

## APPENDIX K: Linking ATUS files to CPS files

One of the strengths of the ATUS is that the data can be linked to data from the Current Population Survey (CPS), a large monthly household survey sponsored by the Bureau of Labor Statistics and conducted by the Census Bureau.

Households selected to participate in the CPS are contacted each month for four months for an interview. They are then contacted again eight months later for another four monthly interviews. For example, a household selected to be interviewed first in April 2002 would be contacted for an interview in April 2002, May 2002, June 2002, July 2002, April 2003, May 2003, June 2003, and July 2003. These monthly interviews are known as the Basic CPS. The month of the last interview is known as the MIS-8 month (or the Month-in-sample 8 month). The CPS rotation scheme is shown in Appendix L.

Additionally, CPS supplement surveys are conducted, usually on a specific subject. These are generally added to the Basic CPS survey for a given month. Some supplements are conducted on a regular basis. For example, the Volunteer Supplement is conducted in September. Some other supplements are conducted on a less regular basis.

ATUS data can be linked to the Basic CPS as well as to CPS supplements. The ATUS files can be linked to the CPS public use files using information on the ATUS-CPS File. There is no one variable that identifies a person or a household uniquely; a number of different variables must be used to link files.

The CPS file format changed in May 2004, and some variables previously used for linking were removed. For this reason, all linking discussed in this document will cover linking prior to May 2004 and linking during or after May 2004. CPS variable names in this document are based on the CPS data dictionaries available at the CPS FTP site ([www.bls.census.gov/ferretftp.htm](http://www.bls.census.gov/ferretftp.htm)).

One thing researchers should keep in mind when linking ATUS data to CPS is that there may be a considerable time difference in when the interviews occurred. The final CPS interview (MIS-8) occurred two to five months prior to the ATUS interview. The fourth CPS interview (MIS-4) occurred a year and two to five months prior to the ATUS interview.

It is possible to link CPS supplements to ATUS data, but the interview dates may be so far apart that researchers may not want to use the information. For example, a CPS household responding to a CPS supplement in their first CPS interview (MIS-1) may not be interviewed for ATUS for another 20 months.

This appendix covers the following topics:

### I. Linking ATUS Files to CPS files

- I.A. Linking the ATUS-CPS file to the final Basic CPS Interview (MIS-8)
- I.B. Linking the ATUS-CPS to a particular Basic CPS month (not necessarily the final interview)
- I.C. Linking the ATUS-CPS file to CPS supplements other than ASEC (March Supplement)
- I.D. Linking the ATUS-CPS file to the ASEC (March Supplement)

### II. Linking CPS files to other CPS files or CPS supplements

- II.A. Linking Basic CPS months to other Basic CPS months or to CPS supplements other than ASEC (March Supplement)
- II.B. Linking two ASEC (March Supplement) files

### III. Common Problems

## I. Linking ATUS files to CPS Files

### I.A. Linking ATUS Public Use Files to the Final Basic CPS Interview (MIS-8)

The ATUS-CPS file contains almost all of the information available on the Basic CPS files, but there are a few exceptions, including geographic variables. To obtain such information, researchers must link the ATUS-CPS file to the final Basic CPS interview. (For a detailed list of variables that are on the Basic CPS files and not on the ATUS-CPS file, please see the ATUS-CPS public use data dictionaries at [www.bls.gov/tus/dictionaries.htm](http://www.bls.gov/tus/dictionaries.htm).)

Because the last CPS interview occurs two to five months prior to the ATUS interview, fifteen months of CPS files—from August to October of the following year—link to a given ATUS-CPS file. For example, CPS files from August 2003 to October 2004 link to the 2004 ATUS-CPS file.

Due to the May 2004 format change in CPS, researchers must use different sets of linking variables, one to link cases whose final CPS interview occurred before May 2004 and one to link cases whose final CPS interview occurred during or after May 2004. The fields HRMONTH and HRYEAR4 identify when the final CPS interview occurred.

#### 1. When the final CPS interview (MIS-8) occurred prior to May 2004

If the following condition is true:

- $\text{HRYEAR4} = 2003 \text{ OR } (\text{HRYEAR4} = 2004 \text{ and } \text{HRMONTH} < 5)$

then use the following variables to link:

- HRHHID
- PULINENO
- HRMONTH
- HRYEAR4
- HRSERSUF

Note that HRSERSUF has different default values in the ATUS-CPS file and the CPS files available for download on the CPS FTP Web site. Failure to account for this will result in many mismatches.

#### 2. When the final CPS interview (MIS-8) occurred during or after May 2004

If the following condition is true:

- $\text{HRYEAR4} > 2004 \text{ or } (\text{HRYEAR4} = 2004 \text{ and } \text{HRMONTH} \geq 5)$

then use the following variables to link:

- HRHHID
- PULINENO
- HRMONTH
- HRYEAR4
- HRHHID2

**Example: Linking the 2004 ATUS-CPS file to Basic CPS (MIS-8)**

Basic CPS files for August 2003 through October 2004 link to the 2004 ATUS. Because of the May 2004 format change, two sets of linking variables are needed.

1. Link the 2004 ATUS-CPS file to Basic CPS months prior to May 2004 using:

- HRHHID
- PULINENO
- HRMONTH
- HRYEAR4
- HRSERSUF

All 2004 ATUS records where  $HRYEAR4 = 2003$  OR ( $HRYEAR4 = 2004$  and  $HRMONTH < 5$ ) will match to a Basic CPS record from August 2003 through April 2004. The ATUS selects only a subset of those who were interviewed for CPS, so there will be many CPS records for which there is no corresponding ATUS record. Delete all records for which there is no match. The remaining records comprise FILE1.

2. Link to CPS months from May 2004 and later using:

- HRHHID
- PULINENO
- HRMONTH
- HRYEAR4
- HRHHID2

All 2004 ATUS records where ( $HRYEAR4 = 2004$  and  $HRMONTH \geq 5$ ) will match to a Basic CPS record from May 2004 through October 2004. The ATUS selects only a subset of those who were interviewed for CPS, so there will be many CPS records for which there is no corresponding ATUS record. Delete all records for which there is no match. The remaining records comprise FILE2.

3. Concatenate FILE1 and FILE2. The remaining file should have exactly the same number of records as does the ATUS-CPS file.

4. To verify that the linking was done correctly, compare the values of age (PRTAGE on the ATUS-CPS file and PEAGE on the Basic CPS), sex (PESEX), and race (PTDTRACE on the ATUS-CPS file and PRDTRACE on the Basic CPS). There should be no differences.

**I.B. Linking the ATUS-CPS File to a particular Basic CPS month (not necessarily the final interview)**

Researchers may wish to know how Basic CPS questions were answered in a given month, especially if they are looking at Supplement responses for that month.

Linking the ATUS-CPS file to a specific non-MIS-8 Basic CPS month requires using different linking variables than does linking to an MIS-8 month.

## 1. When the Basic CPS month occurred prior to May 2004

- Identify all possible MIS-8 months for the particular Basic CPS month. For assistance with this step, refer to Appendix L and the example below.
- If all possible MIS-8 months are prior to May 2004, the Basic CPS file and the ATUS-CPS file can be linked using the following variables:
  - HRHHID
  - PULINENO
  - HUHNUM
  - HRSERSUF
  - HRSAMPLE

Note that HRSERSUF has different default values in the ATUS-CPS file and the CPS files available for download on the CPS FTP Web site. Failure to account for this will result in many mismatches.

- If at least one possible MIS-8 month is May 2004 or later

Several variables must first be defined:

- On the ATUS-CPS file:
  - HUHNUM  
If HRHHID2  $\neq$  -1 then HUHNUM = last digit of HRHHID2  
(The restriction ensures that valid values of HUHNUM are not overwritten)
  - HRSERSUF  
If HRHHID2  $\neq$  -1 then HRSERSUF = digits 3 and 4 of HRHHID2 converted to alpha characters (00 = blank, 01 = A, 02 = B, etc.)  
(The restriction ensures that valid values of HRSERSUF are not overwritten)
  - NumHRSAMPLE  
If HRHHID2  $\neq$  -1 then NumHRSAMPLE = digits 1 and 2 of HRHHID2  
If HRHHID2 = -1 then NumHRSAMPLE = numeric part of HRSAMPLE (characters 2 and 3)
- On the Basic CPS file:
  - NumHRSAMPLE = numeric part of HRSAMPLE (characters 2 and 3)
- Link the Basic CPS to the ATUS-CPS file using the following variables:
  - HRHHID
  - PULINENO
  - HUHNUM
  - HRSERSUF
  - NumHRSAMPLE

Note that HRSERSUF has different default values in the ATUS-CPS file and the CPS files available for download on the CPS FTP website. Failure to account for this will result in many mismatches.

2. When the CPS month is from May 2004 or later, link the Basic CPS file and the ATUS-CPS file using the following variables:

- HRHHID
- PULINENO
- HUUHID2

Example: Linking the 2004 ATUS-CPS file to the 2004 Basic April CPS (not just MIS-8)

1. Since April 2004 is prior to May 2004, first identify all possible MIS-8 months for households receiving the 2004 Basic April CPS.

Refer to Appendix L, and take Year 1 to be 2004. Those households interviewed for April 2004 were D, C, B, A, PP, OO, NN, and MM. Those in MM had an MIS-8 month of April 2004, but all of the other groups had other MIS-8 months. Those in NN, OO, and PP had MIS-8 months of May 2004, June 2004, and July 2004. The other four groups—D, C, B, and A—had MIS-8 months of July 2005, June 2005, May 2005, and April 2005. The 2004 ATUS-CPS file can only be linked to the households that had MIS-8 months in 2004. Therefore, possible MIS-8 months are April 2004, May 2004, June 2004, and July 2004.

Since May 2004, June 2004, and July 2004 are all from May 2004 or later, several variables must be defined before linking can occur.

2. Define three new linking variables on the ATUS-CPS file:

- If HRHHID2  $\neq$  -1 then HUUHNUM = last digit of HRHHID2
- If HRHHID2  $\neq$  -1 then HRSERSUF = digits 3 and 4 of HRHHID2 converted to alpha characters (01 = A, 02 = B, etc.)
- If HRHHID2  $\neq$  -1 then NumHRSAMPLE = digits 1 and 2 of HRHHID2  
If HRHHID2 = -1 then NumHRSAMPLE = numeric part of HRSAMPLE (digits 2 and 3)

3. Define one new linking variable on the Basic April CPS file:

- NumHRSAMPLE= numeric part of HRSAMPLE (digits 2 and 3)

4. Link the Basic April CPS and the ATUS-CPS file using:

- HRHHID
- PULINENO
- HUUHNUM
- HRSERSUF
- NumHRSAMPLE

Many ATUS-CPS records will not have a CPS match. Many CPS records will not have an ATUS-CPS match. Delete all records that do not have a match. The remaining records comprise all members of households that were interviewed for the April 2004 Basic CPS and who were selected to be contacted for an ATUS interview.

The remaining file will have fewer than the total number of records on the ATUS-CPS file with HRYEAR4 = 2004 and HRMONTH = 4, 5, 6, or 7. All records in the ATUS-CPS file with HRYEAR4 = 2004 and HRMONTH = 4—in other words, those records with an MIS-8 month of April—should match to the April 2004 Basic CPS. While every ATUS-CPS record matches to an MIS-8 record, this is not true of other MIS months. A household could have missed its MIS-7 monthly interview, for example, and still have received an MIS-8 monthly interview.

5. To verify that the linking was done correctly, compare values of sex (PESEX) and race (PRDTRACE on the 2004 Basic April CPS and PTDTRACE on the ATUS-CPS file). There should only be a very small number of records that do not match on these two variables.

### I.C. Linking the ATUS-CPS file to CPS Supplements other than ASEC (March Supplement)

Appendix M contains a list of CPS supplements that link to ATUS data.

#### 1. When the CPS Supplement file occurred prior to May 2004

- Identify all possible MIS-8 months that could have received the supplement. For assistance with this step, refer to Appendices L and M and the example below.
- If all possible MIS-8 months are prior to May 2004 then use the following variables to link:
  - HRHHID
  - PULINENO
  - HRSERSUF
  - HUUHNUM
  - HRSAMPLE

Note that HRSERSUF has different default values in the ATUS-CPS file and the CPS files available for download on the CPS FTP website. Failure to account for this will result in many mismatches.

- If at least one possible MIS-8 month is May 2004 or later

Several variables must first be defined:

- On the ATUS-CPS file:
  - HUUHNUM  
If HRHHID2 ≠ -1 then HUUHNUM = last digit of HRHHID2  
(This restriction will ensure that you do not overwrite previous valid values of HUUHNUM)
  - HRSERSUF  
If HRHHID2 ≠ -1 then HRSERSUF = digits 3 and 4 of HRHHID2 converted to alpha characters (00 = blank, 01 = A, 02 = B, etc.)  
(This restriction will ensure that you do not overwrite previous valid values of HRSERSUF)
  - NumHRSAMPLE  
If HRHHID2 ≠ -1 then NumHRSAMPLE = digits 1 and 2 of HRHHID2  
If HRHHID2 = -1 then NumHRSAMPLE = numeric part of HRSAMPLE (characters 2 and 3)
- On the Basic CPS file:
  - NumHRSAMPLE = numeric part of HRSAMPLE (characters 2 and 3)

- Link the Basic CPS to the ATUS-CPS file using the following variables:
  - HRHHID
  - PULINENO
  - HUUHNUM
  - HRSERSUF
  - NumHRSAMPLE

Note that HRSERSUF has different default values in the ATUS-CPS file and the CPS files available for download on the CPS FTP website. Failure to account for this will result in many mismatches.

2. If the CPS Supplement is from May 2004 or later, link the CPS Supplement and the ATUS-CPS file using the following variables:

- HRHHID
- HRHHID2
- PULINENO

Example: Linking the 2003, 2004, and 2005 ATUS-CPS files to the October 2003 Internet and Computer Use Supplement

1. Since October 2003 is prior to May 2004, first identify all possible MIS-8 months that could have received the October 2003 Internet and Computer Use Supplement.

Refer to Appendix L, and take Year 1 to be 2003. Households interviewed in October 2003 were in J, I, H, G, VV, UU, TT, and SS. Those in SS had October 2003 as their MIS-8 month, but all of the other groups had other MIS-8 months. Those in TT, UU, and VV had MIS-8 months of November 2003, December 2003, and January 2004. The other four groups - J, I, H, and G - had MIS-8 months of January 2004, December 2004, November 2004, and October 2004. The 2003 and 2004 ATUS-CPS files can only be linked to the households that had their MIS-8 month in 2003 and 2004. Therefore, possible MIS-8 months are November 2003, December 2003, January 2004, October 2004, November 2004, and December 2004.

Since October 2004, November 2004, and December 2004 are all later than May 2004, several variables must be defined before linking can occur.

2. Define three new linking variables on the 2003-2005 ATUS-CPS file:

- If HRHHID2 ≠ -1 or missing then HUUHNUM = last digit of HRHHID2
- If HRHHID2 ≠ -1 or missing then HRSERSUF = digits 3 and 4 of HRHHID2 converted to alpha characters (01 = A, 02 = B, etc.)
- If HRHHID2 ≠ -1 or missing then NumHRSAMPLE = digits 1 and 2 of HRHHID2  
If HRHHID2 = -1 or missing then NumHRSAMPLE = numeric part of HRSAMPLE (digits 2 and 3)

3. Define one new linking variable on the October 2003 Internet and Computer Use Supplement file:

- NumHRSAMPLE= numeric part of HRSAMPLE (digits 2 and 3)

4. Link the October 2003 Internet and Computer Use Supplement and the 2003-2005 ATUS-CPS file using:

- HRHHID
- PULINENO
- HUHHNUM
- HRSERSUF
- NumHRSAMPLE

Many ATUS-CPS records will not have a match to the October 2003 Internet and Computer Use Supplement. Many October 2003 Internet and Computer Use Supplement records will not have a match to the ATUS-CPS file. Delete all records that do not have a match. The remaining records comprise all members of households that were interviewed for the October 2003 Internet and Computer Use Supplement and who were selected to be contacted for an ATUS interview.

The remaining file will have fewer than the total number of records on the ATUS-CPS file with (HRYEAR = 2004 and HRMONTH = 1, 10, 11, 12) or (HRYEAR4 = 2003 and HRMONTH = 10, 11, or 12). This is because not everyone who could have been interviewed for the October 2003 Internet and Computer Use Supplement actually was interviewed.

5. To verify that the linking process worked correctly, compare the values of sex (PESEX) and race (PTDTRACE on the ATUS-CPS file and PRDTRACE on the October 2003 Internet and Computer Use Supplement). There should be only a very small number of records that do not match on these two variables.

#### I.D. Linking the ATUS-CPS file to the CPS ASEC (March Supplement)

Linking to the ASEC (March Supplement) is particularly difficult for a variety of reasons. Two variables particularly helpful for linking, HRSERSUF and HRSAMPLE, are not included on the ASEC file prior to 2005. In order to increase the sample size, CPS includes two oversamples - known as the SCHIPS and Hispanic oversamples - in the ASEC. The 2003 and 2004 ATUS do not include the CPS SCHIPS or Hispanic oversamples, and steps must be taken to exclude these when matching ATUS files to ASEC files. Because of the complexity of the linking, there are directions on how to link to each year. Note in particular that the value of H-SEQ is different in 2003 and 2004.

##### Linking 2003 ATUS-CPS to 2003 CPS ASEC (March Supplement)

1. Restrict the 2003 ASEC to:

- $5 \leq \text{H-MIS} \leq 8$  AND
- $\text{H-SEQ} \leq 78864$  AND
- $\text{A-FNLWGT} > 0$

The first restriction will limit the ASEC file to the more recent of the two interviews (most households receive the ASEC interview twice). Since the linking variables available on the ASEC file do not uniquely identify households, it also reduces the likelihood of false matches. The last two restrictions ensure that the SCHIPS and Hispanic oversamples are removed; without this step, there will be duplicate household identification and person line numbers.

2. Restrict the 2003 ATUS-CPS file to:

- $\text{HRYEAR4} = 2003$  AND
- $\text{HRMONTH} = 3, 4, 5, \text{ or } 6$



Since the linking variables available on the ASEC file do not uniquely identify households, this step reduces the likelihood of false matches.

3. Link the 2003 ATUS-CPS file and the 2003 ASEC using:

- Household identification number
  - HRHHID on the ATUS-CPS file
  - H-IDNUM on the ASEC file
- Person line number
  - PULINENO on the ATUS-CPS file
  - A-LINENO on the ASEC file
- Household number
  - HUHNUM on the ATUS-CPS file
  - H-HHNUM on the ASEC file

4. After the linking, there will be a small number of duplicates and mismatches. This is because the linking variables available on the ASEC file do not uniquely identify households. Use the following variables to eliminate mismatches and duplicates:

- Age
  - PRTAGE on the ATUS-CPS file
  - A-AGE on the ASEC file

Caution: the information on the ATUS-CPS file may have been collected at a later date than the ASEC file, and so age may be greater
- Sex
  - PESEX on the ATUS-CPS file
  - A-SEX on the ASEC file

In 2003, race is not a good screening variable because new definitions were introduced and it is missing for much of the ATUS-CPS file.

#### Linking 2004 ATUS-CPS to 2004 CPS ASEC (March Supplement)

1. Restrict the 2004 ASEC to:

- $5 \leq H-MIS \leq 8$  AND
- $H-SEQ \leq 78575$  AND
- $A-FNLWGT > 0$

The first restriction will limit the ASEC file to the more recent of the two interviews (each household receives the ASEC questions twice). Since the linking variables available on the ASEC file do not uniquely identify households, it also reduces the likelihood of false matches. The last two restrictions ensure that the SCHIPS and Hispanic oversamples are removed; without this step, there will be duplicate household ID and line numbers.

2. Restrict the 2004 ATUS-CPS file to:

- $HRYEAR4 = 2004$  AND
- $HRMONTH = 3, 4, 5, \text{ or } 6$

Since the linking variables available on the ASEC file do not uniquely identify households, this step reduces the likelihood of false matches.

3. Define HUUHNUM for HRMONTH  $\geq$  5

- If HRMONTH  $\geq$  5 THEN HUUHNUM = last digit of HRHHID2

The variable HUUHNUM is missing (-1) on the ATUS-CPS file for records with HRYEAR4 = 2004 and HRMONTH  $\geq$  5 and must be recreated in order to link properly.

## 4. Link the 2004 ATUS-CPS file and the 2004 ASEC using:

- Household identification number
  - HRHHID on the ATUS-CPS file
  - H-IDNUM on the ASEC file
- Person line number
  - PULINENO on the ATUS-CPS file
  - A-LINENO on the ASEC file
- Household number
  - HUUHNUM on the ATUS-CPS file
  - H-HHNUM on the ASEC file

## 5. After the linking, there will be a small number of duplicates and mismatches. This is because the linking variables available on the ASEC file do not uniquely identify households. Use the following variables to eliminate mismatches and duplicates:

- Age
  - PRTAGE on the ATUS-CPS file
  - A-AGE on the ASEC file

Caution: the maximum value for age changed in May 2004, and so some age mismatches may not really be mismatches. Some of those over age 85 are coded as 80 on one file and 85 on the other. Also, remember that the information on the ATUS-CPS file may have been collected at a later date than the ASEC file, and so age may be greater.
- Sex
  - PESEX on the ATUS-CPS file
  - A-SEX on the ASEC file
- Race
  - PTDTRACE on the ATUS-CPS file
  - PRDTRACE on ASEC file

## Linking 2005 ATUS-CPS to 2005 CPS ASEC (March Supplement)

## 1. Restrict the 2005 ASEC to:

- $5 \leq \text{H-MIS} \leq 8$  AND
- $\text{A-FNLWGT} > 0$

The first restriction will limit the ASEC file to the more recent of the two interviews (each household receives the ASEC questions twice). Since the linking variables available on the ASEC file do not uniquely identify households, it also reduces the likelihood of false matches. The second restriction ensures that the Hispanic oversample is removed.

## 2. Restrict the 2005 ATUS-CPS file to:

- $\text{HRYEAR4} = 2005$  AND
- $\text{HRMONTH} = 3, 4, 5, \text{ or } 6$

Since the linking variables available on the ASEC file do not uniquely identify households, this step reduces the likelihood of false matches.

3. Link the 2005 ATUS-CPS file and the 2005 ASEC using:

- Household identification number (part 1)
  - HRHHID on the ATUS-CPS file
  - H-IDNUM1 on the ASEC file
- Household identification number (part 2)
  - HRHHID on the ATUS-CPS file
  - H-IDNUM2 on the ASEC file
- Person line number
  - PULINENO on the ATUS-CPS file
  - A-LINENO on the ASEC file

Linking 2006 ATUS-CPS to 2006 CPS ASEC (March Supplement)

Follow the above directions for linking the 2006 ATUS-CPS file to the 2006 CPS ASEC. Update the year (HRYEAR4) from 2005 to 2006.

## II. Linking CPS files to other CPS files or supplements

Sometimes ATUS data users have questions related to linking CPS data not to ATUS files, but to other CPS files. The following instructions are included for those who are trying to link only CPS files and not ATUS files. These instructions cover how to link Basic CPS to other months of Basic CPS, Basic CPS to CPS Supplements, Basic CPS to the ASEC (March Supplement), and linking two ASEC files.

### II.A. Linking Basic CPS Months to other Basic CPS Months or to CPS Supplements (other than ASEC/March Supplement)

1. When both CPS files occurred prior to May 2004, then use the following variables to link:
  - HRHHID
  - PULINENO
  - HRSERSUF
  - HUHNUM
  - HRSAMPLE
2. If one file is prior to May 2004 and the other is May 2004 or later, then several variables must be defined:
  - on the later CPS file (May 2004 or later):
    - HUHNUM = last digit of HRHHID2
    - HRSERSUF = digits 3 and 4 of HRHHID2 converted to alpha characters (00 = -1 or default value, 01 = A, 02 = B, etc.)
    - NumHRSAMPLE = digits 1 and 2 of HRHHID2
  - on the earlier CPS file (prior to May 2004):
    - NumHRSAMPLE = numeric part of HRSAMPLE (characters 2 and 3)
  - Use the following variables to link:
    - HRHHID
    - PULINENO
    - HUHNUM
    - HRSERSUF
    - NumHRSAMPLE
3. If both files are May 2004 or later, use the following variables to link:
  - HRHHID
  - HRHHID2
  - PULINENO

Example: Linking the 2004 Basic April CPS file to the 2005 Basic April CPS

1. Since April 2004 is prior to May 2004 and April 2005 is after May 2004, several variables must be defined before linking can occur.

2. Define three new linking variables on the 2005 Basic April CPS file:

- HUHNUM = last digit of HRHHID2
- HRSERSUF = digits 3 and 4 of HRHHID2 converted to alpha characters (00 = -1, 01 = A, 02 = B, etc.)
- NumHRSAMPLE = digits 1 and 2 of HRHHID2

3. Define one new linking variable on the 2004 Basic April CPS file:

- NumHRSAMPLE= numeric part of HRSAMPLE (digits 2 and 3)

4. Link the 2004 Basic April CPS file and the 2005 Basic April CPS file using:

- HRHHID
- PULINENO
- HUHNUM
- HRSERSUF
- NumHRSAMPLE

Many records will not match. This is due to the fact that many of those who were interviewed for the 2004 Basic April CPS were not interviewed for the 2005 Basic April CPS. Delete all records that do not have a match.

5. To verify that the linking was done correctly, compare values of sex (PESEX) and race (PRDTRACE). There should only be a very small number of records that do not match on these two variables.

## II.B. Linking two ASEC (March Supplement) files

The ASEC file is the most difficult of the CPS supplements to link. This is due to a number of factors:

- The linking variables of HRSERSUF and HRSAMPLE are not on the ASEC prior to 2005 (from 2005 onward, the presence of the variable HRHHID2, which incorporates the information that was formerly in HRSERSUF and HRSAMPLE, makes linking much easier). The linking variables that remain - HRHHID, HUHNUM, and PULINENO - do not uniquely identify individuals or households. This results in some incorrect matches.
- In order to increase sample size, the ASEC includes two oversamples—known as the SCHIPS and the Hispanic oversamples. When the oversampled records were included in the ASEC file, the variable identifying the month was changed to 3 to protect confidentiality. This means that—prior to 2005—these oversampled cases could not be identified.
- Some oversampled individuals are interviewed only once for ASEC. These are individuals who are oversampled because of their unusual demographic characteristics (for example, mixed race households with children). As they are interviewed only once for the ASEC, there are no matches for these cases.

## Linking the 2003 ASEC to the 2004 ASEC

1. Link the 2003 Basic March CPS to the 2003 ASEC (March Supplement) using:

- HRHHID
- PULINENO
- HUHNUM

Screen for duplicates and mismatches. These occur because the linking variables available on the ASEC file do not uniquely identify households. Use the following variables to eliminate mismatches and duplicates:

- Age
  - PEAGE on the Basic March CPS file
  - A-AGE on the ASEC file
- Sex
  - PESEX on the Basic March CPS file
  - A-SEX on the ASEC file
- Race
  - PRDTRACE on Basic March CPS file and ASEC file

2. Link the 2004 Basic March CPS to the 2004 ASEC (March Supplement) using:

- HRHHID
- PULINENO
- HUHNUM

Screen for duplicates and mismatches. These occur because the linking variables available on the ASEC file do not uniquely identify households. Use the following variables to eliminate mismatches and duplicates:

- Age
  - PEAGE on the Basic March CPS file
  - A-AGE on the ASEC file
- Sex
  - PESEX on the Basic March CPS file
  - A-SEX on the ASEC file
- Race
  - PRDTRACE on Basic March CPS file and ASEC file

3. Link the two merged Basic-ASEC files using the following variables:

- HRHHID
- PULINENO
- HUHNUM
- HRSAMPLE
- HRSERSUF

Many 2003 ASEC records will not link to 2004 ASEC records. Delete all records that do not have matches.

This method will also work for linking the 2002 ASEC to the 2003 ASEC.

## Linking the 2004 ASEC to the 2005 ASEC

## 1. Link the 2004 Basic March to the 2004 ASEC using:

- HRHHID
- PULINENO
- HUUHNUM

Screen for duplicates and mismatches. These occur because the linking variables available on the ASEC file do not uniquely identify households. Use the following variables to eliminate mismatches and duplicates:

- Age
  - PEAGE on the Basic March CPS file
  - A-AGE on the ASEC file
- Sex
  - PESEX on the Basic March CPS file
  - A-SEX on the ASEC file
- Race
  - PRDTRACE on Basic March CPS file and ASEC file

## 2. On the merged 2004 Basic-ASEC file define:

- NumHRSAMPLE = Numeric part of HRSAMPLE (positions 2 and 3)
- NumHRSERSUF = convert alpha character into a numeric (A=1, B=2, etc.)

## 3. On the 2005 ASEC define:

- NumHRSAMPLE = Digits 1 and 2 of HRHHID2
- NumHRSERSUF = Digits 3 and 4 of HRHHID2

## 4. Link the merged 2004 Basic-ASEC file and the 2005 ASEC using the following variables:

- HRHHID
- PULINENO
- HUUHNUM
- NumHRSAMPLE
- NumHRSERSUF

Many 2004 ASEC records will not link to 2005 ASEC records. Delete all records that do not have matches.

## Linking the 2005 ASEC to the 2006 ASEC

## 1. Link the 2005 ASEC to the 2006 ASEC using the following variables:

- HRHHID
- HRHHID2
- PULINENO

Many 2005 ASEC records will not link to 2006 ASEC records. Delete all records that do not have matches.

(Note that the variable H-MIS, which indicates the record's month-in-sample, is not always accurate on the ASEC file. Records that are given the ASEC because they are part of a CPS oversample may have incorrect values of H-MIS.)

---

**Common Problems**

---

Linking can be a frustrating and lengthy process. Researchers who are having trouble linking should examine four things in particular:

1. **Variable type and length**  
Sometimes files obtained from different places will have different types and/or lengths for linking variables. For example, one file may have HRHHID as a numeric field and another may have it as a character field. Some software packages will not allow you to merge variables of the same name that have different characteristics or lengths.
2. **Leading zeros**  
Some sources of the data have leading zeros on certain fields and some do not. For example, some versions of the CPS files have leading zeros on HRHHID and others don't.
3. **Variable defaults**  
Some sources of the data have different default values for missing variables. For example, the default value of HRSERSUF appears in some datasets as -1 and in others as a blank.
4. **Variable names**  
Some sources of the data have different variable names for the same variable. Variable names in this document are based on the data dictionaries on the CPS FTP site (<http://www.bls.census.gov/ferretftp.htm>).



APPENDIX L: Basic CPS Rotation

This appendix is intended as an aid for researchers who wish to link ATUS and CPS data, but the CPS sample rotation scheme is complicated and is not fully described in this appendix. More detailed information about the CPS rotation and sample can be found in CPS Technical Paper 66 (<http://www.census.gov/prod/2006pubs/tp-66.pdf>).

The CPS rotation scheme follows a 4-8-4 pattern. A given household is in the sample for four consecutive months, not in the sample for the next eight consecutive months, and in the sample again for the next four consecutive months.

The following table illustrates the CPS rotation scheme. Each group - represented by a letter or a letter/number combination - is interviewed for eight months total. For example, a household first interviewed in May (month-in-sample 1 or MIS-1) would be interviewed again in June (MIS-2), July (MIS-3), and August (MIS-4). The household would then be out of the sample for eight months and would not be interviewed again until May of the following year (MIS-5). The final three interviews would occur in June (MIS-6), July (MIS-7), and August (MIS-8) of the following year.

		Month-in-sample (MIS)							
		MIS-1	MIS-2	MIS-3	MIS-4	MIS-5	MIS-6	MIS-7	MIS-8
Year 1	January	A	XX	WW	VV	MM	L1	K1	J1
	February	B	A	XX	WW	NN	MM	L1	K1
	March	C	B	A	XX	OO	NN	MM	L1
	April	D	C	B	A	PP	OO	NN	MM
	May	E	D	C	B	QQ	PP	OO	NN
	June	F	E	D	C	RR	QQ	PP	OO
	July	G	F	E	D	SS	RR	QQ	PP
	August	H	G	F	E	TT	SS	RR	QQ
	September	I	H	G	F	UU	TT	SS	RR
	October	J	I	H	G	VV	UU	TT	SS
	November	K	J	I	H	WW	VV	UU	TT
	December	L	K	J	I	XX	WW	VV	UU
Year 2	January	M	L	K	J	A	XX	WW	VV
	February	N	M	L	K	B	A	XX	WW
	March	O	N	M	L	C	B	A	XX
	April	P	O	N	M	D	C	B	A
	May	Q	P	O	N	E	D	C	B
	June	R	Q	P	O	F	E	D	C
	July	S	R	Q	P	G	F	E	D
	August	T	S	R	Q	H	G	F	E
	September	U	T	S	R	I	H	G	F
	October	V	U	T	S	J	I	H	G
	November	W	V	U	T	K	J	I	H
	December	X	W	V	U	L	K	J	I

## APPENDIX M: Basic CPS files and CPS Supplements

This appendix contains a list of CPS supplements that can be linked to ATUS files. It also contains information about where to obtain these data files.

---

### CPS Supplements that link to ATUS files

---

#### 2006

- September 2006: Volunteer Supplement
- June 2006: Fertility Supplement File
- March/April Match File: Child Support
- March 2006: Annual Social and Economic (ASEC) Supplement
- January 2006: Displaced Worker, Employee Tenure and Occupational Mobility Supplement File

#### 2005

- December 2005: Food Security Supplement
- October 2005: School Enrollment Supplement File
- September 2005: Volunteer Supplement File
- March 2005: Annual Social and Economic (ASEC) Supplement
- February 2005: Contingent Work Supplement File

#### 2004

- 2004: Cell Phone Supplement File
- December 2004: Food Security Supplement File
- November 2004: Voting and Registration Supplement File
- October 2004: School Enrollment, Language Proficiency and Disability Supplement File
- September 2004: Volunteer Supplement File
- June 2004: Fertility Supplement File
- May 2004: Work Schedules and Work at Home Supplement
- March 2004: Annual Social and Economic (ASEC) Supplement
- January 2004: Displaced Workers, Employee Tenure, and Occupational Mobility Supplement File

#### 2003

- December 2003: Food Security Supplement File
- October 2003: School Enrollment and Computer Use Supplement File
- September 2003: Volunteer Supplement File
- August 2003: Veterans Supplement File
- March 2003: Annual Social and Economic (ASEC) Supplement
- February, June, and November 2003: Tobacco Use Supplement File

#### 2002

- December 2002: Food Security Supplement File
- November 2002: Voting and Registration Supplement File
- October 2002: School Enrollment/Library Use Supplement File
- September 2002: Volunteer Supplement File

- August 2002: Public Participation in the Arts Supplement File
- June 2002: Fertility
- March/April 2002 Match File: Child Support
- March 2002: Annual Demographic File
- January 2002: Displaced Worker, Employee Tenure, and Occupational Mobility Supplement File

---

### Obtaining CPS Data and Documentation

---

#### Basic CPS data:

FTP Download Page: [www.bls.census.gov/ferretftp.htm](http://www.bls.census.gov/ferretftp.htm)

DataFerrett: <http://dataferrett.census.gov/>

#### Basic CPS documentation:

FTP Download Page: [www.bls.census.gov/ferretftp.htm](http://www.bls.census.gov/ferretftp.htm)

DataFerrett: <http://dataferrett.census.gov/>

CPS Design and Methodology, Technical Paper 66: [www.census.gov/prod/2006pubs/tp-66.pdf](http://www.census.gov/prod/2006pubs/tp-66.pdf)

#### CPS ASEC (March Supplement) and CPS Internet and Computer Use Supplement data:

FTP Download Page: [www.bls.census.gov/ferretftp.htm](http://www.bls.census.gov/ferretftp.htm)

DataFerrett: <http://dataferrett.census.gov/>

#### CPS ASEC (March Supplement) and CPS Internet and Computer Use Supplement documentation

FTP Download Page: [www.bls.census.gov/ferretftp.htm](http://www.bls.census.gov/ferretftp.htm)

DataFerrett: <http://dataferrett.census.gov/>

CPS Technical Documentation Page: [www.census.gov/apsd/techdoc/cps/cps-main.html](http://www.census.gov/apsd/techdoc/cps/cps-main.html)

#### Other CPS Supplement data

DataFerrett: <http://dataferrett.census.gov/>

#### Other CPS Supplement documentation

DataFerrett: <http://dataferrett.census.gov/>

CPS Technical Documentation Page: [www.census.gov/apsd/techdoc/cps/cps-main.html](http://www.census.gov/apsd/techdoc/cps/cps-main.html)

Basic CPS data and CPS Supplement data can be purchased on CD by calling the Census Bureau Customer Services Center at (301)763-INFO (4636) or through the Census Bureau online catalog ([www.census.gov/mp/www/cat/people\\_and\\_households/006327.html](http://www.census.gov/mp/www/cat/people_and_households/006327.html)).