

Extending Current Population Survey Linkages: Obstacles and Solutions for Linking Monthly Data from 1976 to 1988, Version 2

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

April 2021

Working Paper No. 2020-02 DOI: https://doi.org/10.18128/IPUMS2020-02.v2

Original Version Available at DOI: https://doi.org/10.18128/IPUMS2020-02

[†]Address correspondence to Sarah Flood: floo0017@umn.edu. Support for this research was funded by IPUMS CPS (NIH R01HD067258).



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 Institute for Social Research and Data Innovation, 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 as well as when linking the March basic monthly data to the Annual Social and Economic (ASEC) Supplement in the 1976-1988 period. We describe strategies to address linking obstacles and document linkage rates.

Keywords: Data integration, linking, panel data, Current Population Survey

*Corresponding author: Sarah Flood, Institute for Social Research and Data Innovation, University of Minnesota, Minneapolis, MN, USA. E-mail: floo0017@umn.edu

1. 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, with microdata available 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.

IPUMS¹ (ipums.org) is a leading disseminator of CPS data, streamlining and simplifying access to these vital social, economic, and demographic data (1). IPUMS CPS (cps.ipums.org) delivers Annual Social and Economic (ASEC) Supplement data from 1962 to present and basic monthly survey (BMS) data from 1976 to present along with nearly all topical supplement data.

¹ IPUMS is an organization that provides census and survey data from around the world integrated across time and space. IPUMS originally stood for Integrated Public Use Microdata Series, but as of 2016, we no longer treat IPUMS as an acronym given the growth in our data collection beyond microdata and access conditions in some instances that limit usage. For additional information about IPUMS, please see https://ipums.org/what-is-ipums.

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. This effort dramatically reduces redundant effort across researchers.

Previously we have documented linkages across months from 1989 forward (2) and between the March BMS and the ASEC (3). These linkages can be replicated using the IPUMSconstructed linking keys CPSIDP and MARBASECIDP, respectively; these linking keys are available via IPUMS CPS. While researchers could use original keys to 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 individuals 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 linkages.

2. CPS Rotation Pattern and Linking

The CPS is a rotating panel household-level survey in which 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 a household appears in the CPS is their first "month-in-sample" (MIS²) 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

² We use MIS to indicate month-in-sample; this information is contained in the IPUMS CPS variable MISH.

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 (4), which are insufficient for creating linkages. While they list the variables needed to link observations over time, the documentation is sparse and provides insufficient detail about how to do the linking. Furthermore, the instructions focus on linking within a set of years when linking keys are stable and provide limited direction on how to bridge changes in linking keys over time. Finally, the documentation (5) 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).

3. CPSID(P)

To overcome barriers related to the rotation pattern and limited guidance for linking CPS data over time, IPUMS created CPSID and CPSIDP. CPSID is a linking key that accounts for changes in original CPS linking keys and consistently identifies households across time. CPSIDP is a person-level linking key that consistently identifies persons across time. CPSID(P)³ is created by applying Census Bureau rules for linking across time. CPSID(P) eliminates the need for every researcher to navigate the complex CPS rotation pattern, changes in linking keys over time, and additional data quality 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). A unique CPSIDP value, based on original linking keys, is assigned to a single individual each time

³ Hereafter, we use CPSID(P) to refer to both CPSID, the household-level identifier, and CPSIDP, the person-level identifier.

they appear in the CPS. If one or more of the individuals in the household move out of the household between CPS interviews, individuals who have moved are not followed. If everyone in the household moves, and new people occupy the dwelling, the new people are interviewed the next time the household is included in the CPS. In 1976 to 1988, a household is identified using a household identification number (HRHHID) and a household number (HUHHNUM).⁴ If an entire household is replaced, the household number (HUHHNUM) value changes. Under these circumstances, IPUMS CPS generates a new CPSID value for the new household (2). For example, a household that was first observed in the CPS in December of 1981 may have the following CPSID value 19811203287400; the two 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.⁵

We encountered several problems when linking CPS respondents in the 1976 to 1988 period. Broadly, these included issues with 1) inconsistent household identifiers, and 2) duplicated person identifiers within households. We describe these problems in detail along with solutions we employed to extend the IPUMS CPS variables CPSID(P) and MARBASECID(P) to basic monthly and ASEC data in the 1976 to 1988 period.

4. Creating CPSID(P) for Basic Monthly Files

⁴ 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 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.

⁵ This is described in detail elsewhere (2).

Each BMS data file contains only individuals interviewed in that specific month. In some months, topical supplement data are also collected. Regardless of whether topical supplement 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 (the supplement-containing file) is released later that includes the topical supplement variables appended to the end of the records. The basic monthly portions of the basic monthly-only and supplement-containing files should be identical, though this is not always the case in the 1976-1988 period. We detail the differences we encounter and our efforts to reconcile differences between the basic monthly-only and supplement-containing files in Appendix A. We also encounter challenges linking across months in the 1976 to 1988. The two major obstacles to linking monthly files are inconsistent household identifiers and duplicated person identifiers.

4a. Problem 1: Household Identifiers and Linking across Months

To address problems with the household identifiers that affected linking across months between 1976 and 1988, we made modifications to the original household identifier prior to creating CPSID(P). HHID refers to the original, unmodified household identifier and HRHHID refers to the IPUMS CPS-modified household identifier used for linking. Except in specified months, HHID and HRHHID are the same.

4a1. 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 within and across months. We used these to create HRHHID for the basic monthly-only files as follows: we used the first nine digits of the twelvedigit HHID and replaced the last three digits of HHID with information from the first two digits of the third and sixth sets of columns marked as "blank" in the original CPS codebook (i.e., BLANK3 and BLANK6⁶).

In 1976 and 1977, HRHHID, constructed as just described, yields more plausible households within basic-monthly only files than the original household identifier (HHID). Specifically, households identified using HHID were extremely large and contained multiple household heads; using HRHHID, results in smaller households with only one head per household. HHID uniquely identifies households in the supplement-containing files; accordingly, we gauge the quality of HRHHID in the basic monthly-only files by making comparisons with the supplement-containing files from the same month (in months that contain supplements). In 1976, HRHHID from the basic monthly-only files matched HHID in the supplement-containing files for 75% of individuals. For the remaining quarter of the records, the final three digits of HRHHID (which correspond to BLANK3 and BLANK6) matched across files, but the first nine digits of HRHHID (which correspond to the first nine digits of HHID) did not match. In these instances, the first nine digits differ systematically between the basic monthly-only and supplement-containing files, so we made a second adjustment to HRHHID (see Appendix A). In 1977, neither HHID nor HRHHID yields successful matches between basic monthly-only and supplement-containing files, so we are unable to check the validity of our constructed HRHHID household identifier as we did for 1976. However, HRHHID in the basic monthly-only files results in each household containing only one household head, which is almost always consistent with the organization of households in the supplement-containing files and represents an improvement over HHID.

4a2. Non-numeric Household Identifiers

⁶ 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.

In January 1976 to June 1985, some records contained non-numeric characters in HHID. We made adjustments to the original 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. 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/HRHHID. The December 1978 file included a '/' in the third position of HRHHID for 48 records. We replaced blank spaces with a '9'; '-' characters with an '8'⁷; and '/' characters with a '4'⁸. This character swap preserves the uniqueness of the household identifier within each month of data.

4a3. 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 (5). We elaborate on these known limitations in the 1970s and remark briefly on our investigation in the 1980s.

Unlinkable months within 1977 and between 1976, 1977, and 1978 are the result of a phased-in sample size increase. Starting with the supplement-containing months of 1977, additional households were included in the CPS, resulting in 15,000 to 25,000 more individuals in these months (known as the D-sample). This sample size increase was extended to include all months in 1978. To prevent identification of these additional individuals in 1977, the Census

⁷ 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.

⁸ We looked for the individuals in adjacent months with problematic HHIDs; based on our analysis, we replaced this character with a 4.

Bureau altered the procedures for generating HHID so that supplement-containing months that included the D-sample could not be linked to basic monthly-only months without it (6).

Despite HRHHID uniquely identifying households within the basic monthly-only files in 1976 to 1978, this identifier does not always yield linkages across months within years or across years. Linkages are possible within 1976 and 1978 between months that are basic monthly-only and supplement-containing (see Figure 1a). This is not the case, however, in 1977 due to the Dsample. 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 link to both 1978 basic monthly-only and supplement-containing files in 1977 basic monthly-only file will link to either basic monthly-only or supplement-containing files in 1978.

[figure 1a][figure 1b]

We find that no month before June of 1985 can be linked to any month after June of 1985. This is consistent with Census Bureau technical documentation, which provides no explanation of why linkages are not possible (5). We suspect that this linking barrier may be due to a CPS redesign that began in April of 1984 and concluded by July of 1985 (7). We find a similar linking discontinuity between September and October of 1985. No month after September of 1985 can be linked to any month before September of 1985. This break in linking is due to a change in Census Bureau confidentiality rules beginning in October of 1985 that allows for the identification of smaller individual metropolitan areas in the public use data (12).

As a result of these two barriers to linking in 1985, July, August, and September of that year can only link to one another and to no other months.

4b. Problem 2: Person Identifiers and Linking across Months

4b1. 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, resulting in duplicate values of the person line number (LINENO) within households. This presents a problem for linking individuals across months using only person (LINENO) and household identifiers (HRHHID and HUHHNUM).

In the basic monthly data, if two or more records in a household have identical person identifiers, we do not allow linkages across months for those individuals. CPSIDP is a linking key based solely on Census Bureau identifiers. Accordingly, we do not use additional information about the individuals to try to 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 with the HRHHID value of 003147831701 that appears in three months of the CPS, beginning in December of 1981. [figure 2]

The first two records in the household in December of 1981, shown in Figure 2, 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 (see Figure 3). 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.

[figure 3]

As of January 1982, the household with an HRHHID value of 003147831701 and an HUHHNUM value of one has appeared in the CPS twice and has four CPSIDP values. When we see the same household in February of 1982 (see Figure 4) 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.

[figure 4]

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 different versions of demographic

variables and the Census Bureau documentation on which version to use for these purposes is unclear (see Appendix B for more details). Importantly, these 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. We use the variable that uniquely identifies persons within households for linking.

4b2. Child Records

Children under 14 are not included in basic monthly-only files until 1982. However, children are included in supplement-containing months prior to 1982: October 1976, October 1977, and May 1978 through December 1981 (see Appendix A). Despite inclusion in the data file, we do not attempt link persons under 14 across months in this period, as they do not appear in all months; they have CPSIDP values of 0.

5. Basic Monthly Survey Linkage Rates Across Time

Linkage rates across months in 1976 to 1988 using HRHHID, HUHHNUM, and LINENO as linking keys are lower than in recent years (2) due to comparatively poor data quality as outlined above in the earlier time period. 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 2 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.

[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 eligible records in January are observed in February and nearly all of them are plausible based on age, sex, and race. Plausible links are those which have the same values for SEX and RACE in all time points and whose AGE does not increase by more than two years.

[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, 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 eligible individuals are observed the next year and just over 70% of these links are plausible based on age, sex, and race. [table 4]

Respondents may also 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 January 1987 who appear in the CPS up to eight times. The individuals who are first observed in January 1976 link well in 1976; 93% are also observed in February 1976 and 85% are also observed in February, March, and April of 1976. However, no individuals who started the CPS in January 1976 are linked to January 1977, as it is a supplement-containing month and such linkages are not possible (as described above). The individuals who begin the CPS in January 1987 link well within 1987; 85.91% are observed in all four months between January and April 1987. Linkage rates decrease after the eight month gap between MIS 4 and MIS 5 (see also 2). About 59% of respondents who started the CPS in January 1987 are observed all eight times. In both 1976 and 1987, most of the linkages we make are plausible based on age, sex, and race.

[table 5]

Tables 6-8 show linkage rates for a variety of different linkage scenarios. Table 6 shows those individuals who entered the CPS in January 1976 and January 1987 (MIS 1) and appear in any of the subsequent seven months that their household could have been in the CPS (either MIS 2 or MIS 3 or MIS 4 or MIS 6 or MIS 7 or MIS 8). Most respondents (95% in 1976 and 97% in 1987) appear in at least two months of the CPS. Table 7 shows the percent of individuals in January 1976 (and 1987) who appear in the CPS in MIS 2-4. About three-quarters of individuals who begin the CPS in January 1976 and January 1987 are observed in at least one additional month between February and April of 1976 and 1987, respectively. Finally, Table 8 shows attrition between MIS 4 and 5.

[table 6]

[table 7]

[table 8]

6. Extending CPSID(P) to the Annual Social and Economic Supplement, 1976-1988

The ability to easily link the Annual Social and Economic (ASEC) Supplement with the CPS BMS creates many research possibilities. Information only available in the ASEC may be combined with multiple data points from the BMS or used in combination with CPS topical supplements. To make these linkages easier, we add CPSID(P) to the ASEC files. Research utilizing these linking keys, for example, combines information on union membership from the monthly data with tax and public benefit receipt from the ASEC (8) and analyzes family income and health insurance from the ASEC along with smoking behavior from the Tobacco Use Supplement (9).

This section of the paper details the creation of MARBASECIDH and MARBASECIDP, hereafter MARBASECID(P), for the 1976 to 1988 period.⁹ MARBASECIDH and MARBASECIDP are IPUMS CPS variables that link the March basic monthly data to ASEC data from the same year and enable the addition of CPSID(P) to the ASEC files. Data quality issues in 1976 to 1988 such as differing numbers of records in the household between the March basic monthly and ASEC files and mismatched or duplicated person identifiers necessitated a methodology distinct from the 1989 forward period, which relied on Census Bureau identifiers only (with a few exceptions). The methodology for creating MARBASECID(P) in 1976 to 1988 is also distinct from our CPSID(P) methodology. Because all records from the March BMS should theoretically appear in the ASEC file for a given year, we are more aggressive in our attempts to create linkages between the March BMS and the ASEC than we are for links across months of the BMS. We describe the problems encountered in attempting to link March BMS and ASEC files in this period, detail the solutions we implemented to generate MARBASECID(P), and compare our methodology with alternatives.

6a. 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 (4). 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

⁹ The details of generating MARBASECID(P) for the 1989 forward period are described elsewhere (2).

for each year between 1976 and 1988 (Panel A), the number that merge to the ASEC (Panel B), and details about unmerged individuals (Panel C). Panel B shows that, except in 1977, over 98% of individuals in the March basic monthly file are merged with the ASEC; the majority are located in merge stage 1, which we describe in more detail below in the "Strategy for Linking March basic monthly to ASEC" section. For those unmerged, we differentiate between individuals whose household (using HRHHID) is or is not in the ASEC. In the first instance, the same HRHHID value is in both the ASEC and the March BMS. Either the household in the ASEC contains fewer persons than the same household in the March BMS or we are unable to confidently match records within a household across files 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 is less than 100 per year 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,¹⁰ and we do not attempt to link the basic monthly and ASEC in these years.

6b. Problem 2: Person Identifiers and Auxiliary Linking Keys

We encounter difficulties uniquely identifying records within files and in matching records across BMS and ASEC files from 1976-1988. This is not the case in the 1989 forward period where HRHHID and LINENO are sufficient to uniquely identify almost all records within and link across March BMS and ASEC files (2).

6b1. Difficulty uniquely identifying records

With the exception of the 1982 and 1983 March BMS files, no March BMS or ASEC files contain records that are complete duplicates between 1976 and 1988. However, even though

¹⁰ This is based on a conversation with staff at the U.S. Census Bureau, not official documentation.

entire records are unique in most of the files, we are often unable to find a single set of variables to uniquely identify all records in households that appear in both the March BMS and the ASEC file for a given year. The 1989-onward method of simply using linking keys to match March BMS and ASEC data is insufficient. Furthermore, we are unable to identify a single set of variables to use as linking keys that allow us to link all March basic monthly records to their ASEC counterparts for all years in the 1976-1988 period.

Consider the following example household (HRHHID = 509037594903) in Figure 5. This household is in the March BMS and the ASEC. 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, only person records with LINENO values of 1 and 2 are uniquely identified using LINENO and demographic characteristics, while those with LINENO values of 3, 4, and 5 appear twice. Duplication is not a problem in this household in the ASEC. Because the records with LINENO 1 and 2 in the March BMS and ASEC file have unique linking keys, we successfully link these records. This same household in the ASEC file contains only one record each with LINENO values of 3, 4, and 5; these records also have the same demographic characteristics as their duplicated March BMS counterparts. However, in this instance, we attempt to link only the first two persons in the household to the ASEC file. Because linking keys and demographic characteristics do not uniquely identify the additional persons in the BMS file and it is unclear which BMS records should link to the ASEC, we do not attempt to link these records. [figure 5]

In other cases, person identifiers (LINENO) do not match across the March BMS and ASEC files and the demographic variables are insufficient for uniquely identifying individuals

when the person identifier is omitted. For example, consider the following household (HRHHID = 202962182216) in 1985 shown in Figure 6. This household contains eight persons in each of the March BMS and ASEC files; LINENO uniquely identifies persons in the March BMS file, but not in the ASEC. Without LINENO as a linking key, the remaining demographic variables do not uniquely identify all records. In this instance, the two 11-year-olds and the two 12-year-olds (LINENO values 7 and 8 and 5 and 6, respectively in the March BMS file and LINENO values 3 and 4 and 0 and 4, respectively in the ASEC file) cannot be unambiguously matched. [figure 6]

6b2. Unmatchable records

There are two ways that records may not be linked even if identifiers are unique in 1976 to 1988. Both occur because LINENO values do not match across BMS and ASEC files, which undermines the utility of LINENO as a linking key.¹¹ For example, in this household from 1976 (HRHHID = 046112033020) shown in Figure 7, the two sixteen year-olds can be uniquely identified within their household within both the March BMS and the ASEC file. Notice, however, that the order of the people in the household varies across the BMS and ASEC and the LINENO values are not consecutive in the BMS. Because of the LINENO mismatch, the individuals aged 16 and 18 will not be matched across the March BMS and ASEC files even though they may well represent the same people.

[figure 7 and figure 7 caption]

Second, some individuals are uniquely identified by demographic variables and line number and appear in the same order within the household, but person identifiers are different across the March BMS and ASEC files. This is another instance where despite having unique

¹¹ 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.

person line numbers linkages across March BMS and ASEC files are not possible. 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.

[figure 8 and figure 8 caption]

7. Strategy for Linking March basic monthly to ASEC

Given the problems detailed above, we apply a multi-stage merging process to maximize linkage rates while minimizing spurious matches. Demographic information is indispensable from 1976 to 1988 for making matches between the March BMS and ASEC, though it is only used in a handful of cases in the 1989-2019 period. Reliance on demographic and other auxiliary information for linking represents a departure from our general approach of creating mechanical links for CPSID(P) and MARBASECID(P). However, the unique challenges during this period necessitate a different approach. We identify matches between the March BMS and ASEC in up to six stages: four in 1976-1981, five in 1982-1987, and six in 1988. We use person identifiers and demographic information as linking keys in some stages and exclude them in other stages.

7a. 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 does not appear in both the BMS and ASEC files (see Table 10). Because there are no persons under the age of 14 in the March BMS from 1976 to 1981, we also exclude individuals under age 14 in the ASEC from linking to the March BMS. Individuals excluded from the matching are assigned non-linking MARBASECIDP values and CPSIDP values of 0 in the ASEC file.¹²

We use original Census variables for linking March BMS to ASEC with three exceptions, HRHHID in the 1976 BMS, Race in the 1988 ASEC, and hours worked in the 1982-1988 BMS. The adjustment made to the 1976 BMS household identifier is described in the "Problem 1: Household Identifiers and Linking across Months" section above. In 1988, race codes are different in 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. This missing value code harmonization is available in the IPUMS CPS variable AHRSWORKT.

7b. Linking Methods

Before each linking stage, we set aside all records for later use that are not uniquely identified by that stage's linking keys (Table 9). After linking, we combine the set-aside records with the unlinked records (separately for 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 linking stage in which a given record was linked is available to researchers via IPUMS CPS in the variable

¹² 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.

MARBASECSTAGE, which is available on both the March BMS and ASEC files from 1976 to 1988.

[table 9]

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

Stage 1: 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 (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 (Table 10).

Stage 3: Stage 3 linking adds the number of hours worked last week (HOURS) to the Stage 2 linking keys. We link a handful of records in this stage (Table 10).

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 (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.¹³ 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 shown in Figure 9 (HRHHID = 80079928707), 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 after stage 4. 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.

[figure 9]

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. Figure 10 shows an example of persons merged between March BMS and ASEC files in stage 6.

[figure 10]

[table 10]

8. Comparing multi- and single-stage linking methods

¹³ 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.

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 than both single-stage linkage approaches and higher quality linkages as we illustrate next.

[table 11]

9. 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 12]

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 13]

10. 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. This documentation along with linking keys in the IPUMS CPS data are intended as a resource for the research community. Our goal is

to facilitate linkages between CPS files and to supplement existing documentation about CPS linking. As part of this effort, we also strive to be transparent about our processes and to provide researchers with the flexibility to retain and drop linkages we have made based on their comfort with the linkages we have created. We have described 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 we encountered in linking across BMS files to create CPSID(P) were with household and person identifiers. Our description details the adjustments we made to household identifiers based on extensive investigation and a series of checks to ensure the quality of the adjustments. Our approach does not overcome the problem of duplicate person identifiers for linking across BMS files.

Despite our efforts, BMS linkage rates in the 1976 to 1988 period are lower than observed in the 1989 forward period (2). Some linkages are completely impossible. Particularly complex is 1977, in which linkages between basic monthly-only and supplement-containing months are not possible, though some 1977 months can be linked to adjacent years. Basic monthly-only months in 1977 may only be linked to 1976 and the supplement-containing months in 1977 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. We also created linkages between March BMS and ASEC records. The creation of MARBASECID(P) allows us to add CPSID(P) to the ASEC and enable researchers to use ASEC data linked to any month 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). Though no documentation clarifies the exact reason, we strongly suspect that this is related to the "D sample" which was added to the CPS for the purpose of measuring the efficacy of a jobs training program (see more information in Appendix A).

Our approach for creating CPSID(P) and MARBASECID(P) in the 1976 to 1988 period deviated from that employed in the 1989 forward period. In general, we followed the rules elaborated previously (2) for creating CPSID(P) 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. Using the same logic in 1976 to 1988 as we did with the more recent data misses many linkages, which reduces the value of being able to use the ASEC in combination with monthly CPS data. We have demonstrated that our multi-stage linking approach is superior to the single-stage approaches. 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

[1] Flood S, King M, Rodgers R, Ruggles S, Warren JR. Integrated Public Use Microdata Series, Current Population Survey: Version 8.0 [dataset]. Minneapolis, MN: IPUMS, 2020. https://doi.org/10.18128/D030.V8.0

[2] Drew JA, Flood S, Warren JR. Making full use of the longitudinal design of the Current Population Survey: Methods for linking records across 16 months. *Journal of Economic and Social Measurement*. 2014; 39(3):121-144.

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

[4] U.S. Census Bureau. Current Population Survey Design and Methodology Technical Paper 66. 2006. Washington, D.C. Government Printing Office.

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

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

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

[8] Sojourner, A., Pacas, J. The Relationship Between Union Membership and Net Fiscal Impact. Industrial Relations. 2019; 58(1): 86-107.

[9] Thrul, J., Riehm, K.E., Cohen, J.E., Alexander, G.C., Vernick, J.S., Mojtabai, R. Tobacco Control Policies and Smoking Cessation Treatment Utilization: A Moderated Mediation Analysis [Preprint]. 2020 [cited 2021 January 15]: [25 p.]. Available from https://www.medrxiv.org/content/10.1101/2020.10.16.20213900v1

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

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

[12] U.S. Census Bureau. Current Population Survey, October 1985: School Enrollment [codebook]. 1986.

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 supplementcontaining 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 we detail some choices we made in handling 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 (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 supplementcontaining 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 supplementcontaining 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 (6, 10). 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 supplementcontaining 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 supplementcontaining 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 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 (11). 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.¹⁴ 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 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

¹⁴ 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 (Table B4). 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 (Table B5).¹⁵

[table_b5]

¹⁵ 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 1C for a breakdown of these categories across years.

[table 1c]

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.

This work was supported by the Minnesota Population Center at the University of Minnesota (grant number P2CH041023) and IPUMS CPS (grant number R01HD067258), both funded through grants from the Eunice Kennedy Shriver National Institute for Child Health and Human Development (NICHD). Thanks to Kari Williams, Matt Sobek, and Rob Warren for their helpful comments.

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

						19	76							19	77	
	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
						19	87							19	88	
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

Note: Table reports unweighted sample sizes for the number of people participating in the CPS in each calendar month, by month-in-sample group.

	Y	ear 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
Retentio	n rate			92.94%	91.13%		93.95%	91.55%

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

Note: Table reports the unweighted number and percentage of CPS repondents in January of one year (the shaded box) who responded to the CPS in February of that year. Under "Year X," entries report the month and year in which respondents were in MIS1. Because of the rotation group structure, not all respondents in January are eligible to respond in February. The column labeled "plausible" omits apparent matches when respondents' sex or race/ethnicity differs or when their age differs implausibly.

	Y	ear X	,	1976			1987	
				Dec	Dec		Dec	Dec
	Oct	Dec	Oct	(All)	(Plausible)	Oct	(All)	(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	75,337	68,149	65 <i>,</i> 935
Retention rate				73.28%	71.75%		90.46%	87.52%

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

Note: Table reports the unweighted number and percentage of CPS repondents in October of one year (the shaded box) who responded to the CPS in December of that year. Under "Year X," entries report the month and year in which respondents were in MIS1. Because of the rotation group structure, not all respondents in October are eligible to respond in December. The column labeled "plausible" omits apparent matches when respondents' sex or race/ethnicity differs or when their age differs implausibly.

	М	arch		March			March	
				1977	1977		1988	1988
	Year X	Year X+	1976	(All)	(Plausible)	1987	(All)	(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
Retentio	n rate			75.64%	73.18%		74.11%	70.63%

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

Note: Table reports the unweighted number and percentage of CPS repondents in March of one year (the shaded box) who responded to the CPS in March of the next year. Under "Year X," entries report the month and year in which respondents were in MIS1. Because of the rotation group structure, not all respondents in March are eligible to respond the following March. The column labeled "plausible" omits apparent matches when respondents' sex or race/ethnicity differs or when their age differs implausibly.

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%
			1987	
	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%

Note: Separately for people entering the CPS in January 1976 and January 1987, the table reports unweighted sample sizes for the number of people participating in all of the CPS surveys for which they were eligible up through the focal month. For example, among the 12,287 people who began the CPS in January of 1976, there were 10,486 (or 85.34%) who participated in all four surveys between January and April 1976 and 0 who participated in all eight surveys between January 1976 and April 1977 due to the linking discontinuity between samples with and without supplements in these years. The column labeled "plausible" omits apparent matches when respondents' sex or race/ethnicity differs or when their age differs implausibly.

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

		1	976	
	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%
		1	987	
	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%

Note: Separately for people entering the CPS in January 1976 and January 1987, the table reports unweighted sample sizes for the number of people participating in ANY of the CPS surveys for which they were eligible up through April of the following year. For example, among the 12,287 people who began the CPS in January of 1976, there were 11,775 (or 95.83%) who participated in at least one more survey between February 1976 and April 1977. The column labeled "plausible" omits apparent matches when respondents' sex or race/ethnicity differs or when their age differs implausibly.

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

		1	976	
	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%
		1	987	
	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%

Note: Separately for people entering the CPS in January 1976 and January 1987, the table reports unweighted samples sizes for the number of people participating in any of the CPS surveys for which they were eligible between January and April of the following year. For example, among the 12,287 people who began the CPS in January of 1976, there were 9,475 (or 77.11%) who participated in at least one more survey between February 1976 and April 1977. The column labeled "plausible" omits apparent matches when respondents' sex or race/ethnicity differs or when their age differs implausibly.

Table 6. Sample size and retention rate, CFS respondents in month-in-sample 4 linked to month-in-sample 5

						Ye	ar X							Year	⁻ Х+1	
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr
MIS4	Oct _{X-1}	Nov _{X-1}	Dec _{X-1}	Jan _x	Feb _x	Mar _x	Apr _x	May _x	Jun _x	Jul _x	Aug _x	Sep _x	Oct _x	Nov _x	Dec _x	Jan _{X+1}
MIS5	Jan _{x-1}	Feb _{x-1}	Mar _{x-1}	Apr _{X-1}	May _{x-1}	Jun _{x-1}	Jul _{x-1}	Aug _{X-1}	Sep _{X-1}	Oct _{X-1}	Nov _{X-1}	Dec _{x-1}	Jan _x	Feb _x	Mar _x	Apr _x
						19	976							19	77	
	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%
						19	987							19	88	
	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%

Note: Table reports the number and percentage of CPS repondents in month-in-sample four who responded to the CPS in month-in-sample five nine months later. Under "Year X," entries report the month and year in which respondents were in MIS1. The rows labeled "plausible" omit apparent matches when respondents' sex or race/ethnicity differs or when their age differs implausibly.

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

	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988
A. Total BMS													
Ν	97559		113992	113689	133914	133488	153195	153431	152614	152705	149440	147543	147937
B. Merged BMS													
Ν	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													
Ν	48		57	28	25	71	178	230	196	239	2057	1949	2896
HHID in ASEC but													
individual is not													
Ν	36		29	18	10	21	145	168	155	189	205	187	469
HHID not in ASEC													
Ν	12		28	10	15	50	33	62	41	50	1852	1762	2427

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

					Single-Stage	
			Single-Stage	Additional	Merge,	Additional
			Merge,	Matches from	Including	Matches from
		Multi-Stage	Identifiers	Multi-Stage	Demographic	Multi-Stage
Year	Total BMS	Merge	Only	Merge	Variables	Merge
1976	97,559	97,511	94,820	2,691	91,988	5,523
1977						
1978	113,992	113,935	110,211	3,724	106,720	7,215
1979	113,689	113,661	110,190	3,471	113,606	55
1980	133,914	133,889	129,438	4,451	133,828	61
1981	133,488	133,417	129,002	4,415	133,394	23
1982	153,195	153,017	149,583	3,434	152,888	129
1983	153,431	153,201	151,444	1,757	152,869	332
1984	152,614	152,418	150,181	2,237	151,402	1,016
1985	152,705	152,466	150,086	2,380	151,338	1,128
1986	149,440	147,383	145,047	2,336	145,956	1,427
1987	147,543	145,594	143,275	2,319	144,140	1,454
1988	147,937	145,041	144,043	998	139,068	5,973

Table 11. Comparison Between Multi Stage and Single Stage Merges, 1976-1988

	Table 12. Col	mparison Between	Multi Stage and	d Single Stage Mer	ges, 1976-1988
--	---------------	------------------	-----------------	--------------------	----------------

			Single-Stage	Merge, Idei	ntifiers Only		Single-Stage Merge, Including D			emographic Variables			Multi-Stage Merge		
		BMS			Valid N	1erges	BMS			Valid N	1erges			Valid N	/lerges
		Duplicates	Successfu	I Merges	(age, se	x, race)	Duplicates	Successfu	l Merges	(age, se	x, race)	Successfu	l Merges	(age, se	x, race)
Year	Total BMS		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 13. Validation on Demographic Characteristics Broken Out by Merge Stage

	Merges							
	by stage	Ma	Matches by Characteristic					
		1982 (N=	=153195)					
					AGE, SEX,			
		AGE	SEX	RACE	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%			

	by stage	Matches by Characteristic				
		1983 (N=	=153431)			
					AGE, SEX,	
		AGE	SEX	RACE	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%	

Merges

	1984 (N=152614)					
					AGE, SEX,	
	_	AGE	SEX	RACE	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)					
					AGE, SEX,	
		AGE	SEX	RACE	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)		
					AGE, SEX,
		AGE	SEX	RACE	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)						
					AGE, SEX,		
		AGE	SEX	RACE	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)					
					AGE, SEX,	
		AGE	SEX	RACE	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%	

1988					
			Supplement-		
		Basic monthly-	containing file		
		only file sets of	sets of	Partial	
		Duplicate	Duplicate	Duplicate	Unmatchable
Year	Month	Records	Records	Records	records*
	January	0	0	0	1
	April	0	0	0	0
1076	May	0	0	0	1
1570	June	0	0	348	0
	October	0	0	0	0
	November	0	0	11	0
	January	na	0	0	na
	May	na	0	0	na
1077	June	na	0	0	na
1311	August	na	0	0	na
	October	na	2	0	na
	December	na	0	0	na
	January	0	0	0	0
	May	0	0	0	0
1978	September	0	0	2	0
	October	0	0	0	0
	November	0	1	0	0
	May	0	8	0	0
	June	0	7	2	0
1070	September	0	7	2	0
1979	October	0	4	2	0
	November	0	1	0	0
	December	0	5	0	0
	May	0	14	0	0
	June	0	8	0	0
1980	September	0	2	0	0
	October	0	4	2	0
	November	0	21	0	0
	January	0	6	0	-
	May	0	3	0	0
1001	June	0	8	0	0
1301	September	0	0	0	0
	October	0	2	2	0
	December	0	3	0	2
	June	3	3	0	1
1007	September	2	1	2	0
1307	October	2	2	2	0
	November	2	2	0	0

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

	January	1	1	0	0
	April	4	4	0	0
	May	2	2	0	2
1002	June	3	3	0	0
1902	September	3	2	2	2
	October	2	1	2	0
	November	3	3	0	0
	December	2	2	0	0
	January	0	0	0	0
	May	0	0	0	0
	June	0	0	0	0
1984	September	0	0	0	0
	October	0	6	2	0
	November	0	0	0	0
	December	0	0	0	0
	April	0	0	0	0
	May	0	0	0	0
1095	June	0	11	6	11
1903	September	0	0	0	6
	October	0	2	2	0
	December	0	0	0	0
	January	0	0	0	0
1986	June	0	0	0	0
1980	October	0	8	0	0
	November	0	0	0	2
	January	0	0	0	0
	June	0	0	0	0
1987	October	0	0	0	13
	November	0	0	0	0
	December	0	68	41	4
	January	0	14	0	1
	May	0	0	6	3
1088	June	0	0	0	4
1900	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

			Number of Reco	ords	Supplement Characteristics			
Year	Month	Basic Monthly	Supplement	Excess Supplement Records	Торіс	Duplicate Record sets	Partial Duplicate Records	Includes Children Under 14
-	January	98,276	108,116	9,840	Tenure	0	0	
	April	98,770	108,696	9,926	Food Security	0	0	
1076	May	98,921	108,958	10,037	Work Schedules	0	0	
1970	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	
	January	99,341	129,031	29,690	Tenure	0	0	
	May	99,667	129,658	29,991	Work Schedules	0	0	
1077	June	99,469	129,401	29,932	Fertility & Marriage	0	0	
1977	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	
	January	128,788	116,421	-12367	Job Tenure	0	0	
	1400	126 676	126,676/	0/	Adult Education /			
1070	iviay	120,070	108,338	-18,338	Work Schedules	0	0	
1978	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
	May	127,376	160,544	33,168	Work Schedules	8	0	Yes
	June	127,660	160,921	33,261	Fertility & Marriage	7	2	Yes
1070	September	128,375	160,442	32,067	Immunization	7	2	Yes
1979	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
	May	135,483	188,729	53,246	Work Schedules	14	0	Yes
	June	134,912	188,201	53,289	Fertility & Marriage	8	0	Yes
1980	September	135,045	187,533	52 <i>,</i> 488	Immunization	2	0	Yes
	October	135,471	188,318	52,847	Education	4	2	Yes
	November	136,122	189,074	52 <i>,</i> 952	Voter	21	0	Yes
	January	135,461	188,764	53 <i>,</i> 303	Job Tenure	6	0	
	May	122,843	170,954	48,111	Work Schedules	3	0	Yes
4004	June	122,524	170,359	47,835	Fertility & Marriage	8	0	Yes
1981	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
	June	169,440	169,436	-4	Fertility & Marriage	0	0	Yes
	September	169,066	169,066	0	Immunization	0	2	Yes
1982	October	169,099	169,098	-1	Education	0	2	Yes
	November	169,324	169,324	0	Voter	0	0	Yes
			·		Job Tenure / Job			
	January	168,124	168,124	0	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
1983	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
					Telephone			
	November	168,189	168,189	0	Availability	0	0	Yes
	December	167,906	167,906	0	Agricultural Worker	0	0	Yes

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

	Januarv	167.985	167.984	-1	Displaced Worker /			Yes
	••••••			_	Job Training	0	0	
	May	165,991	165,991	0	Adult Education	0	0	Yes
	June	166,055	166,055	0	Fertility & Marriage	0	0	Yes
1984	September	165,316	165,316	0	Immunization	0	0	Yes
1901					Education /			
	October	165,490	165,490	0	Computer & Internet			Yes
					Use	6	2	
	November	165,992	165,992	0	Voter	0	0	Yes
	December	166,590	166,590	0	Child Care	0	0	Yes
	Anril	167 213	167 213	0	Veterans / Houshold			Yes
	Abu	107,213	107,213	Ū	Composition	0	0	105
	May	165,498	165,498	0	Work Schedules	0	0	Yes
1025	June	165,471	165,471	0	Fertility & Marriage	22	6	Yes
1985	Contombor	165 627	165 627	0	Tobacco Use /			
	September	105,027	105,027	0	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
	January	165,360	165,359	-1	Displaced Worker	0	0	Yes
1096	June	165,201	165,201	0	Fertility & Marriage	0	0	Yes
1900	October	163,303	161,750	-1,553	Education	4	0	Yes
	November	163,178	163,178	0	Voter	0	0	Yes
	January	162,875	162,875	0	Job Tenure	0	0	Yes
	June	163,541	163,541	0	Fertility & Marriage	0	0	Yes
1987	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
	January	163,357	163,371	14	Displaced Worker	14	0	Yes
	May	152,850	152,850	0	Employee Benefit	0	3	Yes
		·			Fortility & Marriago /			
	June	152,460	152,460	0	Immigration	-	-	Yes
1988						0	0	
	•			-	Retiree Health		c	
	August	152,539	152,539	0	Insurance	0	0	Yes
	October	154,224	154,224	0	Education	0	0	Yes
	November	154,311	154,311	0	Voter	0	0	Yes

Year	Month			Basic					Supplement	t			Supp	lement Sur	plus	
			Non-	Non-				Non-	Non-				Non-	Non-		
		Adults,	Interview	Interview	Armed	Children	Adults,	Interview	Interview	Armed	Children	Adults,	Interview	Interview	Armed	Children
		14+			Forces	ennaren	14+			Forces	erniaren	14+		Type B/C	Forces	erniaren
	lanuany	09 276		туре в/с			09 276	2 027	7 00/				2 027	7 004		
	January	30,270	0	0	-	-	90,270 00 770	2,057	7,004	-	-	0	2,037	7,004	-	-
	April	98,770	0	0	-	-	98,770	2,155	7,771	-	-	0	2,155	7,771	-	-
1976	iviay	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	-	-
	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	-	-
4077	June	99,469	0	0	-	-	116,960	2,493	9,949	-	-	17,491	2,493	9,949	-	-
1977	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	_	_
	January	116 / 21	0	0			116 / 21	0	0			0	0	0		
	May*	11/ 207	2 715	0 75/	_	_	11/ 207	2 715	0 75/	_	_	0	0	0	_	_
1070	lvidy	114,207	2,713	0.047	-	-	120 5 60	2,713	0.047	-	-	22 670	0	0	-	-
1978	September	115,881	2,329	9,847	-	-	139,500	2,329	9,847	-	-	23,079	0	0	-	-
	October	116,//8	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	-	-
	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	-	-
1070	September	115,941	2,326	10,108	-	-	148,008	2,326	10,108	-	-	32,067	0	0	-	-
1979	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	-	-
	Mav	135.483	0	0	-	-	174.070	2.543	12.116	-	-	38.587	2,543	12.116	-	-
	June	134 912	0	0	-	-	173 220	2 817	12 155	-	-	38 317	2,817	12 155	-	-
1920	Sentember	135,015	n	n	_	-	172 000	2,017	11 005	-	_	37 925	2,017	11 905	_	-
1500	Octobor	125,043	0	0	_	-	172,000	2,020	11 0/0	_	-	20 517	2,020	11 040	-	_
	Nevershar	135,471	0	0	-	-	173,988	2,488	11,842	-	-	38,517	2,488	11,842	-	-
	November	136,122	0	0	-	-	1/4,822	2,428	11,824	-	-	38,700	2,428	11,824	-	
	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	-	-
1981	June	122,524	0	0	-	-	156,921	2,546	10,892	-	-	34,397	2,546	10,892	-	-
1901	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	-	-
	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
1982	October	156 526	2 1 2 9	10 444	0	0	156 526	2 1 2 9	10 444	0	0	0	0	0	0	0
	November	156 09/	1 070	10,444	0	0	156 09/	1 070	10,771	0	0	0	0	0	0	0
	lanuary	155,004	2,570	10,370	0	0	155,004	2,570	10,370	0	0		0	0	0	
	January	155,029	2,574	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
1983	June	154,694	2,574	10,747	0	0	154,694	2,574	10,747	0	0	0	0	0	0	0
2000	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
	January	120,303	2,669	11,016	649	33,348	120,303	2,669	11,016	649	33,348	0	0	0	0	0
	Mav	119,296	2.767	10.628	657	32,643	119,296	2.767	10.628	657	32.643	0	0	0	0	0
	lune	119 508	2 812	10 574	680	32 481	119 508	2 812	10 574	680	32 481	0	0	0	0	0
109/	Sentember	110 152	2,012	10 250	606	22,401	110 150	2,012	10 250	606	27 /20	0	n	n	n	0
1904	Octobor	110 405	2,000	10,332	600	22,430	110 405	2,000	10,332	600	22,430	0	0	0	0	0
	Nevershare	110 400	2,331	10,430	009	JZ,427	110,405	2,551	10,438	009	JZ,4Z/	0	0	0	0	0
	November	119,463	2,668	10,724	005	52,472	119,463	2,668	10,724	605	32,472	U	U	U	U	0
	December	120,181	2,588	10,591	647	32,583	120,181	2,588	10,591	64/	32,583	0	0	U	0	0
	April	120,259	2,914	11,181	651	32,208	120,259	2,914	11,181	651	32,208	0	0	0	0	0
	Мау	118,982	2,929	11,058	635	31,894	118,982	2,929	11,058	635	31,894	0	0	0	0	0
1005	June	118,741	2,982	11,299	630	31,819	118,741	2,982	11,299	630	31,819	0	0	0	0	0
1900	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
	Januarv	118.833	2,846	11.154	640	31,887	118.833	2.846	11.154	640	31,887	0	0	0	0	0
	June	118 282	2,888	11.623	669	31.738	118 283	2,888	11.623	669	31,738	0	0	0	0	0
1986	October	117 /59	2,000 2 /122	11 207	637	31 /127	117 /59	2,000 2 /22	11 207	637	31/127	n	n	n	n	n
	November	117 200	2,432 7 E1E	11 010	654	31,402 31 /00	117 200	2,432 2 E1F	11 210	654	21 100	0	0	0	0	0
	19dimevovi	117,290	2,515	11,218	050	51,499	117,290	2,515	11,218	020	31,499	0	0	U	0	0
	January	116,504	2,834	11,458	669	31,410	116,504	2,834	11,458	669	31,410	0	U	U	U	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
1987	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
	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
1988	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	2 200	11 207	654	29 828	110 216	2 200	11 207	654	29 838	n	0	0	n n	n
	November	110,210	2,303	11 046	644	29,000	110,210	2,305	11 046	644	29,000	0	n	n	n	n
		±±0,+JZ	2,002	±±,070	0-1-1		±±0,+02	2,002	±±,0+0	0-1-1	,.0/	0	0	0	0	

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

* 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 k	evs for basic	monthly-on	v and sup	plement-con	taining file merge	1976
			1			

	5 - / · · · · · · · · · · · · · · · · · ·
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

		cige stuge			
					Unlinkable
	Merge	Merge	Merge	Unlinkable	supplement
Month	Stage 1	Stage 2	Stage 3	bms records	records
January	76,041	22,234	0	1	0
April	76,528	22,242	0	0	0
May	76,456	22,283	181	1	0
June	76,477	22,241	174	0	348**
October*	76,962	22,573	185	0	0
November	77,147	22,521	186	0	18**

Table A5. Supplement-containing records merged to basic-monthly only records in 1976 by merge stage

* 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.

	Jan 1976 (N= 98,276)		276)	, Apr 1976	5 (N= 98,7	770)	May 1976 (N= 98,921)			Jun 1976 (N= 99,240)			Oct 1976 (N= 124,360)*			Nov 1976 (N= 99,872)				
	Expected	I																		
	merge																			
Linked	rate		Supplement	Basic	Gain	Supplement	Basic	Gain	Supplement	t Basic	Gain	Supplement	Basic	Gain	Supplement	Basic	Gain	Supplement	Basic	Gain
forward	750/	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
1 month	75%	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%	
forward	50%	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
2 months	50%	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%	
forward	250/	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
3 months	25%	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%	
forward	F.09/	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
months	50%	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%	
forward	2004	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
13 months	38%	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%	
forward	250/	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
months	25%	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%	
forward	12%	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
months	13/0	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%	

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

* 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.

	AGE New		SE	X	RACE			
	New Demographic Unedited		New		New			
	Demographic	Unedited	Demographic	Unedited	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		

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

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	200	Ο	405
lun	0	100	12	300	0	
Juli	0	628	12	504	0	531
Jul	0	628 364	12 16 10	504 328	0	531 341
Jul Aug	0 0 0	628 364 370	12 16 10 17	504 328 321	0 0 0	531 341 406
Jul Aug Sep	0 0 0 0	628 364 370 523	12 16 10 17 9	504 328 321 342	0 0 0 0	531 341 406 475
Jul Aug Sep Oct	0 0 0 0 0	628 364 370 523 432	12 16 10 17 9 8	504 328 321 342 383	0 0 0 0 0	531 341 406 475 368
Jul Aug Sep Oct Nov	0 0 0 0 0 0	628 364 370 523 432 360	12 16 10 17 9 8 13	504 328 321 342 383 284	0 0 0 0 0 0	531 341 406 475 368 278

^ 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	New																		
		Edit	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	1	2		2	8		1	0		1	5		50	0		33	3		6	2	

		108/			1085			1086			1087			1022	
		1504			1303			1500			1587			1,000	
	Unedited	Basic	New												
		Edit	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	I														
ASEC		41			50			1852			1762			2427	

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

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

	Merges by stage	Matche	es by Chara	cteristic (U	nedited)		Merges by stage	Matche	s by Chara	cteristic (Ba	asic Edit)
		19	976 (N=975	59)			.,	1976 (N	=97559)		
					AGE, SEX,						AGE, SEX,
		AGE	SEX	RACE	and RACE			AGE	SEX	RACE	and RACE
Stage 1	91693	91693	91693	91693	91693	Stage 1	91988	91988	91988	91988	91988
Stage 2	5495	5495	5495	5495	5495	Stage 2	5505	5505	5505	5505	5505
Stage 3	20	20	20	20	20	Stage 3	18	18	18	18	18
Stage 4	287	238	52	265	0	Stage 4	0	0	0	0	0
Stage 5	8	4	2	7	0	Stage 5	0	0	0	0	0
Stage 6	0	0	0	0	0	Stage 6	0	0	0	0	0
N matches	97503	97450	97262	97480	97208	N matches	97511	97511	97511	97511	97511
% matches		99.95%	99.75%	99.98%	99.70%	% matches		100.00%	100.00%	100.00%	100.00%
		1978 (N:	=113992)					1978 (N:	=113992)		
					AGE, SEX,						AGE, SEX
		AGE	SEX	RACE	and RACE			AGE	SEX	RACE	and RACE
Stage 1	106278	106278	106278	106278	106278	Stage 1	106720	106720	106720	106720	106720
Stage 2	7186	7186	7186	7186	7186	Stage 2	7191	7191	7191	7191	7191
Stage 3	26	26	26	26	26	Stage 3	24	24	24	24	24
Stage 4	409	329	93	368	0	Stage 4	0	0	0	0	0
Stage 5	18	10	4	12	1	Stage 5	0	0	0	0	0
Stage 6	0	0	0	0	0	Stage 6	0	0	0	0	0
N matches	113917	113829	113587	113870	113491	N matches	113935	113935	113935	113935	113935
% matches		99.92%	99.71%	99.96%	99.63%	% matches		100.00%	100.00%	100.00%	100.00%
		1979 (N:	=113689)					1979 (N:	=113689)		
					AGE, SEX,						AGE, SEX,
		AGE	SEX	RACE	and RACE			AGE	SEX	RACE	and RACE
Stage 1	110899	110899	110899	110899	110899	Stage 1	113606	113606	113606	113606	113606
Stage 2	50	50	50	50	50	Stage 2	55	55	55	55	55
Stage 3	0	0	0	0	0	Stage 3	0	0	0	0	0
Stage 4	2698	2573	2367	420	0	Stage 4	0	0	0	0	0
Stage 5	4	2	1	1	0	Stage 5	0	0	0	0	0
Stage 6	0	0	0	0	0	Stage 6	0	0	0	0	0
N matches	113651	113524	113317	111370	110949	N matches	113661	113661	113661	113661	113661
% matches		99.89%	99.71%	97.99%	97.62%	% matches		100.00%	100.00%	100.00%	100.00%
		1980 (N:	=133914)					1980 (N:	=133914)		
					AGE, SEX,						AGE, SEX
		AGE	SEX	RACE	and RACE			AGE	SEX	RACE	and RACE
Stage 1	130804	130804	130804	130804	130804	Stage 1	133828	133828	133828	133828	133828
Stage 2	52	52	52	52	52	Stage 2	59	59	59	59	59
Stage 3	2	2	2	2	2	Stage 3	2	2	2	2	2
Stage 4	3016	2918	2820	271	0	Stage 4	0	0	0	0	0
Stage 5	1	0	1	1	0	Stage 5	0	0	0	0	0
Stage 6	0	0	0	0	0	Stage 6	0	0	0	0	0
N matches	133875	133776	133679	131130	130858	N matches	133889	133889	133889	133889	133889
% matches		99.93%	99.85%	97.95%	97.75%	% matches		100.00%	100.00%	100.00%	100.00%
		1981 (N:	=133488)					1981 (N:	=133488)		
					AGE, SEX,						AGE, SEX,
		AGE	SEX	RACE	and RACE			AGE	SEX	RACE	and RACE

					AGE, SEX,
		AGE	SEX	RACE	and RACE
Stage 1	133394	133394	133394	133394	133394
Stage 2	21	21	21	21	21
Change 2	2	2	2	2	h

Juage J	2	2	2	2	2	Juge J	2	2	2	2	2
Stage 4	3133	3050	2978	201	0	Stage 4	0	0	0	0	0
Stage 5	2	1	2	1	0	Stage 5	0	0	0	0	0
Stage 6	0	0	0	0	0	Stage 6	0	0	0	0	0
N matches	133395	133311	133240	130462	130260	N matches	133417	133417	133417	133417	133417
% matches		99.94%	99.88%	97.80%	97.65%	% matches		100.00%	100.00%	100.00%	100.00%

		1982 (N=153195)										
-					AGE, SEX,							
		AGE	SEX	RACE	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%							

С

Stage 1 130239

Stage 2

Stage 2

			1982 (N=153195)									
AGE, SEX,	-					AGE, SEX,						
and RACE			AGE	SEX	RACE	and RACE						
148582	Stage 1	152888	152888	152888	152888	152888						
43	Stage 2	48	48	48	48	48						
2	Stage 3	2	2	2	2	2						
0	Stage 4	78	35	68	53	0						
0	Stage 5	1	0	1	1	0						
0	Stage 6	0	0	0	0	0						
148627	N matches	153017	152973	153007	152992	152938						
97.14%	% matches		99.97%	99.99%	99.98%	99.95%						

		1983 (N=	=153195)			_	
-					AGE, SEX,	_	
		AGE	SEX	RACE	and RACE		
Stage 1	148005	148005	148005	148005	148005	Stage 1	152869
Stage 2	128	128	128	128	128	Stage 2	187
Stage 3	9	9	9	9	9	Stage 3	12
Stage 4	4949	4663	4522	634	0	Stage 4	128
Stage 5	39	3	8	5	0	Stage 5	5
Stage 6	0	0	0	0	0	Stage 6	0
N matches	153130	152808	152672	148781	148142	N matches	15320
% matches		99.79%	99.70%	97.16%	96.74%	% matches	

		1983 (N=153195)											
_					AGE, SEX,								
		AGE	SEX	RACE	and RACE								
ge 1	152869	152869	152869	152869	152869								
ge 2	187	187	187	187	187								
ge 3	12	12	12	12	12								
ge 4	128	52	87	114	0								
ge 5	5	0	4	5	0								
ge 6	0	0	0	0	0								
tches	153201	153120	153159	153187	153068								
tches		99.95%	99.97%	99.99%	99.91%								

	Merges by stage Matches by Characteristic (Unedited)			Merges by stage	Matche	s by Charac	teristic (B	asic Edit)	Merges by stage Matches by Characteristic (New Edit)			lew Edit)			
		198	34 (N=1526	14)			198	34 (N=1526	14)			198	84 (N=1526	514)	
					AGE, SEX,					AGE, SEX,					AGE, SEX,
		AGE	SEX	RACE	and RACE		AGE	SEX	RACE	and RACE		AGE	SEX	RACE	and RACE
Stage 1	147626	147626	147626	147626	147626	151488	151488	151488	151488	151488	151402	151402	151402	151402	151402
Stage 2	91	91	91	91	91	147	147	147	147	147	873	873	873	873	873
Stage 3	13	13	13	13	13	12	12	12	12	12	14	14	14	14	14
Stage 4	4581	4320	4178	587	0	761	37	87	715	0	120	48	80	107	0
Stage 5	44	1	2	1	0	1	0	1	1	0	9	0	9	9	0
Stage 6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
N matches	152355	152051	151910	148318	147730	152409	151684	151735	152363	151647	152418	152337	152378	152405	152289
% matches		99.80%	99.71%	97.35%	96.96%		99.52%	99.56%	99.97%	99.50%		99.95%	99.97%	99.99%	99.92%
		4.00		a a)			10		4 A			10			
		198	35 (N=1526	14)			198	35 (N=1526	14)	ACE SEV		198	85 (N=1526	514)	ACE SEV
		ACE	CEV	DACE	and PACE		ACE	CEV	DACE	AGE, SEA,		ACE	CEV	DACE	and PACE
Stage 1	1/7510	1/7E10	147510	147E10	147510	151540	151540	151540	1E1E40	151540	151220	1E1220	1E1220	1E1220	1E1220
Stage 1	14/516	14/516	14/510	14/510	14/518	151549	151549	151549	170	151549	131330	121220	121220	121220	121220
Stage 2	100	100	100	11	11	1/9	0	0	0	0	980	96U 0	96U 0	980 o	96U 0
Stage 3	1600	1100	1205	LL 610	11	0 716	0 22	0/	0 602	0	0 122	0 26	05	0 112	0
Stage 4	4090	4400	4295	010	0	10	1	94 10	10	1	122	30 2	17	115	1
Stage 5	/1	4	0	0	0	10	1	10	10	1	18	2	1/	10	1
N matches	152208	152022	151025	1/12/55	147620	152462	151750	1519/0	152/28	151727	152466	15226/	152/28	152/57	152227
% matches	152550	132033 00 76%	131935 00 70%	1402JJ 07 28%	96 87%	152402	131739 00 5/%	131840 00 50%	00 08%	99 52%	152400	00 03%	00 08%	102407	132327 00 01%
70 matches		55.7070	55.7070	57.2070	50.8770		55.5470	55.5570	55.5070	55.5270		55.5570	55.5070	55.5570	55.5170
		1986 (N=	=149440)				1986 (N=	=149440)				1986 (N=	=149440)		
					AGE, SEX,					AGE, SEX,					AGE, SEX,
<u>.</u>		AGE	SEX	RACE	and RACE		AGE	SEX	RACE	and RACE		AGE	SEX	RACE	and RACE
Stage 1	146004	146004	146004	146004	146004	146095	146095	146095	146095	146095	145956	145956	145956	145956	145956
Stage 2	97	97	97	97	97	184	184	184	184	184	971	971	971	971	971
Stage 3	12	12	12	12	12	10	10	10	10	10	10	10	10	10	10
Stage 4	1139	617	835	759	0	431	53	399	409	0	440	59	401	418	0
Stage 5	55	0	5	2	0	612	0	3	3	0	6	0	6	6	0
Stage 6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
N matches	147307	146730	146953	146874	146113	147332	146342	146691	146701	146289	147383	146996	147344	147361	146937
% matches		99.61%	99.76%	99.71%	99.19%		99.33%	99.56%	99.57%	99.29%		99.74%	99.97%	99.99%	99.70%

		1987 (N=147543)					1987 (N=147543)				1987 (N=147543)				
					AGE, SEX,					AGE, SEX,					AGE, SEX,
		AGE	SEX	RACE	and RACE		AGE	SEX	RACE	and RACE		AGE	SEX	RACE	and RACE
Stage 1	144149	144149	144149	144149	144149	144330	144330	144330	144330	144330	144140	144140	144140	144140	144140
Stage 2	108	108	108	108	108	179	179	179	179	179	953	953	953	953	953
Stage 3	28	28	28	28	28	26	26	26	26	26	28	28	28	28	28
Stage 4	1198	613	885	824	0	451	42	424	436	0	457	58	415	441	0
Stage 5	64	2	7	6	1	590	0	4	4	0	16	1	15	16	0
Stage 6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
N matches	145547	144900	145177	145115	144286	145576	144577	144963	144975	144535	145594	145180	145551	145578	145121
% matches		99.56%	99.75%	99.70%	99.13%		99.31%	99.58%	99.59%	99.28%		99.72%	99.97%	99.99%	99.68%
		1988 (N=	=147937)				1988 (N=	=147937)				1988 (N=	=147937)		
					AGE, SEX,					AGE, SEX,					AGE, SEX,
		AGE	SEX	RACE	and RACE		AGE	SEX	RACE	and RACE		AGE	SEX	RACE	and RACE
Stage 1	142354	142354	142354	142354	142354	142107	142107	142107	142107	142107	143058	143058	143058	143058	143058
Stage 2	660	660	660	660	660	632	632	632	632	632	387	387	387	387	387
Stage 3	12	12	12	12	12	12	12	12	12	12	0	0	0	0	0
Stage 4	1541	513	1081	1171	0	1233	210	867	1090	0	1246	199	881	1102	0
Stage 5	77	7	11	13	2	663	4	17	28	2	12	1	7	11	0
Stage 6	334	334	334	334	334	338	338	338	338	338	338	338	338	338	336
N matches	144978	143880	144452	144544	143362	144985	143303	143973	144207	143091	145041	143983	144671	144896	143781
% matches		99.24%	99.64%	99.70%	98.89%		98.84%	99.30%	99.46%	98.69%		99.27%	99.74%	99.90%	99.13%

Matches by	Charact	eristic (Ba	asic Edit)	Matches I	by Characteristic (New Edit)				
		1984				1984			
	AGE	SEX	RACE		AGE	SEX	RACE		
Civilians, 14+	116,708	116,708	116,708	Civilians, 14+	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		
		1985				1985			
	AGE	SEX	RACE		AGE	SEX	RACE		
Civilians 14+	117 232	117 232	117 232	Civilians 14+	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		
	0.)017	0.,000	0.)010		0.,001	0.)011	0 .)0 _0		
		1986				1986			
	AGE	SEX	RACE		AGE	SEX	RACE		
Civilians, 14+	113,305	113,594	113,594	Civilians, 14+	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		
		1987				1987			
	AGE	SEX	RACE		AGE	SEX	RACE		
Civilians, 14+	112,002	112,327	112,327	Civilians, 14+	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		
		4000				4000			
		1988				1988	DACE.		
0: ::: 4.4	AGE	SEX	RACE		AGE	SEX	RACE		
Civilians, 14+	110,945	111,532	111,751	Civilians, 14+	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 B4. Validation of March	Basic Monthly-to-ASEC merg	ges by record type, 1984-1988
-------------------------------	----------------------------	-------------------------------

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

Table CI. Split Households in 1970-1987 ASEC mes							
		One	Both				
		Contains	Contain	Total Split			
Year	All Empty	People	People	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			

Table C1. Split Households in 1976-1987 ASEC files
Fig. 7. Caption:

This household also contains four records with age values of less than 14. These are excluded from the table and from consideration for linking in 1976-1981, as the March BMS files do not contain records under 14 years of age in those years.

Fig. 8 Caption:

This household also contains three records with age values of less than 14. These are excluded from the table and from consideration for linking in 1976-1981, as the March BMS files do not contain records under 14 years of age in those years.

Fig. 1a. Linkages Possible between File Types within Years, 1976-1978



Fig. 1b. Linkages Possible between File Types across Years, 1976-1978



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

Fig. 2. Household from the December 1981 Current Population Survey File

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

Fig. 3. Household from the January 1982 Current Population Survey File

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

Fig. 4. Household from the February 1982 Current Population Survey File

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

Fig. 5. Household 509037594903 in the 1982 March BMS and ASEC BMS

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	

בועוס						
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

Fig. 6. Household 202962182216 in the 1985 March BMS and ASEC BMS

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	

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	

Fig. 7. Household 046112033020 in the 1976 March BMS and ASEC
50.46

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	

ASEC

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	

Fig. 8. Household 50007137050 in the 1976 March BMS and ASEC BMS

A C	
7.7	

······································										
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					

Fig. 9. March BMS-ASEC merge stages for household 80079928707

										Merge
LI	NENO	AGE		SEX		RACE		HOURS		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

HRHHID=80079928707, HUHHNUM=1

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

Fig. 10. An example of persons linked between March BMS and ASEC in stage 6 Household from March BMS 1988 linked to ASEC in Stage 6

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