IPUMS
Working Papers

Extending Current Population Survey Linkages: Obstacles and Solutions for Linking Monthly Data from 1976 to 1988

Sarah Flood†
University of Minnesota

Renae Rodgers
University of Minnesota

Jose Pacas
University of Minnesota

Devon Kristiansen
University of Minnesota

Ben Klaas
University of Minnesota

Extending Current Population Survey Linkages: Obstacles and Solutions for Linking Monthly Data from 1976 to 1988

Sarah Flood, Renae Rodgers, José Pacas, Devon Kristiansen, and Ben Klaas
University of Minnesota

Abstract:

The Current Population Survey (CPS) has been the nation's primary source of information about employment and unemployment for decades. The data are widely used by social scientists and policy makers to study labor force participation, poverty, and other high-priority topics. An underutilized feature of the CPS is its short-run panel component. This paper discusses the unique challenges encountered when linking basic monthly data in the 1976-1988 period as well as linking the March basic monthly data to the Annual Social and Economic Supplement (ASEC) in the 1976-1988 period. We describe strategies to address linking obstacles and document linkage rates.

Introduction

The Current Population Survey (CPS) is the primary source of information about employment and unemployment in the United States. It has been a key data resource for the social science research and policy making community for decades, providing monthly snapshots of the civilian labor force since 1976. The survey has a short-run panel component that is largely unknown and underutilized by the research community. CPS respondents participate in the survey eight times, answering surveys for four consecutive months, rotating out of the survey for eight months, then answering for four more consecutive months. Linked CPS data present numerous opportunities for charting and understanding dynamics of short-run change over the last half century. Examples of research that could be conducted using linked CPS data include analyzing the demographic and employment correlates of families transitioning in and out of poverty; examining the extent to which engagement in volunteering changes following transitions out of employment among older adults; and studying how families' employment arrangements are stable or change following the birth of a child. Other investigations might focus on how individuals organize their work and family lives in response to recessions, the effects of policy changes on employment, or the labor force participation of veterans from different wars.

IPUMS (ipums.org) is a leading disseminator of CPS data, streamlining and simplifying access to these vital social, economic, and demographic data. IPUMS CPS (cps.ipums.org) delivers Annual Social and Economic Supplement (ASEC) data from 1962 to present and basic monthly survey (BMS) data from 1976 to present along with nearly all topical supplement data. IPUMS codes variables consistently across time, provides access to original unrecoded versions of the variables, documents all variables for ease of use, and enables users to create customized

datasets that include only variables from the months and years of data they want to analyze thereby dramatically reducing redundant effort across researchers.

Part of the IPUMS effort entails informing the research community about opportunities for linking the CPS over time and enabling analyses using the CPS panel component. Previously we have documented linkages across months from 1989 forward (Drew, Flood, & Warren, 2014) and between the March BMS and the ASEC (Flood & Pacas, 2017). These linkages rely on CPSIDP and MARBASECIDP, respectively, which are linking keys constructed by IPUMS to simplify use of the CPS panel component. While researchers could independently link CPS files over time, IPUMS-constructed linking keys and accompanying documentation provide the research community with a common starting point, save individual researchers hours of effort, and reduce potential for error in the linking process. In this paper, we document the specific challenges linking across months and between the ASEC and March BMS in the 1976 to 1988 period. We outline the solutions we implemented and provide linkage rates against which researchers can check their own linkages.

CPS Rotation Pattern and Linking

The CPS is a rotating panel household-level survey, which means that all individuals residing in a household are surveyed for four months, rotate out of the survey for eight months, and then rotate back into the survey for another four months (known as the 4-8-4 rotation pattern). The first time the household appears in the CPS is their first "month-in-sample" (MIS) and this is indicated in the data (MIS coded as 1). Each subsequent month that a household is included in the CPS, their MIS value increments by one. A household's MIS value does not increment in months when the household is out of rotation (the eight-month break). Thus, the final month of the 4-8-4 rotation pattern in the CPS has an MIS value of eight.

The Census Bureau provides a very brief set of guidelines for linking individual observations across CPS data files (U.S. Census Bureau, 2006). While they list the variables researchers should use to link observations over time, the documentation is sparse and provides insufficient detail about how to do the linking. Furthermore, the instructions are focused on linking within a set of years when linking keys are stable. Little attention is given to how to bridge changes in linking keys over time. Finally, the documentation (U.S. Census Bureau, 1994) also indicates that it is not possible to link some years of data together due to changes in the survey (i.e., 1976 to 1977 and 1985 to 1986).

CPSID(P)

To overcome barriers to linking CPS data over time related to the rotation pattern and limited guidance, IPUMS created CPSID, a linking key that handles changes in linking keys in the CPS to consistently identify households across time, and CPSIDP, a person-level linking key. CPSID(P)[1] is created by applying Census Bureau rules for linking across time within a strictly enforced set of eligible months as dictated by the CPS rotation pattern. CPSID(P) eliminates the need for every researcher to navigate the complex CPS rotation pattern, changes over time in linking keys, and additional challenges that arise. CPSID(P) provides a common starting point for the research community, which should increase reproducibility of research using linked CPS data.

The CPS is a *household* survey, which has implications for the creation of CPSID(P). In 1976 to 1988, a household is defined by its household identification number (HRHHID) and its household number (HUHHNUM)[2]. If one or more of the individuals in the household move out

---

[1] We use CPSID(P) to refer to both CPSID, the household-level identifier, and CPSIDP, the person-level identifier.
[2] A note on terminology: Prior to 1989, there are often variables without proper variable names in codebooks; rather, variables were referenced as "items" such as "Item 18A. Line no." For convenience, we refer to the

of the household between CPS interviews, no attempt is made to follow the individuals who have moved. All individuals living in the surveyed CPS household at the time of the interview are surveyed. If everyone in the household moves out, and new people move into the household, the new people are interviewed the next time the household is interviewed for the CPS. If an entire household is replaced in this way, the household number (HUHHNUM) value changes, and a new CPSID value is generated for the new household (see Drew et al., 2014). CPSIDP identifies the same individuals, based on linking keys, over time and assigns a unique CPSIDP value to every individual in the data. For example, a household with two individuals that was first observed in the CPS in December of 1981 may have the following CPSID value 19811203287400; the individuals living in the household would have CPSIDP values of 19811203287401 and 19811203287402, respectively. A third individual who joins the household in January of 1982 would be assigned a CPSIDP value of 19811203287403; the two returning household members would be assigned the CPSDIP values they were first assigned in December.[3]

We encountered several problems creating CPSID(P) in the 1976 to 1988 period. Broadly, these included issues with 1) inconsistent household identifiers, 2) duplicated person identifiers within households, and 3) problematic linkages between BMS and ASEC files. We describe these problems in detail along with solutions we employed to create CPSID(P) for monthly and ASEC data in the 1976 to 1988 period.

Creating CPSID(P) for Basic Monthly Files

---

variables that should uniquely identify households and persons within households in the original census CPS files as HRHHID and HUHHNUM and LINENO, respectively. These are IPUMS CPS variable names.

[3] This is described in detail elsewhere (Drew et al 2014).

An important distinction between the monthly CPS data files is whether or not they are accompanied by a topical supplement. Regardless of whether topical supplemental data are collected, a file containing only the basic monthly variables (hereafter referred to as the basic monthly-only file) is always released. In months where a topical supplement is also fielded, a second file is released after the release of the basic monthly file; the topical supplement variables are appended to the end of the records in the basic monthly-only file (the supplement-containing file). Contrary to expectations, in the 1976-1988 period, the basic monthly portions of basic monthly-only and supplement-containing files are not always identical. Prior to creating CPSID(P), we attempt to reconcile differences—duplicate records, unmatchable supplement records, and different numbers of records between files—in these two versions of the monthly data (see Appendix A).

After reconciling within-month differences between basic monthly-only and supplement-containing versions of the CPS data files, we often encounter challenges linking across months that extend beyond those documented for the 1989 forward period (Drew et al., 2014). The two major obstacles to linking monthly files that must be overcome are inconsistent household identifiers and duplicated person identifiers.

Problem 1: Household Identifiers and Linking across Months

We made modifications to the original household identifier to create CPSID(P) between 1976 and 1988. HHID refers to the original unmodified household identifier and HRHHID refers to the modified household identifier used for linking. Except in specified months, HHID and HRHHID are the same.

A. *Non-unique Household Identifiers*

In 1976 and 1977, the original household identifier (HHID) does not uniquely identify all households in basic monthly-only files. Through trial and error, we located a set of variables that together uniquely identify households in these months. We created HRHHID, which uniquely identifies households within the basic monthly-only files, as follows: we used the first nine digits of the twelve-digit HHID and replaced the last three digits of HHID with information from the third and sixth sets of columns marked as "blank" in the original CPS codebook (i.e., BLANK3 and BLANK6[4]).

In 1976 and 1977, HRHHID yields more plausible households in basic-monthly only files than the original household identifier (HHID). Households identified using HHID that had been extremely large and contained multiple household heads are separated into smaller households with no more than one head per household when households are identified using HRHHID. We gauge the quality of HRHHID in the basic monthly-only files by making comparisons with the supplement-containing files from the same month (where possible) in which HHID uniquely identifies households. In 1976, three-quarters of the person records matched across basic monthly-only and supplement-containing files using HRHHID from the basic monthly-only files and HHID in the supplement-containing files. For the remaining quarter of the records, BLANK3 and BLANK6 matched across files, but the first nine digits of HHID did not match. In these households, the first nine digits differ systematically between the basic monthly-only and supplement-containing files, so we reconciled these differences in HRHHID (see Appendix A). In 1977, neither HHID nor HRHHID yields successful matches between basic monthly-only and supplement-containing files. However, HRHHID in the basic monthly-only files results in each household containing only one household head, which is almost always consistent with the

---

[4] Blank3 is found in columns 25-27 in the original data. Only the first two columns, 25-26 are used to uniquely identify households; the third column is blank for all records. Blank 6 is found in column 107 in the original data.

organization of households in the supplement-containing files and represents an improvement over HHID.

B. *Non-numeric Household Identifiers*

In January 1976 to June 1985, some records contained non-numeric characters in the original household identifier. We made adjustments to the household identifier that allow us to treat HRHHID as a numeric variable in all months. In 1976 and 1977 basic monthly-only files, a few records in each monthly file, usually less than 10, had a '-' in the final position of the 12-digit HRHHID variable. These strings are introduced into HRHHID through the adjustments described in the previous section to ensure that households are uniquely identified within files. In 1978-1985, about 12,000 to 14,000 cases in each month contained a blank space or a '-' in the second position of HHID. The December 1978 file included a '/' in the third position of HHID for 48 records. We replaced blank spaces with a '9'; '-' characters with an '8'[5]; and '/' characters with a '4'[6]. This character swap preserves the uniqueness of the household identifier within each month of data.

C. *Unlinkable Months*

Census Bureau documentation indicates that it is not possible to link some months of CPS data across years (1976 to 1977 and 1985 to 1986) due to changes in the survey (U.S. Census Bureau, 1994). We elaborate on these known limitations in the 1970s and remark briefly on our investigation in the 1980s.

---

[5] We used an inductive approach to replace spaces and '-' with numbers taking care to ensure that the numbers we selected would result in the same number of unique HHID values before and after the character swap. Converting HHID to be numeric does not change the number of linkages that result; numeric characters are simpler for researchers to use.

[6] We looked for the individuals in adjacent months with problematic HHIDs; based on our analysis, we replaced this character with a 4.

Basic monthly-only and supplement-containing files within month in 1977 cannot be linked and for the same reasons there are implications for linking across years. Starting with the supplement-containing months of 1977, additional households were included in the CPS, resulting in an additional 15,000 to 25,000 individuals in these months (known as the D-sample). This sample size increase was extended to include basic monthly-only months in 1978. To prevent identification of these additional individuals, the Census Bureau altered the procedures for generating original household identifiers so that supplement-containing months that included the D-sample could not be linked to basic monthly-only months without it (McIntire, 1980).

Despite HRHHID uniquely identifying households within the basic monthly-only files in 1976 to 1978, this identifier does not always yield expected linkages across months and (non)linkages are systematic. Within each year from 1976 to 1978, linkages across basic monthly-only and supplement-containing months are possible in 1976 and 1978 (see Figure 1a). Due to the D-sample, in 1977 months that contain supplements and those that are basic monthly-only cannot be linked. Figure 1b shows linkages between types of files across years. In 1976, both basic monthly-only and supplement-containing files can be linked to 1977 basic monthly-only files, but not to 1977 supplement-containing files. No linkages are made between 1976 and 1978 even when possible given the 4-8-4 rotation pattern (i.e., MIS 1 in October, November, and December 1976). In 1977, supplement-containing files can be linked to both 1978 basic monthly-only and supplement-containing files; no 1977 basic monthly-only file will link to either basic monthly-only or supplement-containing files in 1978.

Figure 1a. Linkages Possible between
File Types within Years, 1976-1978
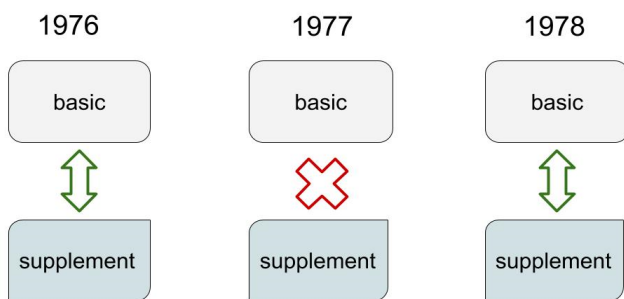
Figure 1b. Linkages Possible between
File Types across Years, 1976-1978
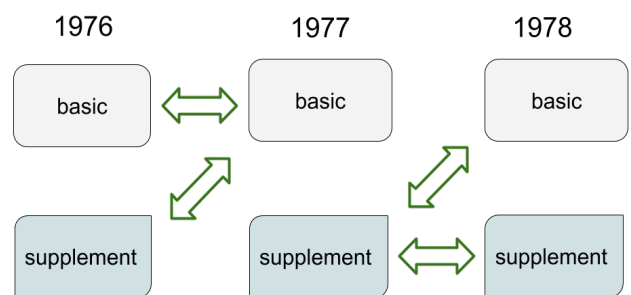
Census Bureau documentation says that 1985 and 1986 cannot be linked. We find that no

month before June of 1985 can be linked to any month after June of 1985. While we have found

no Census Bureau technical documentation on this issue, this linking barrier may be due to a

CPS redesign that began in April of 1984 and concluded by July of 1985 (Creighton and

Wilkinson, 1984).

Problem 2: Person Identifiers and Linking across Months

A. *Non-Unique Person Identifiers*

Person identifiers are required along with household identifiers to match individuals across

months of the CPS. While person identifiers should be unique within household in a given data

file, the same person identifier is sometimes assigned to multiple individuals in the same

household in months between 1976 and 1983. This presents a problem for linking individuals

across months using only person and household identifiers (LINENO, HRHHID, and

HUHHNUM, respectively).

If two or more records in a household have identical person identifiers, we do not allow

linkages across months for those individuals. Without using additional information about the

individuals, we cannot ascertain which record with the duplicated person identifier should be

linked to a single record with the same identifier in a surrounding month. When we encounter

pairs of records with duplicate LINENO values, we assign them both a new (and unique)

CPSIDP value. Consider the following household of individuals that appears in three months of

the CPS, beginning in December of 1981.

December 1981

HRHHID	HUHHNUM	MIS	LINENO	CPSIDP
003147831701	1	1	01	19811203287401
003147831701	1	1	01	19811203287402
003147831701	1	1	02	19811203287403

The first two records in the household in December of 1981 have the same identifiers

(HRHHID, HUHHNUM, and LINENO). Though the identifiers are the same, there is variation

between these records on the other variables, so the records are not complete duplicates. Since

the first two records have the same identifiers, we assign them unique CPSIDP values.

This household is also in the January 1982 CPS. In January, HRHHID and HUHHUM

values match and the LINENO values are 01 and 02, respectively. Without additional

information, we do not know which of the people in this household in December 1981 with

LINENO=01 is the same person in January of 1982 with LINENO=01. Therefore, when we see

the individual with LINENO=01 in January 1982, we assign them a new CPSIDP value

incrementing by one from the highest CPSIDP value ever assigned to a person in that household.

Note that LINENO=02 is assigned the same CPSIDP value in both months.

January 1982

HRHHID	HUHHNUM	MIS	LINENO	CPSIDP
003147831701	1	2	01	19811203287404
003147831701	1	2	02	19811203287403

As of January 1982, the household with an HRHHID value of 003147831701 and an

HUHHNUM value of one has four CPSIDP values. When we see the same household in

February of 1982 with LINENO values of 01 and 02, we assign the CPSIDP values to them that

they were assigned in January of 1982. This occurs because we create CPSIDP by looking to the

previous month for a match and continuing backward in time through all linkable months until a

match is found.

February 1982

HRHHID	HUHHNUM	MIS	LINENO	CPSIDP
003147831701	1	3	01	19811203287404
003147831701	1	3	02	19811203287403

Our approach is conservative and undoubtedly misses some plausible links across

months. However, CPSIDP is a mechanical match using *only* household and person identifiers to

assign CPSIDP values. Duplicate person line numbers are uncommon, representing less than 1%

of cases each month from January 1976 to December 1983.

Duplicate person line numbers are not an issue in the 1984 to 1988 period. The challenge

in this period is that the CPS basic monthly files include three person line number variables.

However, only one of the person line number variables (Item 18A in columns 541-542) uniquely

identifies persons within households.[7] We use this version for linking.

B. *Child Records*

Children under 14 are not included in basic monthly-only files until 1982. However, children are

included in the following supplement-containing months: October 1976, October 1977, and May

1978 through December 1981 (see Appendix A). Despite inclusion in the data file, we do not

link persons under 14 across months.

Basic Monthly Survey Linkage Rates Across Time

---

[7] For more information on these versions of person line number, see Appendix B.

Linkage rates across months in 1976 to 1988 using HRHHID, HUHHNUM, and LINENO as linking keys are lower than in more recent years (see Drew et al., 2014) due to comparatively poor data quality. We provide sample sizes and retention rates before and after validating links using AGE, SEX, and RACE for CPS data collected in 1976-1977 and 1987-1988 (see Drew et al., 2014 for linkages in 1994-1995 and 2009-2010). Table 1 shows the total number of records in each month-in-sample group from January 1976 to April 1977 and January 1987 to April 1988. Plausible links are those that have the same SEX and RACE in all time points and whose AGE does not increase by more than two years.

[table 1]

Based on the 4-8-4 rotation pattern of the CPS, 75% of respondents are eligible to link between consecutive months (MIS 1-3 and MIS 5-7 in a given month can link to the next month; MIS 4 and 8 rotate out of the survey). Table 2 shows linkages and retention rates between January and February in 1976 and in 1987. More than 90% of records in January eligible to link to February are observed in February and nearly all of them are plausible based on age, sex, and race.

[table 2]

Linkages are also possible across nonconsecutive months. Table 3 shows links two months apart from October to December in 1976 and in 1987. Half of individuals in October are eligible to participate in the CPS in December. Of those who are eligible, only 73% are linked in 1976 compared to 90% in 1987. Most of these linkages are plausible based on comparisons of age, sex, and race.

[table 3]

The CPS rotation pattern also allows for linking the same month across adjacent years (see Table 4). The individuals in MIS 1-4 in 1976 and 1987 are eligible to participate in the CPS in the same month in 1977 and 1988, respectively. About 75% of those eligible are observed the next year and just over 70% appear to be the same person based on age, sex, and race.

[table 4]

Respondents may be linked across up to eight months of participation in the CPS. Table 5 shows the number of people starting the CPS in January 1976 and 1987 who appear in the CPS up to eight times. The individuals who are first observed in 1976 link well in 1976, but do not link to January 1977, which is a supplement-containing month. As described above, 1976 files are not linkable to 1977 supplement-containing files. The individuals who begin the CPS in January 1987 are effectively linked in 1987, but linkage rates decrease after MIS 4 when they are observed again in January 1988 (see also Drew et al., 2014). About 56% of respondents who started the CPS in January 1987 are observed eight times.

[table 5]

Tables 6-8 show linkage rates for a variety of different linkage scenarios. Table 6 shows those people who entered the CPS in January of either 1976 or 1987 (MIS 1) and appear any of the subsequent seven times that their household could have been in the CPS after January (either MIS 2 or MIS 3 or MIS 4 or MIS 6 or MIS 7 or MIS 8). 95% of respondents appear in at least two months of the CPS. Table 7 shows those who appear in the CPS survey in month-in-sample 5-8. Finally, Table 8 shows attrition between MIS 4 and 5.

[table 6]

[table 7]

[table 8]

Extending CPSID(P) to the ASEC, 1976-1988

Adding CPSID(P) to the ASEC significantly lowers the barriers to linking the ASEC with other

months and topical CPS supplements. MARBASECIDH and MARBASECIDP, hereafter

MARBASECID(P), are IPUMS CPS variables that link the March basic monthly data to ASEC

data from the same year. For example, by adding CPSID(P) to the ASEC researchers can easily

link poverty information from the ASEC to information from the food security supplement. The

details of generating MARBASECID(P) for the 1989 forward period are described in detail

elsewhere (Flood & Pacas, 2017). While MARBASECID(P) is constructed using Census Bureau

identifiers only (with a few exceptions) from 1989 through 2019, in the 1976 to 1988 period,

poor data quality required a different methodology for linking. This section describes the

problems encountered in attempting to link March basic monthly and ASEC files in this period,

details the solutions we implemented in linking March basic monthly to ASEC and generating

MARBASECID(P), and compares our methodology to other possible methods.

Problem 1: Missing March basic monthly records in the ASEC

Between 1976 and 1988, the ASEC file should include all March basic monthly households plus

a Hispanic oversample drawn from the previous November CPS. However, in all years during

this period, the March basic monthly file contains individuals who do not appear in the ASEC

file. Table 10 shows the total number of individuals in the March basic monthly survey for each

year between 1976 and 1988, the number that merge to the ASEC, and details about non-merged

individuals. For those unmerged, we differentiate between individuals whose HRHHID is or is

not in the ASEC. In the first instance, an HRHHID value is in both the ASEC and the March

BMS. Either the ASEC does not contain a matching record for the individual or we are unable to

confidently locate the matching record due to duplication or mismatch of variables (described in

the next section). The number of basic monthly records for which the HRHHID does not appear

in the ASEC totals less than 100 in 1976 to 1985 and is in the thousands in 1986 to 1988. In the

1986-1988 period, the Census Bureau scrambled household identifiers in the ASEC file for

privacy reasons,[8] and we do not attempt to link the basic monthly and ASEC in these years.

Problem 2: Low Quality Linking Keys

We encounter difficulties in uniquely identifying records within files and in matching records

across files from 1976-1988. This is not the case in the 1989 forward period where HRHHID and

LINENO are sufficient to uniquely identify all or most records within both March BMS and

ASEC files and to link the March BMS and ASEC files (see Flood & Pacas, 2017).

A. Uniquely identifying records

        The March BMS files from 1982 and 1983 contain records that are entirely identical;

none of the ASEC files contain records that are complete duplicates. However, even though

records are unique in most of the files, we are often unable to find a single set of variables to

uniquely identify all records that appear in both the March BMS and the ASEC file for a given

year. Accordingly, we are both unable to apply the 1989-onward method of linking March BMS

and ASEC data and unable to identify a single set of linking keys that allow us to link all March

basic monthly records to their ASEC counterparts in the 1976-1988 period.

        Consider the following example household from the March BMS. The household

contains several individual records that have the same values for both LINENO and demographic

variables in the March BMS, though they are not completely identical records. Of the eight

records in the household, we only uniquely identify records with LINENO values of 1 and 2

while those with LINENO values of 3, 4, and 5 appear twice in the file. In this instance, we

---

[8] This is based on a conversation with staff at the U.S. Census Bureau, not official documentation.

attempt to link the first two persons in the household to the ASEC file; we cannot know which of

the three pairs of records with identical identifiers and demographic variables should be linked to

the corresponding ASEC file. Because the records in the ASEC have unique linking keys, we

link LINENO values of 1 and 2 across the March basic monthly and ASEC.

Household from 1982 March BMS

HRHHID	HUHHNUM	LINENO	AGE	SEX	RACE	HOURS
509037594903	1	1	36	1	1	50
509037594903	1	2	37	2	1	--
509037594903	1	3	7	2	1	BB
509037594903	1	3	7	2	1	BB
509037594903	1	4	6	1	1	BB
509037594903	1	4	6	1	1	BB
509037594903	1	5	2	2	1	BB
509037594903	1	5	2	2	1	BB

Household from 1982 ASEC

HRHHID	HUHHNUM	LINENO	AGE	SEX	RACE	HOURS
509037594903	1	1	36	1	1	50
509037594903	1	2	37	2	1	--
509037594903	1	3	7	2	1	--
509037594903	1	4	6	1	1	--
509037594903	1	5	2	2	1	--

In other cases, records do not match on identifiers and the demographic variables are

insufficient for uniquely identifying individuals when the person identifier is omitted. In a 1985

ASEC household, there are duplicate line numbers (multiple 3s and 4s). We will not look for a

matching record in the ASEC for the two BMS records with the age of 12 (LINENO 5 and 6 in

the BMS and LINENO 3 and 4 in the ASEC) and the two records with the age of 11 (LINENO 7

and 8 in the BMS and the third LINENO 4 and 0 in the ASEC) because LINENO does not match

across files and demographic characteristics are not unique within BMS and ASEC files.

Household from 1985 March BMS

HRHHID	HUHHNUM	LINENO	AGE	SEX	RACE	HOURS
202962182216	1	1	40	1	1	60
202962182216	1	2	41	2	1	40
202962182216	1	5	12	1	1	BB
202962182216	1	3	18	2	1	--
202962182216	1	7	11	2	1	BB
202962182216	1	6	12	1	1	BB
202962182216	1	4	16	2	1	--
202962182216	1	8	11	2	1	BB

Household from 1985 ASEC

HRHHID	HUHHNUM	LINENO	AGE	SEX	RACE	HOURS
202962182216	1	1	40	1	1	60
202962182216	1	2	41	2	1	40
202962182216	1	3	18	2	1	--
202962182216	1	4	16	2	1	--
202962182216	1	3	12	1	1	--
202962182216	1	4	12	1	1	--
202962182216	1	4	11	2	1	--
202962182216	1	0	11	2	1	--

B. Unmatchable records

Among records that can be uniquely identified, some still cannot be matched across files in

1976-1988. For example, even when records are uniquely identified within a household,

LINENO values can differ across March BMS and ASEC files, which undermines the utility of

LINENO as a linking key.[9] In some instances, individuals within the household appear to be

sorted differently. For example, in this household from 1976, the position of the 18-year old and

the first 16-year old are switched between the March BMS and the ASEC file. Though the two

sixteen year-olds can be uniquely identified within their household and in each file, these

---

[9] See Appendix B for more information on versions of the person identifier in 1984-1988. Only one version of the line number variable exists in the ASEC files.

records—though almost certainly representing the same people in both files—will not be

matched across the March BMS and the ASEC.

Household from 1976 March BMS

HRHHID	HUHHNUM	LINENO	AGE	SEX	RACE	HOURS
046112033020	1	1	38	1	2	--
046112033020	1	2	37	2	2	40
046112033020	1	3	16	1	2	--
046112033020	1	4	18	2	2	--
046112033020	1	6	16	1	2	--

Household from 1976 ASEC

HRHHID	HUHHNUM	LINENO	AGE	SEX	RACE	HOURS
046112033020	1	1	38	1	2	--
046112033020	1	2	37	2	2	40
046112033020	1	3	18	2	2	--
046112033020	1	4	16	1	2	--
046112033020	1	5	16	1	2	--

Similarly, some individuals are uniquely identified by demographic variables and line

number and appear in the same order within the household, but have different person identifiers

in the March BMS and ASEC files. This is another instance where person line numbers, even

when unique within a household, prevent linking across March BMS and ASEC files. We

illustrate this situation in a household from 1976. The households contain the same number of

people with the same age, sex, race, and work hours, but have different LINENO values in the

March BMS and the ASEC. This data quality issue is not present in the data from 1989 forward,

but presents a major obstacle to linking March BMS and ASEC records from 1976-1988.

Household from 1976 March BMS

HRHHID	HUHHNUM	LINENO	AGE	SEX	RACE	HOURS
50007137050	1	1	29	1	1	40
50007137050	1	2	30	2	1	--
50007137050	1	7	46	2	1	--
50007137050	1	8	16	1	1	--
50007137050	1	9	15	1	1	--
50007137050	1	10	15	1	1	--

Household from 1976 ASEC

HRHHID	HUHHNUM	LINENO	AGE	SEX	RACE	HOURS
50007137050	1	1	29	1	1	40
50007137050	1	2	30	2	1	--
50007137050	1	5	46	2	1	--
50007137050	1	6	16	1	1	--
50007137050	1	7	15	1	1	--
50007137050	1	8	15	1	1	--

Problem 3: Unlinkable years

We were unable to generate MARBASECID(P) in 1977. There are no matching HRHHID values

across the March BMS and ASEC files.

Strategy for Linking March basic monthly to ASEC

Due to the problems detailed above, the methodology for merging March BMS and ASEC files

in 1976-1988 differs substantially from that used from 1989 forward (see Flood & Pacas, 2017).

We apply a multi-stage merging process to maximize linkage rates while minimizing spurious

matches. While demographic information is used to make matches between the March BMS and

the ASEC in a handful of cases in the 1989-2019 period, this information becomes indispensable

from 1976-1988. Such heavy reliance on demographic and other auxiliary information in linking

files represents a departure from our usual approach of generating mechanical links for

CPSID(P) and MARBASECID(P). However, the unique challenges of the data from this period

necessitate a different approach to attain reasonable linkage rates between the March BMS and

ASEC. We use person identifiers and demographic information as linking keys in some stages

and exclude them in other stages. We perform matching between the March BMS and ASEC

during four stages in 1976-1981, five stages in 1982-1987, and six stages in 1988 to address

problems detailed in the previous section.

Preliminary Steps

For an individual to be matched across BMS and ASEC files, the household identifier must

appear in both files. Before attempting to link, we exclude those individuals whose HRHHID

value appears in either the BMS or ASEC and not both files (see Table 10 below). Because there

are no persons under the age of 14 in the March BMS from 1976 to 1981, individuals under age

14 in the ASEC are also ineligible to link to the March BMS. In both cases, individuals excluded

from the matching are assigned non-linking MARBASECIDP values and, in the ASEC file,

CPSIDP values of 0.[10]

   We make adjustments to the coding of race and usual hours worked before attempting to

match individuals across March BMS and ASEC files. In 1988, race codes are different between

the March BMS and ASEC; we standardize race codes by recoding values of 3, 4 or 5 in the

ASEC ("American Indian or Aleut Eskimo", "Asian or Pacific Islander", and "Other") to 3,

which represents "Other" in the March BMS. In the "hours worked" variable (AHRSWORKT) in

all March BMS files from 1982-1988, there are two varieties of missing values, one for adults

who were not working last week and one for children. In the ASEC files from these years, all

NIU cases have a value of 0. We recode the two missing values in the March Basic Monthly files

to 0 before matching.

---

[10] As a result of this data quality issue, it will not be possible to distinguish ASEC oversample records from unlinked March basic monthly records in 1976-1988 and those under the age of 14 in 1976-1981.

Linking Methods

As an additional step before each linking stage, we set aside for later use all records that are not uniquely identified by that stage's linking keys (see Table 9). After linking, we combine the set-aside records with the unlinked records (for both the March BMS and ASEC); these files serve as the inputs for the next linking stage. We detail the number of linkages between the March BMS and the ASEC at each stage for every year from 1976 to 1988 in Table 10. The link stage is available to researchers via IPUMS CPS in the variable MARBASECSTAGE, which is available on both the March BMS and ASEC files from 1976 to 1988.

All BMS files in the 1976-1988 period contain multiple versions of the person identifier and age, sex, and race variables. These versions differ slightly from one another (see Appendix B) and we use the version of these variables that yields the highest match rate with the ASEC. These are also the versions of the variable that we offer in the IPUMS CPS variables LINENO, AGE, SEX, and RACE. The ASEC only has one person identifier and one set of demographic variables in all years.

[table 9]

Stage 1: Our objective is to link as many records as possible. Stage 1 records are those that are uniquely identified by and linked using HRHHID, HUHHNUM, LINENO, AGE, SEX, and RACE. The majority of March BMS records are linked to the ASEC in this stage (see Table 10).

Stage 2: In the second stage of linking, we exclude LINENO to mitigate the problem of nonconsecutive and mismatched LINENO values across March BMS and ASEC files. We retain and attempt to link unlinked stage 1 records that are uniquely identified in each file using HRHHID, HUHHNUM, AGE, SEX, and RACE. The result is hundreds or thousands of additional matches (see Table 10).

Stage 3: Stage 3 linking includes the number of hours worked last week (HOURS) as a linking key along with Stage 2 linking keys. We link a handful of records in this stage (see Table10).

Stage 4: In stage 4, we use only household and person identifiers to link records, which is consistent with our approach to assigning MARBASECIDP in 1989 forward. This approach accounts for many successful merges between 1982 and 1988 (see Table 10).

Stage 5: In this stage, we link singletons who are in households that have the same number of individuals in the March BMS and the ASEC and where all other household members have already been linked. Stage 5 uses household identifiers and the number of persons in the household (NUMPER) as linking keys.[11] These singletons have not merged in previous stages due to mismatch of either LINENO or demographic variables but are the only remaining possible matches. For example, in the household below, by the time this merge stage is reached, the three year-old in the March BMS and the zero year-old in the ASEC are the only records in this household that are unmatched. Despite differences on LINENO and AGE, we link them in stage 5, as the only remaining household member that is in both files but still unlinked. This merge stage links individuals in 1982-1988.

HRHHID=80079928707,
HUHHNUM=1

LINENO		AGE		SEX		RACE		HOURS		Merge Stage
BMS	ASEC	BMS	ASEC	BMS	ASEC	BMS	ASEC	BMS	ASEC	
1	1	36	36	1	1	1	1	50	50	1
2	2	36	36	2	2	1	1	25	25	1
3	3	17	17	2	2	1	1	00	00	1
4	5	3	0	1	1	1	1	00	00	5
5	4	7	7	1	1	1	1	00	00	2

---

[11] We generate the count of individuals per household used for linking (NUMPER) based on HHID values. It may differ from the IPUMS CPS variable NUMPREC for some households in the ASEC files during this period. It is necessary to use this generated count rather than the variable from the original data file or NUMPREC due to the fact that some households as defined by HHID are split across multiple households as defined by HHSEQ in the ASEC files. For more information on split households in the ASEC files, see Appendix C.

Stage 6: Stage 6 linking addresses a specific problem with household number (HUHHNUM) in

1988. There are many households in 1988 where HUHHNUM does not match across the March

BMS and ASEC files, but whose members have the same values for AGE, SEX, and RACE

across files. Most of these households have an HUHHNUM value of 0 in the ASEC file, despite

HUHHNUM generally having a minimum value of 1. In this case, we eliminate HUHHNUM as

a linking key and add number of persons in the household (NUMPER) as a linking key. This

yields 336 additional linked records.

Household from March BMS 1988 linked to ASEC in Stage 6

HRHHID	HUHHNUM	LINENO	AGE	SEX	RACE	HOURS	NUMPER
336068000150	1	1	32	1	1	45	3
336068000150	1	2	33	2	1	00	3
336068000150	1	3	2	2	1	00	3

Household from ASEC 1988 linked to March BMS in Stage 6

HRHHID	HUHHNUM	LINENO	AGE	SEX	RACE	HOURS	NUMPER
336068000150	0	1	32	1	1	45	3
336068000150	0	2	33	2	1	00	3
336068000150	0	3	2	2	1	00	3

[table 10]

Comparing multi- and single-stage linking methods

Our multi-stage linking process yields more matches across March basic monthly and

ASEC files than a single-stage linking approach. Table 11 compares the number of linkages

resulting from single-stage linking using HRHHID and LINENO; single-stage linking using

HRHHID, LINENO, AGE, SEX, and RACE; and our multi-stage approach. Our multi-stage

approach yields more linkages and higher quality linkages than both single-stage linkage

approaches.

[table 11]

Validating March Basic and ASEC linkages

Table 12 shows the number of records from our multi-stage linking process that match on AGE,

SEX, and RACE in every year. Note that because AGE, SEX, and RACE were linking keys in

the first three stages, the number of linked records that match on AGE, SEX, and RACE is

extremely high across all years. AGE mismatches are most common followed by SEX and

RACE. Table 13 shows validation broken out by linking stage for 1982-1988. Stage 4, where we

use only household and person identifiers, validates most poorly.

[table 12]

[table 13]

Discussion

Generating a linking key that performs well for merging across months of the CPS and to ASEC

files from 1976 to 1988 is an involved process. Different numbers of people across files and

within households, duplicate and unlinkable records, and different coding schemes in linking

keys across files amplify the complexity of this endeavor. We describe the many problems we

encountered, the several steps we took to resolve problems, and the multi-stage methodologies

we used to create and add a single linking key to all CPS files from 1976 to 1988. By creating

CPSID(P) and MARBASECID(P), IPUMS saves time, eliminates duplication of effort, reduces

errors for individual researchers, and provides the research community with a common starting

point for linking across CPS data files.

The primary issues in the data that we encounter in linking across months are with

household and person identifiers. We made adjustments, which we described in this paper, to the

household identifiers based on extensive investigation into the identifiers and other variables in

the data, and a series of checks to ensure that the adjustments are plausible. Despite our efforts, we are still unable to make linkages between some months. As such, linkage rates in the 1976 to 1988 period are lower than observed in the 1989 forward period (Drew et al., 2014). Of particular concern is 1977 in which linkages between basic monthly-only and supplement-containing files are not possible. In terms of linking to other years from 1977, basic monthly-only files from this year may only be linked to 1976 and supplement-containing files may only be linked to 1978. Extending existing Census Bureau documentation about linking problems between 1985 and 1986, we find that the break occurs between June and July of 1985; no months may be linked across the break, though linkages are possible prior to and following the break. Our efforts result in many linkages that researchers can evaluate and incorporate in their analyses if they wish and additional documentation about linkages across months of the CPS.

We also created linkages between March BMS and ASEC records. These linkages allow us to add CPSID(P) to the ASEC and enable researchers to use ASEC data linked to non-March months of the CPS. We expected the ASEC to contain all individual records in the March BMS, but this was not always the case. In addition, we encountered problematic identifiers; to address this issue, we implemented a multi-stage process to identify matches between the March BMS and the ASEC and make information about the merge stage available to researchers in the IPUMS CPS variable MARBASECSTAGE. We were unable to link the 1977 March BMS and ASEC, meaning that the 1977 ASEC cannot be linked to any other months of CPS data using CPSID(P). We suspect that this is related to the D sample, though we have not located any documentation that confirms this suspicion.

In the creation of CPSID(P), we generally followed the rules elaborated previously (Drew et al., 2014) while addressing the challenges described above. Creating linkages between the

March BMS and ASEC and constructing MARBASECID(P) in 1976 to 1988 deviated

considerably from our approach in 1989 forward. Had we used the same logic as we did with the

more recent data, we would have missed many linkages, which would have reduced the value of

being able to use the ASEC in combination with monthly CPS data. Our multi-stage linking

approach is superior the single-stage approaches as evidenced by our analyses. It strikes a

balance between maximizing the number of possible links while applying a consistent linking

algorithm across years from 1976 to 1988.

  In short, this work serves as a resource for the community of researchers who wish to

leverage the underutilized panel component of the CPS. It sheds light on some of the mysteries

encountered when attempting to systematically link CPS data across months for a period of more

than two decades. It also serves to catalog the challenges we encountered and solutions we

employed when creating MARBASECID(P) and CPSID(P) in the 1976 to 1988 period. With this

information, researchers can make informed choices when balancing the risks and rewards

associated with leveraging the panel component of the CPS back to 1976.

References

Creighton, K. P.; Wilkinson, R. 1984. Redesign of the sample for the Current Population Survey. *Employment and Earnings*, 31(5), 7-10.

Drew JA, Flood S, Warren JR. 2014. 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*, 39(3):121-144.

Flood, SM, Pacas, JD. 2017. Using the Annual Social and Economic Supplement as Part of a Current Population Survey Panel. *Journal of Economic and Social Measurement*, 42(3-4): 225-248.

Hanson, R. H. 1978. *The Current Population Survey: Design and Methodology, Technical Paper 40*. Washington, D.C.: Government Printing Office.

McIntire, Robert J. 1980. The Mechanics of Matching CPS Microdata: Problems and Solutions. Pp. 5-6 in *Current Population Survey as a Longitudinal Data Base*, Bureau of Labor Statistics.

U.S. Census Bureau. 1987. *Current Population Survey, October 1986: School Enrollment* [codebook].

U.S. Census Bureau. 1994. *Current Population Survey, March 1994 Technical Documentation* [codebook].

U.S. Census Bureau. 2006. *Current Population Survey Design and Methodology Technical Paper 66*.

U.S. Census Bureau. 2019. *Current Population Survey Design and Methodology Technical Paper 77*.

Appendix A: Reconciling Multiple CPS File Versions

Basic Monthly-Only versus Supplement-Containing Files

A basic monthly file is available for each month of the CPS since 1976. This basic monthly-only file includes all basic demographic and labor force variables. In some months, supplement data are collected on specific topics (e.g., education, displaced workers). In these months, a second supplement-containing file is available that includes the basic monthly variables along with the supplement-specific variables. In recent years, basic monthly-only and supplement-containing files for a given month are not released simultaneously; accordingly, IPUMS CPS first processes the basic monthly-only file for each month and then adds topical supplement variables from supplement-containing files onto the basic monthly-only file as supplement-containing files are released by the Census Bureau.

While the basic monthly component of the basic monthly-only and supplement-containing files should be identical, this is not always the case in the 1976 to 1988 period. Furthermore, it is often challenging and sometimes impossible to link supplement-containing and basic monthly-only files from the same month. Because the basic monthly-only files are available in all months and are the core part of the CPS, we privilege (and harmonize) the basic monthly variables in the basic monthly-only files over the basic monthly variables in the supplement-containing files except in 1977. Below are detailed some choices we made in dealing with discrepancies between basic monthly-only and supplement-containing files.

Duplicate Records

Supplement-containing files in this period often contain partially or fully duplicated records (see Table A1). Fully duplicated records have the same identifiers and are completely identical across all variables. If a record is duplicated in both the basic monthly-only files and the

supplement-containing files, then both of the duplicated records' supplement variables are attached to the corresponding basic monthly-only records. If a duplicate record is present only in the supplement-containing file, only one of the duplicated pair is retained and combined with the basic monthly-only file. Partially duplicated records in supplement-containing files from 1976 to 1988 have identical basic monthly variables, but differing values for one or more supplement variables. We have no way of adjudicating between the partially duplicated records in the supplement-containing files. In these instances, neither record from the supplement-containing file is appended to the basic monthly-only file.

[table a1]

Unmatchable Supplement Records

Very occasionally, records in a supplement-containing file cannot be paired with any record in the basic monthly-only file (see Table A1). These records are not partial or complete duplicates, but cannot be matched based on uniquely identifying linking keys to a corresponding record in the basic monthly-only file. Because they cannot be linked to a basic monthly-only record, these supplement records are not available in IPUMS CPS data. Occasionally in this period, supplement data can be linked with the basic monthly-only file but are not because the records in the basic monthly-only file are classified as non-interview records.

Differing Record Counts

Record counts often differ between basic monthly-only and supplement-containing files, especially between 1976 and 1981. Some supplement files between 1978 and 1981 contain individuals younger than 14 even though basic monthly-only files only contain individuals whose age is 14 or older. In these instances, we retain the child records from the supplement-containing file, including their basic monthly responses. Table A2 identifies samples from this

period that contain children under the age of 14. Apart from child records where available in

1978-1981, we do not retain surplus records in the supplement-containing file; in the case of

surplus records in the basic monthly-only file, supplement information is simply missing for

these records. Table A3 shows the number of records in both the basic monthly-only and

supplement-containing files, differentiating between household and person records.

[table a2, table a3]

Supplement-containing files in 1977

There is one exception to our decision to prioritize the basic monthly-only over the

supplement-containing file. Beginning with the supplement-containing files in 1977, a new "D

sample" was added to the CPS for the purpose of measuring the efficacy of a jobs training

program (Bureau of the Census, 1978; McIntire, 1980). With the introduction of these new

records, the method for generating household identifiers also changed. However, in months that

had no topical supplement, household identifiers generated using the old method remained. As a

result, it is impossible to link records between supplement-containing and basic monthly-only

files in 1977 and we are unable to follow normal procedure of merging supplement variables

onto basic monthly-only records. In these instances, we deliver the supplement containing-files

only.

Supplement-containing files in 1976

In 1976, after the HHID fix described in section *Problems with Household Identifiers and*

Linking across Months (concatenating the first nine digits of HHID with blank3 and blank6),

roughly 75% of records in the supplement-containing file have household identifiers that

correspond to household identifiers in the basic monthly-only file. The remaining 25% of

supplement-containing file records match on the person identifier, age, sex, and race, and have

household identifiers that differ from the household identifiers of the remaining un-matched records in the basic monthly-only file in systematic ways and match on the person identifier, age, sex, and race. We attach supplement variables to the basic monthly-only file in a three-stage process; HRHHID is the only linking key that changes with each stage. All other keys remain constant within month. Table A4 lists the linking keys for each supplement-containing month in 1976.

[table A4]

Each stage involves a different modification to the constructed HRHHID in the supplement-containing file.

- Stage 1: We attempt to merge all records using the constructed household id (HRHHID). After this stage, most records have successfully merged (see Table A5).

- Stage 2: For household identifiers of supplement-containing records that remain unmerged after stage 1, the fourth digit of the constructed household id is changed from "0" to "1". Most of the remaining supplement-containing records are matched to basic monthly-only records after stage 2.

- Stage 3: Visual examination of the few remaining unmerged records after Stage 2 revealed that the supplement-containing records that remain unmerged after stage 2 have either "12501" or "14600" in the first five digits of the constructed household id while the remaining unmerged records from the basic monthly-only file had values of "94003" or "94409" in these positions. Translating "12501" to "94003" and "14600" to "94409" in the first five digits of the constructed HRHHID of the remaining unmerged supplement-containing records resulted in the linking of all remaining match-able records in the basic monthly-only and supplement-containing files.

[table A5]

At the end of these three stages, all records in supplement-containing files are attached to the basic monthly-only file. As a result of these manipulations, we can attach the supplement variables to the end of the basic monthly records as in other years and we can then use the basic monthly version of the household identifier when linking across months instead of the supplement-containing version. This almost always achieves better linkage rates across time than the supplement-containing version of the household identifier, as shown in Table A6. Linkage rates across months improve by as much as 15 percentage points when the basic monthly-only version of the household identifier. In April of 1976, the basic monthly household identifier performs slightly worse than the supplement-containing version when linking across months. However, in the linkage rates decrease by less than one percentage point.

[table A6]

Appendix B: Versions of Demographic Variables in the March Basic Monthly File, 1976-1988

In 1976-1988 March Basic Monthly files, there are multiple person identifier and demographic variables. Census Bureau and Bureau of Labor Statistics documentation is scant on these variables and differences between them. Understanding the differences between these variables is important to achieving the highest quality link possible since we must include demographic variables in our linking algorithm to create MARBASECID(P) for 1976-1988.

There are up to three versions of the demographic and person variables in this period. We refer to these as the "basic edit", the "new demographic edit", and the "unedited" versions. March Basic Monthly files from 1976-1983 contain the basic edit and the unedited person identifiers and demographic variables. The basic edit versions contain meaningful values for all records in the file during these years; the unedited variables are missing for some records. In 1976-1983, the basic edit version is used to make the link between the March Basic Monthly file and the ASEC. Table B1 shows the differences between the basic edit and new demographic edit versions of AGE, SEX, and RACE for all records in 1976-1988.

[table_b1]

Beginning in 1984, new demographic edit versions are also available. Scant documentation suggests that the basic edit should be used to replicate BLS published estimates and that the new demographic edit variables should be used to replicate BLS family data (U.S. Census Bureau, 1987). Beginning in 1984, more detailed codes were introduced in the variable identifying the relationship to the household head. At this time, editing procedures for age, sex, and race

variables were modified to include more information from parent and spouse pointer variables.[12]

In March Basic Monthly files from 1984-1988, the new demographic edit person line number

and demographic variables are the only versions of these variables that have meaningful values

for all persons in the file. The basic edit versions are missing for all armed forces records.

Discrepancies between the basic edit and the new demographic edit variables for non-armed

forces records are negligible (see Table B1). In 1984-1988, we rely on the new demographic edit

version of these variables.

In all years from 1976-1988, the unedited versions of LINENO, AGE, SEX, and RACE

perform the worst as linking keys. In 1976-1983 we match a few more records successfully

between the March Basic Monthly and the ASEC using the basic edit versions of LINENO,

AGE, SEX, and RACE – between 8 and 29 more. In 1984-1988, we achieve a few more

successful links between the March Basic Monthly and ASEC files using the new demographic

edit versions of LINENO, AGE, SEX, and RACE as compared to the basic edit version –

between 9 and 56. Table B2a shows the difference in successful matches by stage using unedited

and basic edit versions of the person identifier and demographic variables in 1976-1983. Table

B2b shows the difference in successful matches by stage using unedited, basic edit, and new

demographic edit variables in 1984-1988.

[table_b2a]

[table_b2b]

The choice of person identifier and demographic variable version matters most when it

comes to validation of links. The linking keys that yield the largest number of links between

March Basic Monthly and ASEC files also result in the highest validation rates. Table B3a and

---

[12] Based on a conversation with Census Bureau staff.

B3b compares validation by stage for AGE, SEX, and RACE using unedited, basic edit, and new demographic versions of the demographic linking variables. In 1976-1983, all merges in all stages are valid based on AGE, SEX, and RACE using the basic edit version of the linking keys, while not all links made with the unedited versions validate. In 1984-1988, all versions of the linking keys achieve validation rates of over 99%, however the links made with the new demographic edit variables are the highest.

[table_b3a]

[table_b3b]

In 1984-1988, most of the improvement in validation comes from the armed forces records. Validation rates on demographic characteristics are better using the new demographic edit, as armed forces records can be properly validated instead of all being considered invalid due to missing demographic information. Validation rates for non-armed forces records do not differ greatly using the new demographic edit version of the demographic variables.

[table_b4]

The versions of the person identifier and demographic variables that we used to link March Basic Monthly records to ASEC records (basic edit for 1979-1983 and new demographic edit for 1984-1988) are the versions that are available as part of the harmonized variables LINENO, AGE, SEX, and RACE in IPUMS CPS. However, we have also made the unedited and basic edit versions of these variables available through the IPUMS CPS extract system as unharmonized variables. [13]

[table_b5]

---

[13] For more information on IPUMS CPS unharmonized variables, see https://cps.ipums.org/cps/unharmonized_variables.shtml

Appendix C: Split Households in ASEC Files and Assigning

MARBASECIDH

Unlike Basic Monthly files, ASEC files contain two different household identifiers, one that is intended to be used for linking across files (HRHHID), and a second one, generated by the Census Bureau to delineate households within the file (HHSEQ). In the 1976-1987 period, these household identifiers do not always map perfectly to one another; some HRHHID values are "split" across multiple HHSEQ values. This split occurs in three ways: HRHHID values are split such that all HHSEQ values that have the same HRHHID contain no people, one HHSEQ household is empty and the other has people, or both HHSEQ households have people. In this last category, in almost all instances (except for a few in 1986 and 1987), one HHSEQ household contains the household respondent, and the other does not. See Table C1 for a breakdown of these categories across years.

[table c1]

When HRHHID values are split between one empty and one peopled household, the household record for the household that contains people is assigned a linking MARBASECIDH value. The empty household is considered unlinkable.

In general, of the HRHHID-defined households that are split across multiple HHSEQ values where both HHSEQ-defined households contain persons, only one HHSEQ-defined household has a household respondent. There are a few exceptions to this in 1986 and 1987. In instances where only one household has a household respondent, the household record for the household in which the household respondent resides is assigned a linking MARBASECIDH value. For instances where both of the split households contain a household respondent, the

household record for the household with the most people is assigned a linking MARBASECIDH value. The household record for the split household that does not contain a household respondent or that has the fewest people is assigned a non-linking MARBASECIDH value, but all persons within the split household that link to the March Basic Monthly file as described above will have linking values of MARBASECIDP.

Table 1. Number of people responding to the CPS, by calendar month, month-in-sample group, and year

	1976												1977			
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr
MIS1	12,287	12,138	12,216	12,455	12,626	12,232	11,863	12,120	12,406	15,672	12,330	12,679	14,699	12,218	12,173	12,507
MIS2	12,194	12,509	12,209	12,475	12,565	12,795	12,279	12,094	12,385	15,750	12,641	12,508	14,842	12,780	12,197	12,328
MIS3	12,203	12,169	12,405	12,265	12,423	12,536	12,768	12,328	12,185	15,351	12,575	12,758	14,629	12,771	12,686	12,199
MIS4	12,080	12,128	12,147	12,467	12,251	12,341	12,559	12,681	12,313	15,305	12,343	12,638	14,803	12,491	12,636	12,772
MIS5	12,304	11,836	12,167	12,435	12,505	12,164	12,115	12,502	12,664	15,070	12,227	11,986	14,676	12,401	12,414	12,541
MIS6	12,336	12,481	11,837	12,347	12,432	12,580	12,255	12,189	12,654	16,074	12,218	12,353	14,491	12,482	12,371	12,451
MIS7	12,382	12,306	12,387	11,865	12,253	12,351	12,493	12,268	12,221	15,725	12,874	12,303	14,338	11,986	12,412	12,356
MIS8	12,490	12,351	12,191	12,461	11,866	12,241	12,267	12,515	12,241	15,413	12,664	12,849	14,496	12,349	11,989	12,443
Total	98,276	97,918	97,559	98,770	98,921	99,240	98,599	98,697	99,069	124,360	99,872	100,074	116,974	99,478	98,878	99,597
	1987												1988			
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr
MIS1	18,684	18,373	18,547	19,087	18,005	18,088	18,399	18,278	18,679	18,730	18,273	18,594	18,269	18,294	18,455	17,287
MIS2	18,254	18,955	18,482	19,060	19,412	18,290	18,467	18,908	18,645	19,239	18,838	18,673	18,991	18,476	18,570	17,725
MIS3	18,857	18,290	18,912	18,580	19,023	19,321	18,106	18,488	19,122	18,962	19,025	18,814	18,658	18,973	18,381	17,391
MIS4	18,514	18,767	18,127	18,824	18,484	18,840	19,022	18,076	18,496	19,191	18,735	18,829	18,633	18,682	18,939	17,192
MIS5	17,691	18,105	18,677	18,619	18,567	18,466	18,435	18,691	18,421	18,674	18,675	18,167	18,514	18,328	18,582	17,696
MIS6	18,855	17,975	18,119	18,974	18,741	18,623	18,643	18,717	18,928	18,694	18,638	18,879	18,312	18,632	18,319	17,486
MIS7	18,806	18,832	17,915	18,063	18,935	18,591	18,434	18,613	18,745	19,099	18,552	18,535	18,891	18,288	18,491	16,929
MIS8	18,922	18,689	18,764	17,905	17,984	18,801	18,521	18,495	18,637	18,785	18,845	18,338	18,501	18,893	18,200	17,221
Total	148,583	147,986	147,543	149,112	149,151	149,020	148,027	148,266	149,673	151,374	149,581	148,829	148,769	148,566	147,937	138,927

Table 2. Sample size and retention rate, CPS respondents linked across two consecutive calendar months

| | Year X | | 1976 | | | 1987 | | |
	Jan	Feb	Jan	Feb (All)	Feb (Plausible)	Jan	Feb (All)	Feb (Plausible)
MIS1	Jan_X	Feb_X	12,287	-	-	18,684	-	-
MIS2	Dec_{X-1}	Jan_X	12,194	11,434	11,168	18,254	17,652	17,125
MIS3	Nov_{X-1}	Dec_{X-1}	12,203	11,385	11,180	18,857	17,241	16,803
MIS4	Oct_{X-1}	Nov_{X-1}	-	11,443	11,261	-	17,869	17,437
MIS5	Jan_{X-1}	Feb_{X-1}	12,304	-	-	17,691	-	-
MIS6	Dec_{X-2}	Jan_{X-1}	12,336	11,423	11,187	18,855	16,562	16,105
MIS7	Nov_{X-2}	Dec_{X-2}	12,382	11,346	11,111	18,806	17,585	17,193
MIS8	Oct_{X-2}	Nov_{X-2}	-	11,474	11,264	-	17,510	17,097
Total			73,706	68,505	67,171	111,147	104,419	101,760
Retention rate				92.94%	91.13%		93.95%	91.55%

Table 3. Sample size and retention rate, CPS respondents linked across two non-consecutive calendar months

	Year X		1976			1987		
	Oct	Dec	Oct	Dec (All)	Dec (Plausible)	Oct	Dec (All)	Dec (Plausible)
MIS1	Oct_X	Dec_X	15,672	-	-	18,730	-	-
MIS2	Sep_X	Nov_X	15,750	-	-	19,239	-	-
MIS3	Aug_X	Oct_X	-	11,485	11,180	-	17,193	16,533
MIS4	Jul_X	Sep_X	-	11,675	11,449	-	17,544	17,003
MIS5	Oct_{X-1}	Dec_{X-1}	15,070	-	-	18,674	-	-
MIS6	Sep_{X-1}	Nov_{X-1}	16,074	-	-	18,694	-	-
MIS7	Aug_{X-1}	Oct_{X-1}	-	10,959	10,762	-	16,740	16,185
MIS8	Jul_{X-1}	Sep_{X-1}	-	11,730	11,499	-	16,672	16,214
Total			62,566	45,849	44,890	0 75,337	68,149	65,935
Retention rate				73.28%	71.75%		90.46%	87.52%

Table 4. Sample size and retention rate, CPS respondents linked in March across two consecutive years

	March		March			March		
	Year X	Year X+	1976	1977 (All)	1977 (Plausible)	1987	1988 (All)	1988 (Plausible)
MIS1	Mar_X	Mar_{X+1}	12,216	-	-	18,547	-	-
MIS2	Feb_X	Feb_{X+1}	12,209	-	-	18,482	-	-
MIS3	Jan_X	Jan_{X+1}	12,405	-	-	18,912	-	-
MIS4	Dec_{X-1}	Dec_X	12,147	-	-	18,127	-	-
MIS5	Mar_{X-1}	Mar_X	-	9,277	8,941	-	13,687	12,991
MIS6	Feb_{X-1}	Feb_X	-	9,212	8,927	-	13,583	12,970
MIS7	Jan_{X-1}	Jan_X	-	9,412	9,132	-	13,984	13,345
MIS8	Dec_{X-2}	Dec_{X-1}	-	9,144	8,843	-	13,636	13,008
Total			48,977	37,045	35,843	74,068	54,890	52,314
Retention rate				75.64%	73.18%		74.11%	70.63%

Table 5. Number and percentage of people responding to subsequent CPS surveys among those beginning the CPS in January 1976 and 1987

	1976			
	All links	Retention rate	Plausible links	Retention rate
Began in CPS in MIS 1 in Jan 1976…	12,287	-	-	-
…and also responded in February 1976	11,434	93.06%	11,168	90.89%
…and responded on all 3 occasions through March 1976	10,938	89.02%	10,636	86.56%
…and responded on all 4 occasions through April 1976	10,486	85.34%	10,209	83.09%
…and responded on all 5 occasions through January 1977	0	0.00%	0	0.00%
…and responded on all 6 occasions through February 1977	0	0.00%	0	0.00%
…and responded on all 7 occasions through March 1977	0	0.00%	0	0.00%
…and responded on all 8 occasions through April 1977	0	0.00%	0	0.00%
	1988			
	All links	Retention rate	Plausible links	Retention rate
Began in CPS in MIS 1 in Jan 1987…	18,684	-	-	-
…and also responded in February 1987	17,652	94.48%	17,125	91.66%
…and responded on all 3 occasions through March 1987	16,781	89.81%	16,212	86.77%
…and responded on all 4 occasions through April 1987	16,052	85.91%	15,553	83.24%
…and responded on all 5 occasions through January 1988	12,985	69.50%	12,426	66.51%
…and responded on all 6 occasions through February 1988	12,546	67.15%	12,123	64.88%
…and responded on all 7 occasions through March 1988	12,101	64.77%	11,665	62.43%
…and responded on all 8 occasions through April 1988	11,012	58.94%	10,627	56.88%

Table 6. Number and percentage of people responding to subsequent CPS surveys among those beginning the CPS in January 1976 and 1987

	1976			
	All links	Retention rate	Plausible links	Retention rate
Began in CPS in MIS 1 in Jan 1976…	12,287	-	-	-
…and also responded in ANY of the subsequent surveys between February 1976 and April 1977	11,775	95.83%	11,574	94.20%
	1987			
	All links	Retention rate	Plausible links	Retention rate
Began in CPS in MIS 1 in Jan 1987…	18,684	-	-	-
…and also responded in ANY of the subsequent surveys between February 1987 and April 1988	18,135	97.06%	17,757	95.04%

Table 7. Number and percentage of people responding to subsequent CPS surveys among those beginning the CPS in January 1976 and 1988

	1976			
	All links	Retention rate	Plausible links	Retention rate
Began in CPS in MIS 1 in Jan 1976…	12,287	-	-	-
…and also responded in ANY of the subsequent surveys between February 1976 and April 1977	9,475	77.11%	9,057	73.71%
	1988			
	All links	Retention rate	Plausible links	Retention rate
Began in CPS in MIS 1 in Jan 1987…	18,684	-	-	-
…and also responded in ANY of the subsequent surveys between February 1987 and April 1988	14,676	78.55%	13,855	74.15%

Table 8. Sample size and retention rate, CPS respondents in month-in-sample 4 linked to month-in-sample 5

	Year X												Year X+1			
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr
MIS4																
MIS5																

	1976												1977			
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr
MIS4	12,080	12,128	12,147	12,467	12,251	12,341	12,559	-	-	-	-	-	-	-	-	-
MIS5 (All)	-	-	-	-	-	-	-	-	-	9,438	9,494	9,379	0	9,493	9,763	9,767
Retention rate (All)	-	-	-	-	-	-	-	-	-	78.13%	78.28%	77.21%	0.0%	77.49%	79.11%	77.77%
MIS5 (Plausible)	-	-	-	-	-	-	-	-	-	9,206	9,255	9,131	0	9,247	9,511	9,523
Retention rate (Plausible)	-	-	-	-	-	-	-	-	-	76.21%	76.31%	75.17%	0.0%	75.48%	77.07%	75.83%

	1987												1988			
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr
MIS4	18,514	18,767	18,127	18,824	18,484	18,840	19,022	-	-	-	-	-	-	-	-	-
MIS5 (All)	-	-	-	-	-	-	-	-	-	14,345	14,552	14,148	14,634	14,402	14,786	14,077
Retention rate (All)	-	-	-	-	-	-	-	-	-	77.48%	77.54%	78.05%	77.74%	77.92%	78.48%	74.00%
MIS5 (Plausible)	-	-	-	-	-	-	-	-	-	13,753	14,000	13,632	14,087	13,875	14,183	13,539
Retention rate (Plausible)	-	-	-	-	-	-	-	-	-	74.28%	74.60%	75.20%	74.84%	75.06%	75.28%	71.18%

Table 9. Multi-stage Merge Linking Keys, 1976-1988

Merge Stage	Linking Keys	Years
1	HRHHID, HUHHNUM, LINENO, AGE, SEX, RACE	1976-1988
2	HRHHID, HUHHNUM, AGE, SEX, RACE	1976-1988
3	HRHHID, HUHHNUM, AGE, SEX, RACE, HOURS	1976-1988
4	HRHHID, HUHHNUM, LINENO	1982-1988
5	HRHHID, HUHHNUM, NUMPER	1982-1988
6	HRHHID, NUMPER, AGE, SEX, RACE	1988

Table 10. Multi-stage Merge of March Basic Monthly and ASEC Files, 1976-1988

	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988
A. Total BMS													
N	97559		113992	113689	133914	133488	153195	153431	152614	152705	149440	147543	147937
B. Merged BMS													
N	97511		113935	113661	133889	133417	153017	153201	152418	152466	147383	145594	145041
%	99.95%		99.95%	99.98%	99.98%	99.95%	99.88%	99.85%	99.87%	99.84%	98.62%	98.68%	98.04%
Merge Stage													
1	91988		106720	113606	133828	133394	152888	152869	151402	151338	145956	144140	143058
2	5505		7191	55	59	21	48	187	873	980	971	953	387
3	18		24	0	2	2	2	12	14	8	10	28	0
4	-		-	-	-	-	78	128	120	122	440	457	1246
5	-		-	-	-	-	1	5	9	18	6	16	12
6	-		-	-	-	-	-	-	-	-	-	-	338
C. Unmerged BMS													
N	48		57	28	25	71	178	230	196	239	2057	1949	2896
HHID in ASEC but individual is not													
N	36		29	18	10	21	145	168	155	189	205	187	469
HHID not in ASEC													
N	12		28	10	15	50	33	62	41	50	1852	1762	2427

Table 11. Comparison Between Multi Stage and Single Stage Merges

	total bms	bms duplicates	single-stage merge, identifiers only				bms duplicates	single-stage merge, incl. demographic variables				multi-stage merge			
			successful merges		valid merges (age, sex, race)			successful merges		valid merges (age, sex, race)		successful merges		valid merges (age, sex, race)	
			n	%	n	%		n	%	n	%	n	%	n	%
1976	97,559	250	94830	97.20%	91,870	94.17%	14	91,988	94.29%	91,988	94.29%	97,511	99.95%	97,511	99.95%
1977															
1978	113,992	593	110,222	96.69%	106,434	93.37%	12	106,720	93.62%	106,720	93.62%	113,935	99.95%	113,935	99.95%
1979	113,689	392	113,226	99.59%	113,226	99.59%	18	113,606	99.93%	113,606	99.93%	113,661	99.98%	113,661	99.98%
1980	133,914	652	133,185	99.46%	133,185	99.46%	12	133,828	99.94%	133,828	99.94%	133,889	99.98%	133,889	99.98%
1981	133,488	598	132,818	99.50%	132,818	99.50%	22	133,394	99.93%	133,394	99.93%	133,417	99.95%	133,417	99.95%
1982	153,195	939	152,137	99.31%	152,058	99.26%	80	152,888	99.80%	152,888	99.80%	153,017	99.88%	152,938	99.83%
1983	153,431	1725	151,436	98.70%	151,310	98.62%	109	152,869	99.63%	152,869	99.63%	153,201	99.85%	153,068	99.76%
1984	152,614	0	150,806	98.82%	150,681	98.73%	0	151,402	99.21%	151,402	99.21%	152,418	99.87%	152,289	99.79%
1985	152,705	0	150,673	98.67%	150,541	98.58%	0	151,338	99.10%	151,338	99.10%	152,466	99.84%	152,327	99.75%
1986	149,440	0	145,635	97.45%	145,184	97.15%	0	145,956	97.67%	145,956	97.67%	147,383	98.62%	146,937	98.33%
1987	147,543	0	143,843	97.49%	143,368	97.17%	0	144,140	97.69%	144,140	97.69%	145,594	98.68%	145,121	98.36%
1988	147,937	0	144,701	97.81%	143,058	96.70%	0	143,058	96.70%	143,058	96.70%	145,041	98.04%	143,783	97.19%

Table 12. March Basic Monthly to ASEC Merges Validation on AGE, SEX, RACE, 1976-1978

	AGE		SEX		RACE	
	Match	Nonmatch	Match	Nonmatch	Match	Nonmatch
1976	97,511	0	97,511	0	97,511	0
1977						
1978	113,935	0	113,935	0	113,935	0
1979	113,627	2	113,629	0	113,629	0
1980	133,861	0	133,861	0	133,861	0
1981	133,388	0	133,388	0	133,388	0
1982	152,916	47	152,951	12	152,938	25
1983	153,120	78	153,157	41	153,184	14
1984	151,695	79	151,744	30	151,762	12
1985	151,764	95	151,844	15	151,852	7
1986	146,379	384	146,729	34	146,742	21
1987	144,590	413	144,976	27	144,988	15
1988	143,326	1,059	144,000	385	144,240	145

Table 13. Validation on Demographic Characteristics Broken Out by Merge Stage

1982 (N=153195)	Merges by stage	Matches by characteristic			
		AGE	SEX	RACE	AGE, SEX, and RACE
Stage 1	152888	152888	152888	152888	152888
Stage 2	48	48	48	48	48
Stage 3	2	2	2	2	2
Stage 4	78	35	68	53	0
Stage 5	1	0	1	1	0
Stage 6	0	0	0	0	0
N matches	153017	152973	153007	152992	152938
% matches	--	99.97%	99.99%	99.98%	99.95%

1983 (N=153431)	Merges by stage	Matches by characteristic			
		AGE	SEX	RACE	AGE, SEX, and RACE
Stage 1	152869	152869	152869	152869	152869
Stage 2	187	187	187	187	187
Stage 3	12	12	12	12	12
Stage 4	128	52	87	114	0
Stage 5	5	0	4	5	0
Stage 6	0	0	0	0	0
N matches	153201	153120	153159	153187	153068
% matches	--	99.95%	99.97%	99.99%	99.91%

1984 (N=152614)	Merges by stage	Matches by characteristic			
		AGE	SEX	RACE	AGE, SEX, and RACE
Stage 1	151402	151402	151402	151402	151402
Stage 2	873	873	873	873	873
Stage 3	14	14	14	14	14
Stage 4	120	48	80	107	0
Stage 5	9	0	9	9	0
Stage 6	0	0	0	0	0
N matches	152418	152337	152378	152405	152289
% matches	--	99.95%	99.97%	99.99%	99.92%

1985 (N=152614)	Merges by stage	Matches by characteristic			
		AGE	SEX	RACE	AGE, SEX, and RACE
Stage 1	151338	151338	151338	151338	151338
Stage 2	980	980	980	980	980
Stage 3	8	8	8	8	8
Stage 4	122	36	95	113	0
Stage 5	18	2	17	18	1
Stage 6	0	0	0	0	0
N matches	152466	152364	152438	152457	152327
% matches	--	99.93%	99.98%	99.99%	99.91%

1986 (N=149440)	Merges by stage	Matches by characteristic			
		AGE	SEX	RACE	AGE, SEX, and RACE
Stage 1	145956	145956	145956	145956	145956
Stage 2	971	971	971	971	971
Stage 3	10	10	10	10	10
Stage 4	440	59	401	418	0
Stage 5	6	0	6	6	0
Stage 6	0	0	0	0	0
N matches	147383	146996	147344	147361	146937
% matches	--	99.74%	99.97%	99.99%	99.70%

1987 (N=147543)	Merges by stage	Matches by characteristic			
		AGE	SEX	RACE	AGE, SEX, and RACE
Stage 1	144140	144140	144140	144140	144140
Stage 2	953	953	953	953	953
Stage 3	28	28	28	28	28
Stage 4	457	58	415	441	0
Stage 5	16	1	15	16	0
Stage 6	0	0	0	0	0
N matches	145594	145180	145551	145578	145121
% matches	--	99.72%	99.97%	99.99%	99.68%

1988 (N=147937)	Merges by stage	Matches by characteristic			
		AGE	SEX	RACE	AGE, SEX, and RACE
Stage 1	143058	143058	143058	143058	143058
Stage 2	387	387	387	387	387
Stage 3	0	0	0	0	0
Stage 4	1246	199	881	1102	0
Stage 5	12	1	7	11	0
Stage 6	338	338	338	338	338
N matches	145041	143983	144671	144896	143783
% matches	--	99.27%	99.74%	99.90%	99.13%

Table A1. Duplicate records in basic monthly-only and supplement-containing files, 1976-1988

Year	Month	Basic monthly-only file sets of Duplicate Records	Supplement-containing file sets of Duplicate Records	Partial Duplicate Records	Unmatchable records*
1976	January	0	0	0	1
	April	0	0	0	0
	May	0	0	0	1
	June	0	0	348	0
	October	0	0	0	0
	November	0	0	11	0
1977	January	na	0	0	na
	May	na	0	0	na
	June	na	0	0	na
	August	na	0	0	na
	October	na	2	0	na
	December	na	0	0	na
1978	January	0	0	0	0
	May	0	0	0	0
	September	0	0	2	0
	October	0	0	0	0
	November	0	1	0	0
1979	May	0	8	0	0
	June	0	7	2	0
	September	0	7	2	0
	October	0	4	2	0
	November	0	1	0	0
	December	0	5	0	0
1980	May	0	14	0	0
	June	0	8	0	0
	September	0	2	0	0
	October	0	4	2	0
	November	0	21	0	0
1981	January	0	6	0	-
	May	0	3	0	0
	June	0	8	0	0
	September	0	0	0	0
	October	0	2	2	0
	December	0	3	0	2
1982	June	3	3	0	1
	September	2	1	2	0
	October	2	2	2	0
	November	2	2	0	0
1983	January	1	1	0	0
	April	4	4	0	0
	May	2	2	0	2
	June	3	3	0	0
	September	3	2	2	2
	October	2	1	2	0
	November	3	3	0	0
	December	2	2	0	0
1984	January	0	0	0	0
	May	0	0	0	0
	June	0	0	0	0
	September	0	0	0	0
	October	0	6	2	0
	November	0	0	0	0
	December	0	0	0	0
1985	April	0	0	0	0
	May	0	0	0	0
	June	0	11	6	11
	September	0	0	0	6
	October	0	2	2	0
	December	0	0	0	0
1986	January	0	0	0	0
	June	0	0	0	0
	October	0	8	0	0
	November	0	0	0	2
1987	January	0	0	0	0
	June	0	0	0	0
	October	0	0	0	13
	November	0	0	0	0
	December	0	68	41	4
1988	January	0	14	0	1
	May	0	0	6	3
	June	0	0	0	4
	August	0	0	0	3
	October	0	0	0	0
	November	0	0	0	1

* Does not include duplicate or partial duplicate records
- We have been unable to adequately reconcile these files

Table A2. Basic monthly-only and supplement-containing file characteristics, 1976-1993

Year	Month	Number of Records			Supplement Characteristics			
		Basic Monthly	Supplement	Excess Supplement Records	Topic	Duplicate Record sets	Partial Duplicate Records	Includes Children Under 14
1976	January	98,276	108,116	9,840	Tenure	0	0	
	April	98,770	108,696	9,926	Food Security	0	0	
	May	98,921	108,958	10,037	Work Schedules	0	0	
	June	99,240	109,173	9,933	Fertility & Marriage	0	348	
	October	99,720	133,697	33,977	Education	0	0	Yes
	November	99,872	109,203	9,331	Voter	0	11	
1977	January	99,341	129,031	29,690	Tenure	0	0	
	May	99,667	129,658	29,991	Work Schedules	0	0	
	June	99,469	129,401	29,932	Fertility & Marriage	0	0	
	August	99,259	129,072	29,813	Food Security	0	0	
	October	100,246	146,292	46,046	Education	2	0	Yes
	December	99,605	129,173	29,568	Agricultural Worker	0	0	
1978	January	128,788	116,421	-12367	Job Tenure	0	0	
	May	126,676	126,676/108,338	0/-18,338	Adult Education / Work Schedules	0	0	
	September	128,057	151,736	23,679	Immunization	0	2	Yes
	October	129,162	139,436	10,274	Education	0	0	Yes
	November	129,083	162,827	33,744	Voter	1	0	Yes
1979	May	127,376	160,544	33,168	Work Schedules	8	0	Yes
	June	127,660	160,921	33,261	Fertility & Marriage	7	2	Yes
	September	128,375	160,442	32,067	Immunization	7	2	Yes
	October	128,230	160,666	32,436	Education	4	2	Yes
	November	127,824	160,255	32,431	Language	1	0	Yes
	December	128,475	161,612	33,137	Agricultural Worker	5	0	Yes
1980	May	135,483	188,729	53,246	Work Schedules	14	0	Yes
	June	134,912	188,201	53,289	Fertility & Marriage	8	0	Yes
	September	135,045	187,533	52,488	Immunization	2	0	Yes
	October	135,471	188,318	52,847	Education	4	2	Yes
	November	136,122	189,074	52,952	Voter	21	0	Yes
1981	January	135,461	188,764	53,303	Job Tenure	6	0	
	May	122,843	170,954	48,111	Work Schedules	3	0	Yes
	June	122,524	170,359	47,835	Fertility & Marriage	8	0	Yes
	September	122,720	170,173	47,453	Immunization	0	0	Yes
	October	123,130	170,566	47,436	Education	2	2	Yes
	December	122,741	169,813	47,072	Agricultural Worker	3	0	Yes
1982	June	169,440	169,436	-4	Fertility & Marriage	0	0	Yes
	September	169,066	169,066	0	Immunization	0	2	Yes
	October	169,099	169,098	-1	Education	0	2	Yes
	November	169,324	169,324	0	Voter	0	0	Yes
1983	January	168,124	168,124	0	Job Tenure / Job Training	0	0	Yes
	April	168,594	168,594	0	Immigration	0	0	Yes
	May	168,216	168,216	0	Pension & Retirement	0	0	Yes
	June	168,015	168,015	0	Fertility & Marriage	0	0	Yes
	September	167,419	167,419	0	Immunization	0	2	Yes
	October	167,502	167,502	0	Education	0	2	Yes
	November	168,189	168,189	0	Telephone Availability	0	0	Yes
	December	167,906	167,906	0	Agricultural Worker	0	0	Yes
1984	January	167,985	167,984	-1	Displaced Worker / Job Training	0	0	Yes
	May	165,991	165,991	0	Adult Education	0	0	Yes
	June	166,055	166,055	0	Fertility & Marriage	0	0	Yes
	September	165,316	165,316	0	Immunization	0	0	Yes
	October	165,490	165,490	0	Education / Computer & Internet Use	6	2	Yes
	November	165,992	165,992	0	Voter	0	0	Yes
	December	166,590	166,590	0	Child Care	0	0	Yes
1985	April	167,213	167,213	0	Veterans / Houshold Composition	0	0	Yes
	May	165,498	165,498	0	Work Schedules	0	0	Yes
	June	165,471	165,471	0	Fertility & Marriage	22	6	Yes
	September	165,627	165,627	0	Tobacco Use / Immunization	6	0	Yes
	October	165,995	165,995	0	Education	4	2	Yes
	December	166,383	166,383	0	Agricultural Worker	0	0	Yes
1986	January	165,360	165,359	-1	Displaced Worker	0	0	Yes
	June	165,201	165,201	0	Fertility & Marriage	0	0	Yes
	October	163,303	161,750	-1,553	Education	4	0	Yes
	November	163,178	163,178	0	Voter	0	0	Yes
1987	January	162,875	162,875	0	Job Tenure	0	0	Yes
	June	163,541	163,541	0	Fertility & Marriage	0	0	Yes
	October	165,519	163,816	-1,703	Education	0	0	Yes
	November	163,806	163,806	0	Veterans	0	0	Yes
	December	163,325	163,325	0	Agricultural Worker	0	41	Yes
1988	January	163,357	163,371	14	Displaced Worker	14	0	Yes
	May	152,850	152,850	0	Employee Benefit	0	3	Yes
	June	152,460	152,460	0	Fertility & Marriage / Immigration	0	0	Yes
	August	152,539	152,539	0	Retiree Health Insurance	0	0	Yes
	October	154,224	154,224	0	Education	0	0	Yes
	November	154,311	154,311	0	Voter	0	0	Yes

Table A3. Basic monthly-only and supplement-containing file records by record type, 1976-1988

Year	Month	Basic					Supplement					Supplement Surplus				
		1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
1976	January	98,276	0	0	-	-	98,276	2,037	7,804	-	-	0	2,037	7,804	-	-
	April	98,770	0	0	-	-	98,770	2,155	7,771	-	-	0	2,155	7,771	-	-
	May	98,921	0	0	-	-	98,921	2,174	7,864	-	-	0	2,174	7,864	-	-
	June	99,240	0	0	-	-	99,240	2,205	7,728	-	-	0	2,205	7,728	-	-
	October	99,720	0	0	-	-	124,360	1,814	7,523	-	-	24,640	1,814	7,523	-	-
	November	99,872	0	0	-	-	99,872	1,902	7,429	-	-	0	1,902	7,429	-	-
1977	January	99,341	0	0	-	-	116,974	2,311	9,747	-	-	17,633	2,311	9,747	-	-
	May	99,667	0	0	-	-	117,441	2,399	9,818	-	-	17,774	2,399	9,818	-	-
	June	99,469	0	0	-	-	116,960	2,493	9,949	-	-	17,491	2,493	9,949	-	-
	August	99,259	0	0	-	-	116,694	2,425	9,953	-	-	17,435	2,425	9,953	-	-
	October	100,246	0	0	-	-	146,292	0	0	-	-	46,046	0	0	-	-
	December	99,605	0	0	-	-	117,294	2,108	9,771	-	-	17,689	2,108	9,771	-	-
1978	January	116,421	0	0	-	-	116,421	0	0	-	-	0	0	0	-	-
	May*	114,207	2,715	9,754	-	-	114,207	2,715	9,754	-	-	0	0	0	-	-
	September	115,881	2,329	9,847	-	-	139,560	2,329	9,847	-	-	23,679	0	0	-	-
	October	116,778	2,319	10,065	-	-	129,354	2,016	8,066	-	-	12,576	-303	-1,999	-	-
	November	116,686	2,495	9,902	-	-	150,430	2,495	9,902	-	-	33,744	0	0	-	-
1979	May	114,336	2,780	10,260	-	-	147,504	2,780	10,260	-	-	33,168	0	0	-	-
	June	114,764	2,662	10,234	-	-	148,025	2,662	10,234	-	-	33,261	0	0	-	-
	September	115,941	2,326	10,108	-	-	148,008	2,326	10,108	-	-	32,067	0	0	-	-
	October	115,785	2,249	10,196	-	-	148,221	2,249	10,196	-	-	32,436	0	0	-	-
	November	115,275	2,447	10,102	-	-	147,706	2,447	10,102	-	-	32,431	0	0	-	-
	December	116,142	2,272	10,061	-	-	149,279	2,272	10,061	-	-	33,137	0	0	-	-
1980	May	135,483	0	0	-	-	174,070	2,543	12,116	-	-	38,587	2,543	12,116	-	-
	June	134,912	0	0	-	-	173,229	2,817	12,155	-	-	38,317	2,817	12,155	-	-
	September	135,045	0	0	-	-	173,000	2,628	11,905	-	-	37,955	2,628	11,905	-	-
	October	135,471	0	0	-	-	173,988	2,488	11,842	-	-	38,517	2,488	11,842	-	-
	November	136,122	0	0	-	-	174,822	2,428	11,824	-	-	38,700	2,428	11,824	-	-
1981	January	135,461	0	0	-	-	174,021	2,665	12,078	-	-	38,560	2,665	12,078	-	-
	May	122,843	0	0	-	-	157,496	2,520	10,938	-	-	34,653	2,520	10,938	-	-
	June	122,524	0	0	-	-	156,921	2,546	10,892	-	-	34,397	2,546	10,892	-	-
	September	122,720	0	0	-	-	157,000	2,312	10,861	-	-	34,280	2,312	10,861	-	-
	October	123,130	0	0	-	-	157,636	2,231	10,700	-	-	34,506	2,231	10,700	-	-
	December	122,741	0	0	-	-	156,860	2,268	10,685	-	-	34,119	2,268	10,685	-	-
1982	June	156,398	2,392	10,650	0	0	156,393	2,392	10,650	0	0	-5	0	0	0	0
	September	156,318	2,222	10,526	0	0	156,318	2,222	10,526	0	0	0	0	0	0	0
	October	156,526	2,129	10,444	0	0	156,526	2,129	10,444	0	0	0	0	0	0	0
	November	156,984	1,970	10,370	0	0	156,984	1,970	10,370	0	0	0	0	0	0	0
1983	January	155,029	2,374	10,721	0	0	155,029	2,374	10,721	0	0	0	0	0	0	0
	April	155,267	2,451	10,876	0	0	155,267	2,451	10,876	0	0	0	0	0	0	0
	May	155,055	2,478	10,683	0	0	155,055	2,478	10,683	0	0	0	0	0	0	0
	June	154,694	2,574	10,747	0	0	154,694	2,574	10,747	0	0	0	0	0	0	0
	September	154,161	2,349	10,909	0	0	154,161	2,349	10,909	0	0	0	0	0	0	0
	October	154,588	2,300	10,614	0	0	154,588	2,300	10,614	0	0	0	0	0	0	0
	November	155,261	2,354	10,574	0	0	155,261	2,354	10,574	0	0	0	0	0	0	0
	December	154,705	2,445	10,756	0	0	121,237	2,445	10,756	0	33,468	-33,468	0	0	0	33,468
1984	January	120,303	2,669	11,016	649	33,348	120,303	2,669	11,016	649	33,348	0	0	0	0	0
	May	119,296	2,767	10,628	657	32,643	119,296	2,767	10,628	657	32,643	0	0	0	0	0
	June	119,508	2,812	10,574	680	32,481	119,508	2,812	10,574	680	32,481	0	0	0	0	0
	September	119,152	2,686	10,352	696	32,430	119,152	2,686	10,352	696	32,430	0	0	0	0	0
	October	119,405	2,531	10,438	689	32,427	119,405	2,531	10,438	689	32,427	0	0	0	0	0
	November	119,463	2,668	10,724	665	32,472	119,463	2,668	10,724	665	32,472	0	0	0	0	0
	December	120,181	2,588	10,591	647	32,583	120,181	2,588	10,591	647	32,583	0	0	0	0	0
1985	April	120,259	2,914	11,181	651	32,208	120,259	2,914	11,181	651	32,208	0	0	0	0	0
	May	118,982	2,929	11,058	635	31,894	118,982	2,929	11,058	635	31,894	0	0	0	0	0
	June	118,741	2,982	11,299	630	31,819	118,741	2,982	11,299	630	31,819	0	0	0	0	0
	September	119,284	2,729	11,026	635	31,953	119,284	2,729	11,026	635	31,953	0	0	0	0	0
	October	119,984	2,561	10,997	639	31,814	119,984	2,561	10,997	639	31,814	0	0	0	0	0
	December	119,911	2,662	11,203	652	31,955	119,911	2,662	11,203	652	31,955	0	0	0	0	0
1986	January	118,833	2,846	11,154	640	31,887	118,833	2,846	11,154	640	31,887	0	0	0	0	0
	June	118,283	2,888	11,623	669	31,738	118,283	2,888	11,623	669	31,738	0	0	0	0	0
	October	117,458	2,432	11,297	634	31,482	117,458	2,432	11,297	634	31,482	0	0	0	0	0
	November	117,290	2,515	11,218	656	31,499	117,290	2,515	11,218	656	31,499	0	0	0	0	0
1987	January	116,504	2,834	11,458	669	31,410	116,504	2,834	11,458	669	31,410	0	0	0	0	0
	June	117,004	2,751	11,770	654	31,362	117,004	2,751	11,770	654	31,362	0	0	0	0	0
	October	117,323	2,491	11,654	629	33,422	117,323	2,491	11,654	629	31,719	0	0	0	0	-1,703
	November	117,259	2,625	11,600	635	31,687	117,259	2,625	11,600	635	31,687	0	0	0	0	0
	December	116,763	2,766	11,726	621	31,449	116,763	2,766	11,726	621	31,449	0	0	0	0	0
1988	January	116,503	2,744	11,843	682	31,585	116,509	2,744	11,845	684	31,589	6	0	2	2	4
	May	109,192	2,505	11,451	647	29,055	109,192	2,505	11,451	647	29,055	0	0	0	0	0
	June	108,583	2,637	11,625	650	28,965	108,583	2,637	11,625	650	28,965	0	0	0	0	0
	August	108,571	2,592	11,513	641	29,222	108,571	2,592	11,513	641	29,222	0	0	0	0	0
	October	110,216	2,309	11,207	654	29,838	110,216	2,309	11,207	654	29,838	0	0	0	0	0
	November	110,452	2,382	11,046	644	29,787	110,452	2,382	11,046	644	29,787	0	0	0	0	0

* May 1978 has two separate supplement containing files. The Work Schedules supplement is shown in the table. The Adult education supplement file has 108,338 records of type 1.

Table A4. Linking keys for basic monthly-only and supplement-containing file merge, 1976

Month	Linking keys (IPUMS CPS variable names)
January	HRHHID*, UH_LINENO_B1, UH_AGE_B1, UH_SEX_B1, UH_RACE_B1, UH_MARSTAT_B1, UH_DOCNO_B1, UH_JRCERR_B1
April	HRHHID*, UH_LINENO_B1, UH_AGE_B1, UH_SEX_B1, UH_RACE_B1, UH_MARSTAT_B1, UH_DOCNO_B1, UH_JRCERR_B1, UH_HOURS_B1
May	HRHHID*, UH_LINENO_B1, UH_AGE_B1, UH_SEX_B1, UH_RACE_B1, UH_MARSTAT_B1, UH_DOCNO_B1, UH_JRCERR_B1, UH_HOURS_B1, UH_IND_B1
June	HRHHID*, UH_LINENO_B1, UH_AGE_B1, UH_SEX_B1, UH_RACE_B1, UH_MARSTAT_B1, UH_DOCNO_B1, UH_JRCERR_B1, UH_HOURS_B1
October	HRHHID*, UH_LINENO_B1, UH_AGE_B1, UH_SEX_B1, UH_RACE_B1, UH_MARSTAT_B1, UH_DOCNO_B1, UH_JRCERR_B1
November	HRHHID*, UH_LINENO_B1, UH_AGE_B1, UH_SEX_B1, UH_RACE_B1, UH_MARSTAT_B1, UH_DOCNO_B1, UH_JRCERR_B1

* HRHHID is modified in the supplement-containing file at each merge stage, all other keys remain constant across stages

Table A5. Supplement-containing records merged to basic-monthly only records by stage

Month	Merge Stage 1	Merge Stage 2	Merge Stage 3	Unlinkable bms records	Unlinkable supplement records
Jan 1976	76,041	22,234	0	1	0
Apr 1976	76,528	22,242	0	0	0
May 1976	76,456	22,283	181	1	0
June 1976	76,477	22,241	174	0	348**
Oct 1976*	76,962	22,573	185	0	0
Nov 1976	77,147	22,521	186	0	18**

* In October, there are 24,640 child records that cannot be merged to the basic monthly-ony file. These are added to the basic monthly-only file the same way child records in May 1978-Dec 1981

**These records are partial duplicates in the supplement-containing file and are not merged to the basic monthly-only file

Table A6. 1976 merge rates across months using supplement vs basic monthly household identifiers

Expected merge rate	Linked To		Jan 1976 (N= 98,276)			Apr 1976 (N= 98,770)			May 1976 (N= 98,921)			Jun 1976 (N= 99,240)			Oct 1976 (N= 124,360)*			Nov 1976 (N= 99,872)		
			Supplement	Basic	Gain	Supplement	Basic	Gain	Supplement	Basic	Gain	Supplement	Basic	Gain	Supplement	Basic	Gain	Supplement	Basic	Gain
75.00%	forward 1 month	Successful links	53,122	68,505	15,383	67,908	67,791	-117	68,211	68,211	0	52,628	68,049	15,421	69,800	69,800	0	54,227	70,096	15,869
		Actual merge rate	54.05%	69.71%		68.75%	68.64%		68.96%	68.96%		53.03%	68.57%		56.13%	56.13%		54.30%	70.19%	
50.00%	forward 2 months	Successful links	34,625	44,640	10,015	44,290	44,225	-65	34,308	44,351	10,043	33,951	44,040	10,089	35,480	45,849	10,369	0	0	0
		Actual merge rate	35.23%	45.42%		44.84%	44.78%		34.68%	44.83%		34.21%	44.38%		28.53%	45.98%		0.00%	0.00%	
25.00%	forward 3 months	Successful links	21,857	21,857	0	16,730	21,531	4,801	16,685	21,653	4,968	16,170	21,007	4,837	0	0	0	16,910	21,883	4,973
		Actual merge rate	22.24%	22.24%		16.94%	21.80%		16.87%	21.89%		16.29%	21.17%		0.00%	0.00%		16.93%	21.91%	
50.00%	forward 12 months	Successful links	0	0	0	28,779	37,184	8,405	0	0	0	0	0	0	0	0	0	28,743	37,034	8,291
		Actual merge rate	0.00%	0.00%		29.14%	37.65%		0.00%	0.00%		0.00%	0.00%		0.00%	0.00%		28.78%	37.08%	
37.50%	forward 13 months	Successful links	21,034	27,171	6,137	0	0	0	0	0	0	21,302	27,345	6,043	21,464	27,556	6,092	0	0	0
		Actual merge rate	21.40%	27.65%		0.00%	0.00%		0.00%	0.00%		21.47%	27.55%		17.26%	27.63%		0.00%	0.00%	
25.00%	forward 14 months	Successful links	13,736	17,833	4,097	0	0	0	14,118	18,114	3,996	0	0	0	0	0	0	0	0	0
		Actual merge rate	13.98%	18.15%		0.00%	0.00%		14.27%	18.31%		0.00%	0.00%		0.00%	0.00%		0.00%	0.00%	
12.50%	forward 15 months	Successful links	6,758	8,799	2,041	6,904	8,786	1,882	0	0	0	6,583	8,501	1,918	0	0	0	0	0	0
		Actual merge rate	6.88%	8.95%		6.99%	8.90%		0.00%	0.00%		6.63%	8.57%		0.00%	0.00%		0.00%	0.00%	

* October 1976 contains children under 14 while no other months in this year do. This is the reason for the high record count in October 1976.

Table B1. Number of Persons with Different Values in Unedited and New Demographic Edit Versions of Demographic Variables as compared to the Basic Edit Version (non-armed forces), 1976-1988 BMS

	AGE		SEX		RACE	
	New Demographic	Unedited	New Demographic	Unedited	New Demographic	Unedited
1976						
jan	--	247	--	356	--	115
feb	--	261	--	323	--	116
mar	--	55	--	248	--	25
apr	--	269	--	383	--	105
may	--	301	--	436	--	163
jun	--	266	--	418	--	141
jul	--	238	--	362	--	100
aug	--	265	--	377	--	110
sep	--	248	--	365	--	120
oct*	--	243	--	387	--	148
nov	--	184	--	340	--	96
dec	--	198	--	326	--	99
1977^						
jan	--	--	--	--	--	--
feb	--	228	--	371	--	121
mar	--	46	--	198	--	17
apr	--	280	--	389	--	131
may	--	--	--	--	--	--
jun	--	--	--	--	--	--
jul	--	205	--	362	--	107
aug	--	--	--	--	--	--
sep	--	241	--	382	--	119
oct	--	--	--	--	--	--
nov	--	213	--	332	--	105
dec	--	--	--	--	--	--
1978						
jan	--	306	--	480	--	159
feb	--	257	--	464	--	149
mar	--	95	--	341	--	49
apr	--	267	--	462	--	146
may	--	328	--	502	--	175
jun	--	266	--	479	--	145
jul	--	285	--	426	--	142
aug	--	265	--	360	--	108
sep*	--	286	--	474	--	139
oct*	--	274	--	389	--	159
nov*	--	270	--	410	--	132
dec	--	254	--	413	--	153
1979						
jan	--	460	--	328	--	215
feb	--	375	--	275	--	204
mar	--	128	--	335	--	79
apr	--	340	--	649	--	157
may*	--	377	--	730	--	217
jun*	--	340	--	429	--	187
jul	--	306	--	407	--	136
aug	--	314	--	449	--	141

sep*	--	344	--	457	--	152
oct*	--	289	--	405	--	139
nov*	--	289	--	428	--	149
dec*	--	256	--	389	--	113
1980						
jan	--	353	--	393	--	123
feb	--	336	--	368	--	185
mar	--	105	--	199	--	126
apr	--	302	--	366	--	171
may*	--	289	--	379	--	216
jun*	--	371	--	451	--	292
jul	--	294	--	380	--	214
aug	--	370	--	468	--	319
sep*	--	401	--	406	--	224
oct*	--	339	--	342	--	212
nov*	--	272	--	369	--	146
dec	--	322	--	458	--	184
1981						
jan	--	289	--	394	--	138
feb	--	323	--	373	--	148
mar	--	84	--	155	--	97
apr	--	298	--	342	--	163
may*	--	273	--	343	--	153
jun*	--	253	--	358	--	163
jul	--	229	--	339	--	152
aug	--	196	--	312	--	132
sep*	--	266	--	333	--	177
oct*	--	222	--	298	--	162
nov	--	233	--	316	--	154
dec*	--	236	--	339	--	133
1982						
jan	--	281	--	307	--	214
feb	--	251	--	333	--	213
mar	--	128	--	255	--	579
apr	--	297	--	306	--	198
may	--	312	--	312	--	201
jun	--	376	--	351	--	241
jul	--	304	--	345	--	200
aug	--	349	--	325	--	247
sep	--	348	--	339	--	255
oct	--	250	--	269	--	311
nov	--	360	--	326	--	354
dec	--	314	--	355	--	579
1983						
jan	--	327	--	449	--	347
feb	--	318	--	376	--	225
mar	--	349	--	483	--	686
apr	--	320	--	357	--	298
may	--	336	--	324	--	412
jun	--	392	--	381	--	362
jul	--	324	--	311	--	256
aug	--	346	--	308	--	269
sep	--	401	--	337	--	306
oct	--	528	--	372	--	491

nov	--	336	--	329	--	240
dec	--	345	--	360	--	333
1984						
jan	0	395	9	411	0	345
feb	0	323	8	319	0	260
mar	0	323	10	448	0	316
apr	0	339	8	349	0	313
may	0	402	5	444	0	427
jun	0	365	5	322	0	387
jul	0	448	12	388	0	574
aug	0	421	2	363	0	438
sep	0	431	9	436	0	372
oct	0	538	7	453	0	545
nov	0	416	10	389	0	351
dec	0	407	7	327	0	327
1985						
jan	0	439	8	383	0	329
feb	0	394	8	353	0	305
mar	0	384	13	473	0	411
apr	0	370	7	373	0	324
may	0	472	15	434	0	415
jun	0	478	7	439	0	527
jul	0	365	12	275	0	383
aug	0	412	13	277	0	402
sep	0	389	13	301	0	399
oct	0	346	10	283	0	336
nov	0	353	10	269	0	271
dec	0	326	15	282	0	290
1986						
jan	0	346	10	270	0	356
feb	0	338	10	290	0	340
mar	0	324	5	388	0	449
apr	0	344	9	344	0	376
may	0	319	12	267	0	277
jun	0	431	10	369	0	391
jul	0	335	12	305	0	351
aug	0	264	10	266	0	281
sep	0	273	11	260	0	285
oct	0	346	13	309	0	343
nov	0	303	7	271	0	293
dec	0	260	10	283	0	307
1987						
jan	0	318	8	317	0	379
feb	0	283	10	265	0	276
mar	0	359	16	377	0	426
apr	0	257	10	277	0	259
may	0	278	16	284	0	279
jun	0	389	11	361	0	358
jul	0	355	15	302	0	314
aug	0	268	13	253	0	245
sep	0	290	10	243	0	239
oct	0	598	15	369	0	400
nov^^	0	148946	11	148946	0	148946
dec	0	410	12	324	0	381

1988

jan	0	452	18	364	0	427
feb	0	453	14	293	0	344
mar	0	578	16	410	0	437
apr	0	438	15	360	0	380
may	0	433	12	388	0	405
jun	0	628	16	504	0	531
jul	0	364	10	328	0	341
aug	0	370	17	321	0	406
sep	0	523	9	342	0	475
oct	0	432	8	383	0	368
nov	0	360	13	284	0	278
dec	0	301	18	331	0	300

^ 1977 has multiple versions of demographic variables in months without supplements only.

* Children in months with supplements do not have values for unedited variables and are excluded from the comparison.

^^ Unedited variables are empty in this month.

Table B2a. March Basic-to-ASEC merges using different versions of person identifiers and demographic variables, 1976-1983

	1976			1978			1979			1980			1981			1982			1983		
	unedited	basic edit	new edit	unedited	basic edit	new edit	unedited	basic edit	new edit	unedited	basic edit	new edit	unedited	basic edit	new edit	unedited	basic edit	new edit	unedited	basic edit	new edit
total BMS		97559			113992			113689			133914			133488			153195			153431	
N merged BMS	97503	97511	--	113917	113935	--	113651	113661	--	133875	133889	--	133395	133417	--	152999	153017	--	153130	153201	--
% merged BMS	99.94%	99.95%	--	93.23%	99.95%	--	99.97%	99.98%	--	99.97%	99.98%	--	99.93%	99.95%	--	99.87%	99.88%	--	99.80%	99.85%	--
Stage 1	91693	91988	--	106278	106720	--	110899	113606	--	130804	133828	--	130239	133394	--	148582	152888	--	148005	152869	--
Stage 2	5495	5505	--	7186	7191	--	50	55	--	52	59	--	19	21	--	43	48	--	128	187	--
Stage 3	20	18	--	26	24	--	0	0	--	2	2	--	2	2	--	2	2	--	9	12	--
Stage 4	287	0	--	409	0	--	2698	0	--	3016	0	--	3133	0	--	4368	78	--	4949	128	--
Stage 5	8	0	--	18	0	--	4	0	--	1	0	--	2	0	--	4	1	--	39	5	--
Stage 6	0	0	--	0	0	--	0	0	--	0	0	--	0	0	--	0	0	--	0	0	--
total unmerged BMS	56	48	--	75	57	--	38	28	--	39	25	--	93	71	--	196	178	--	301	230	--
unmerged with hhid in ASEC	44	36	--	47	29	--	28	18	--	24	10	--	43	21	--	163	145	--	239	168	--
hhid does not appear in ASEC		12	--		28	--		10	--		15	--		50	--		33	--		62	--

Table B2b. March Basic-to-ASEC merges using different versions of person identifiers and demographic variables, 1984-1988

	1984			1985			1986			1987			1988		
	unedited	basic edit	new edit	unedited	basic edit	new edit	unedited	basic edit	new edit	unedited	basic edit	new edit	unedited	basic edit	new edit
total BMS		152614			152705			149440			147543			147937	
N merged BMS	152355	152409	152418	152398	152462	152466	147307	147332	147383	145547	145576	145594	144978	144985	145041
% merged BMS	99.83%	99.87%	99.87%	99.80%	99.84%	99.84%	98.57%	98.59%	98.62%	98.65%	98.67%	98.68%	98.00%	98.00%	98.04%
Stage 1	147626	151488	151402	147518	151549	151338	146004	146095	145956	144149	144330	144140	142354	142107	143058
Stage 2	91	147	873	100	179	980	97	184	971	108	179	953	660	632	387
Stage 3	13	12	14	11	8	8	12	10	10	28	26	28	12	12	0
Stage 4	4581	761	120	4698	716	122	1139	431	440	1198	451	457	1541	1233	1246
Stage 5	44	1	9	71	10	18	55	612	6	64	590	16	77	663	12
Stage 6	0	0	0	0	0	0	0	0	0	0	0	0	334	338	338
total unmerged BMS	259	205	196	307	243	239	2133	2108	2057	1996	1967	1949	2959	2952	2896
unmerged with hhid in ASEC	218	164	155	257	193	189	281	256	205	234	205	187	532	525	469
hhid does not appear in ASEC		41			50			1852			1762			2427	

Table B3a. Validated March Basic Monthly-to-ASEC merges by stage, 1976-1983

1976 (N=97559)

Matches by characteristic (unedited)

	Merges by stage	AGE	SEX	RACE	AGE, SEX, and RACE
Stage 1	91693	91693	91693	91693	91693
Stage 2	5495	5495	5495	5495	5495
Stage 3	20	20	20	20	20
Stage 4	287	238	52	265	0
Stage 5	8	4	2	7	0
Stage 6	0	0	0	0	0
N matches	97503	97450	97262	97480	97208
% matches	--	99.95%	99.75%	99.98%	99.70%

Matches by characteristic (basic edit)

	Merges by stage	AGE	SEX	RACE	AGE, SEX, and RACE
Stage 1	91988	91988	91988	91988	91988
Stage 2	5505	5505	5505	5505	5505
Stage 3	18	18	18	18	18
Stage 4	0	0	0	0	0
Stage 5	0	0	0	0	0
Stage 6	0	0	0	0	0
N matches	97511	97511	97511	97511	97511
% matches	--	100.00%	100.00%	100.00%	100.00%

1978 (N=113992)

Matches by characteristic (unedited)

	Merges by stage	AGE	SEX	RACE	AGE, SEX, and RACE
Stage 1	106278	106278	106278	106278	106278
Stage 2	7186	7186	7186	7186	7186
Stage 3	26	26	26	26	26
Stage 4	409	329	93	368	0
Stage 5	18	10	4	12	1
Stage 6	0	0	0	0	0
N matches	113917	113829	113587	113870	113491
% matches	--	99.92%	99.71%	99.96%	99.63%

Matches by characteristic (basic edit)

	Merges by stage	AGE	SEX	RACE	AGE, SEX, and RACE
Stage 1	106720	106720	106720	106720	106720
Stage 2	7191	7191	7191	7191	7191
Stage 3	24	24	24	24	24
Stage 4	0	0	0	0	0
Stage 5	0	0	0	0	0
Stage 6	0	0	0	0	0
N matches	113935	113935	113935	113935	113935
% matches	--	100.00%	100.00%	100.00%	100.00%

1979 (N=113689)

Matches by characteristic (unedited)

	Merges by stage	AGE	SEX	RACE	AGE, SEX, and RACE
Stage 1	110899	110899	110899	110899	110899
Stage 2	50	50	50	50	50
Stage 3	0	0	0	0	0
Stage 4	2698	2573	2367	420	0
Stage 5	4	2	1	1	0
Stage 6	0	0	0	0	0
N matches	113651	113524	113317	111370	110949
% matches	--	99.89%	99.71%	97.99%	97.62%

Matches by characteristic (basic edit)

	Merges by stage	AGE	SEX	RACE	AGE, SEX, and RACE
Stage 1	113606	113606	113606	113606	113606
Stage 2	55	55	55	55	55
Stage 3	0	0	0	0	0
Stage 4	0	0	0	0	0
Stage 5	0	0	0	0	0
Stage 6	0	0	0	0	0
N matches	113661	113661	113661	113661	113661
% matches	--	100.00%	100.00%	100.00%	100.00%

1980 (N=133914)

Matches by characteristic (unedited)

	Merges by stage	AGE	SEX	RACE	AGE, SEX, and RACE
Stage 1	130804	130804	130804	130804	130804
Stage 2	52	52	52	52	52
Stage 3	2	2	2	2	2
Stage 4	3016	2918	2820	271	0
Stage 5	1	0	1	1	0
Stage 6	0	0	0	0	0
N matches	133875	133776	133679	131130	130858
% matches	--	99.93%	99.85%	97.95%	97.75%

Matches by characteristic (basic edit)

	Merges by stage	AGE	SEX	RACE	AGE, SEX, and RACE
Stage 1	133828	133828	133828	133828	133828
Stage 2	59	59	59	59	59
Stage 3	2	2	2	2	2
Stage 4	0	0	0	0	0
Stage 5	0	0	0	0	0
Stage 6	0	0	0	0	0
N matches	133889	133889	133889	133889	133889
% matches	--	100.00%	100.00%	100.00%	100.00%

1981 (N=133488)

Matches by characteristic (unedited)

	Merges by stage	AGE	SEX	RACE	AGE, SEX, and RACE
Stage 1	130239	130239	130239	130239	130239
Stage 2	19	19	19	19	19
Stage 3	2	2	2	2	2
Stage 4	3133	3050	2978	201	0
Stage 5	2	1	2	1	0
Stage 6	0	0	0	0	0
N matches	133395	133311	133240	130462	130260
% matches	--	99.94%	99.88%	97.80%	97.65%

Matches by characteristic (basic edit)

	Merges by stage	AGE	SEX	RACE	AGE, SEX, and RACE
Stage 1	133394	133394	133394	133394	133394
Stage 2	21	21	21	21	21
Stage 3	2	2	2	2	2
Stage 4	0	0	0	0	0
Stage 5	0	0	0	0	0
Stage 6	0	0	0	0	0
N matches	133417	133417	133417	133417	133417
% matches	--	100.00%	100.00%	100.00%	100.00%

1982 (N=153195)

Matches by characteristic (unedited)

	Merges by stage	AGE	SEX	RACE	AGE, SEX, and RACE
Stage 1	148582	148582	148582	148582	148582
Stage 2	43	43	43	43	43
Stage 3	2	2	2	2	2
Stage 4	4368	4242	4115	354	0
Stage 5	4	3	4	0	0
Stage 6	0	0	0	0	0
N matches	152999	152872	152746	148981	148627
% matches	--	99.92%	99.83%	97.37%	97.14%

Matches by characteristic (basic edit)

	Merges by stage	AGE	SEX	RACE	AGE, SEX, and RACE
Stage 1	152888	152888	152888	152888	152888
Stage 2	48	48	48	48	48
Stage 3	2	2	2	2	2
Stage 4	78	35	68	53	0
Stage 5	1	0	1	1	0
Stage 6	0	0	0	0	0
N matches	153017	152973	153007	152992	152938
% matches	--	99.97%	99.99%	99.98%	99.95%

1983 (N=153195)

Matches by characteristic (unedited)

	Merges by stage	AGE	SEX	RACE	AGE, SEX, and RACE
Stage 1	148005	148005	148005	148005	148005
Stage 2	128	128	128	128	128
Stage 3	9	9	9	9	9
Stage 4	4949	4663	4522	634	0
Stage 5	39	3	8	5	0
Stage 6	0	0	0	0	0
N matches	153130	152808	152672	148781	148142
% matches	--	99.79%	99.70%	97.16%	96.74%

Matches by characteristic (basic edit)

	Merges by stage	AGE	SEX	RACE	AGE, SEX, and RACE
Stage 1	152869	152869	152869	152869	152869
Stage 2	187	187	187	187	187
Stage 3	12	12	12	12	12
Stage 4	128	52	87	114	0
Stage 5	5	0	4	5	0
Stage 6	0	0	0	0	0
N matches	153201	153120	153159	153187	153068
% matches	--	99.95%	99.97%	99.99%	99.91%

Table B3b. Validated March Basic Monthly-to-ASEC merges by stage, 1984-1988

1984 (N=152614)

Matches by characteristic (unedited)

	Merges by stage	AGE	SEX	RACE	AGE, SEX, and RACE
Stage 1	147626	147626	147626	147626	147626
Stage 2	91	91	91	91	91
Stage 3	13	13	13	13	13
Stage 4	4581	4320	4178	587	0
Stage 5	44	1	2	1	0
Stage 6	0	0	0	0	0
N matches	152355	152051	151910	148318	147730
% matches	--	99.80%	99.71%	97.35%	96.96%

Matches by characteristic (basic edit)

	Merges by stage	AGE	SEX	RACE	AGE, SEX, and RACE
Stage 1	151488	151488	151488	151488	151488
Stage 2	147	147	147	147	147
Stage 3	12	12	12	12	12
Stage 4	761	37	87	715	0
Stage 5	1	0	1	1	0
Stage 6	0	0	0	0	0
N matches	152409	151684	151735	152363	151647
% matches	--	99.52%	99.56%	99.97%	99.50%

Matches by characteristic (new edit)

	Merges by stage	AGE	SEX	RACE	AGE, SEX, and RACE
Stage 1	151402	151402	151402	151402	151402
Stage 2	873	873	873	873	873
Stage 3	14	14	14	14	14
Stage 4	120	48	80	107	0
Stage 5	9	0	9	9	0
Stage 6	0	0	0	0	0
N matches	152418	152337	152378	152405	152289
% matches	--	99.95%	99.97%	99.99%	99.92%

1985 (N=152614)

Matches by characteristic (unedited)

	Merges by stage	AGE	SEX	RACE	AGE, SEX, and RACE
Stage 1	147518	147518	147518	147518	147518
Stage 2	100	100	100	100	100
Stage 3	11	11	11	11	11
Stage 4	4698	4400	4295	618	0
Stage 5	71	4	11	8	0
Stage 6	0	0	0	0	0
N matches	152398	152033	151935	148255	147629
% matches	--	99.76%	99.70%	97.28%	96.87%

Matches by characteristic (basic edit)

	Merges by stage	AGE	SEX	RACE	AGE, SEX, and RACE
Stage 1	151549	151549	151549	151549	151549
Stage 2	179	179	179	179	179
Stage 3	8	8	8	8	8
Stage 4	716	22	94	692	0
Stage 5	10	1	10	10	1
Stage 6	0	0	0	0	0
N matches	152462	151759	151840	152438	151737
% matches	--	99.54%	99.59%	99.98%	99.52%

Matches by characteristic (new edit)

	Merges by stage	AGE	SEX	RACE	AGE, SEX, and RACE
Stage 1	151338	151338	151338	151338	151338
Stage 2	980	980	980	980	980
Stage 3	8	8	8	8	8
Stage 4	122	36	95	113	0
Stage 5	18	2	17	18	1
Stage 6	0	0	0	0	0
N matches	152466	152364	152438	152457	152327
% matches	--	99.93%	99.98%	99.99%	99.91%

1986 (N=149440)

Matches by characteristic (unedited)

	Merges by stage	AGE	SEX	RACE	AGE, SEX, and RACE
Stage 1	146004	146004	146004	146004	146004
Stage 2	97	97	97	97	97
Stage 3	12	12	12	12	12
Stage 4	1139	617	835	759	0
Stage 5	55	0	5	2	0
Stage 6	0	0	0	0	0
N matches	147307	146730	146953	146874	146113
% matches	--	99.61%	99.76%	99.71%	99.19%

Matches by characteristic (basic edit)

	Merges by stage	AGE	SEX	RACE	AGE, SEX, and RACE
Stage 1	146095	146095	146095	146095	146095
Stage 2	184	184	184	184	184
Stage 3	10	10	10	10	10
Stage 4	431	53	399	409	0
Stage 5	612	0	3	3	0
Stage 6	0	0	0	0	0
N matches	147332	146342	146691	146701	146289
% matches	--	99.33%	99.56%	99.57%	99.29%

Matches by characteristic (new edit)

	Merges by stage	AGE	SEX	RACE	AGE, SEX, and RACE
Stage 1	145956	145956	145956	145956	145956
Stage 2	971	971	971	971	971
Stage 3	10	10	10	10	10
Stage 4	440	59	401	418	0
Stage 5	6	0	6	6	0
Stage 6	0	0	0	0	0
N matches	147383	146996	147344	147361	146937
% matches	--	99.74%	99.97%	99.99%	99.70%

1987 (N=147543)

Matches by characteristic (unedited)

	Merges by stage	AGE	SEX	RACE	AGE, SEX, and RACE
Stage 1	144149	144149	144149	144149	144149
Stage 2	108	108	108	108	108
Stage 3	28	28	28	28	28
Stage 4	1198	613	885	824	0
Stage 5	64	2	7	6	1
Stage 6	0	0	0	0	0
N matches	145547	144900	145177	145115	144286
% matches	--	99.56%	99.75%	99.70%	99.13%

Matches by characteristic (basic edit)

	Merges by stage	AGE	SEX	RACE	AGE, SEX, and RACE
Stage 1	144330	144330	144330	144330	144330
Stage 2	179	179	179	179	179
Stage 3	26	26	26	26	26
Stage 4	451	42	424	436	0
Stage 5	590	0	4	4	0
Stage 6	0	0	0	0	0
N matches	145576	144577	144963	144975	144535
% matches	--	99.31%	99.58%	99.59%	99.28%

Matches by characteristic (new edit)

	Merges by stage	AGE	SEX	RACE	AGE, SEX, and RACE
Stage 1	144140	144140	144140	144140	144140
Stage 2	953	953	953	953	953
Stage 3	28	28	28	28	28
Stage 4	457	58	415	441	0
Stage 5	16	1	15	16	0
Stage 6	0	0	0	0	0
N matches	145594	145180	145551	145578	145121
% matches	--	99.72%	99.97%	99.99%	99.68%

1988 (N=147937)

Matches by characteristic (unedited)

	Merges by stage	AGE	SEX	RACE	AGE, SEX, and RACE
Stage 1	142354	142354	142354	142354	142354
Stage 2	660	660	660	660	660
Stage 3	12	12	12	12	12
Stage 4	1541	513	1081	1171	0
Stage 5	77	7	11	13	2
Stage 6	334	334	334	334	334
N matches	144978	143880	144452	144544	143362
% matches	--	99.24%	99.64%	99.70%	98.89%

Matches by characteristic (basic edit)

	Merges by stage	AGE	SEX	RACE	AGE, SEX, and RACE
Stage 1	142107	142107	142107	142107	142107
Stage 2	632	632	632	632	632
Stage 3	12	12	12	12	12
Stage 4	1233	210	867	1090	0
Stage 5	663	4	17	28	2
Stage 6	338	338	338	338	338
N matches	144985	143303	143973	144207	143091
% matches	--	98.84%	99.30%	99.46%	98.69%

Matches by characteristic (new edit)

	Merges by stage	AGE	SEX	RACE	AGE, SEX, and RACE
Stage 1	143058	143058	143058	143058	143058
Stage 2	387	387	387	387	387
Stage 3	0	0	0	0	0
Stage 4	1246	199	881	1102	0
Stage 5	12	1	7	11	0
Stage 6	338	338	338	338	336
N matches	145041	143983	144671	144896	143781
% matches	--	99.27%	99.74%	99.90%	99.13%

Table B4. Validation of March Basic Monthly-to-ASEC merges by record type, 1984-1988

validated records, basic edit line number and demographics (1984)					validated records, new edit line number and demographics (1984)			
Record type	AGE	SEX	RACE		Record type	AGE	SEX	RACE
Civilians, 15+	116,708	116,708	116,708		Civilians, 15+	116,708	116,698	116,708
Armed forces	0	0	610		Armed forces	642	644	643
Children	34,976	35,027	35,045		Children	34,987	35,036	35,054

validated records, basic edit line number and demographics (1985)					validated records, new edit line number and demographics (1985)			
Record type	AGE	SEX	RACE		Record type	AGE	SEX	RACE
Civilians, 15+	117,232	117,232	117,232		Civilians, 15+	117,232	117,219	117,232
Armed forces	0	0	590		Armed forces	601	608	606
Children	34,527	34,608	34,616		Children	34,531	34,611	34,619

validated records, basic edit line number and demographics (1986)					validated records, new edit line number and demographics (1986)			
Record type	AGE	SEX	RACE		Record type	AGE	SEX	RACE
Civilians, 15+	113,305	113,594	113,594		Civilians, 15+	113,305	113,589	113,594
Armed forces	0	1	0		Armed forces	617	620	619
Children	33,037	33,096	33,107		Children	33,074	33,135	33,148

validated records, basic edit line number and demographics (1987)					validated records, new edit line number and demographics (1987)			
Record type	AGE	SEX	RACE		Record type	AGE	SEX	RACE
Civilians, 15+	112,002	112,327	112,327		Civilians, 15+	112,002	112,311	112,327
Armed forces	0	0	0		Armed forces	590	591	590
Children	32,575	32,636	32,648		Children	32,588	32,649	32,661

validated records, basic edit line number and demographics (1988)					validated records, new edit line number and demographics (1988)			
Record type	AGE	SEX	RACE		Record type	AGE	SEX	RACE
Civilians, 15+	110,945	111,532	111,751		Civilians, 15+	110,971	111,574	111,797
Armed forces	0	0	1		Armed forces	656	656	655
Children	32,324	32,406	32,420		Children	32,342	32,427	32,442

Table B5. Basic Edit, New Demographic Edit, and Unedited Variables Available through IPUMS CPS, 1979-1988

		Basic Edit	New Demographic Edit	Unedited
1976-1983	PID	LINENO	--	UH_LINENOX_1
	AGE	AGE	--	UH_AGEX_1
	SEX	SEX	--	UH_SEXX_1
	RACE	RACE	--	UH_RACEX_1
1984-1988	PID	UH_LINENO_1	LINENO	UH_LINENOX_1
	AGE	UH_AGE_2, UH_AGE_3	AGE	UH_AGEX_1
	SEX	UH_SEX_1	SEX	UH_SEXX_1
	RACE	UH_RACE_1	RACE	UH_RACEX_1

Table C1. Split Households in 1976-1987 ASEC files

Year	All Empty	One Contains People	Both Contain People	Total Split households
1976	6	80	22	108
1977	24	152	22	198
1978	37	112	28	177
1979	54	146	0	200
1980	58	186	2	246
1981	56	154	2	212
1982	40	122	2	164
1983	35	122	4	161
1984	28	177	8	213
1985	50	207	20	277
1986	36	166	12	214
1987	38	281	16	335