76.77%		9,548	27.66%
MIS 3, 7	April	34,488	33,090	95.95%	25,430	73.74%		8,818	25.57%
MIS 2, 6	May	34,018	31,891	93.75%	24,640	72.43%		8,686	25.53%
MIS 1, 5	June	33,288	30,314	91.07%	23,469	70.50%		7,978	23.97%
	Total	136,315	129,816	95.23%	100,042	73.39%		35,030	25.70%
2006									
MIS 4, 8		33,875	33,875	100.00%	26,116	77.10%		9,931	29.32%
MIS 3, 7	April	33,957	32,760	96.47%	25,315	74.55%		9,546	28.11%
MIS 2, 6	May	33,749	31,458	93.21%	24,476	72.52%		9,161	27.14%
MIS 1, 5	June	33,447	30,229	90.38%	23,431	70.05%		8,740	26.13%
	Total	135,028	128,322	95.03%	99,338	73.57%		37,378	27.68%
2007									
MIS 4, 8		33,443	33,443	100.00%	25,940	77.56%		10,166	30.40%
MIS 3, 7	April	33,737	32,603	96.64%	25,214	74.74%		9,579	28.39%
MIS 2, 6	May	33,324	31,407	94.25%	24,337	73.03%		9,352	28.06%
MIS 1, 5	June	33,313	30,537	91.67%	23,720	71.20%		8,953	26.88%
	Total	133,817	127,990	95.65%	99,211	74.14%		38,050	28.43%
2008									
MIS 4, 8		33,520	33,520	100.00%	26,135	77.97%		10,299	30.72%
MIS 3, 7	April	33,556	32,446	96.69%	25,305	75.41%		9,897	29.49%
MIS 2, 6	May	32,806	30,732	93.68%	24,086	73.42%		9,372	28.57%

March MIS	Month of MIS 4 or 8	Linking Within a Year					Linking Across Years	
		(1) Number of Cases in March Basic	(2) Number Linked to MIS 4 and 8	(3) Percentage Linked to March (Col 2/Col 3)	(4) Number Linked Cases with Earnings Data	(5) Percentage Linked with Earnings Data (Col 4/Col 2)	(6) Number Linked Cases with Earnings Data	(7) Percentage Linked with Earnings Data (Col 6/Col 2)
MIS 1, 5	June	33,273	30,521	91.73%	23,863	71.72%	9,207	27.67%
Total		133,155	127,219	95.54%	99,389	74.64%	38,775	29.12%
2009								
MIS 4, 8		33,250	33,250	100.00%	26,049	78.34%	10,343	7.68%
MIS 3, 7	April	34,352	33,189	96.61%	25,887	75.36%	10,020	7.44%
MIS 2, 6	May	33,263	31,310	94.13%	24,591	73.93%	9,431	7.00%
MIS 1, 5	June	33,785	31,227	92.43%	24,384	72.17%	9,185	6.82%
Total		134,650	128,976	95.79%	100,911	74.94%	38,979	28.95%