APPENDIX A: LINKING HOUSEHOLD MEMBERS TOGETHER

To link records for married couples that include the householder, first identify the relevant variables for the analysis. In this example, the variables HHT (household type), RELP (relationship to householder), and MAR (marital status) can be used to identify married householders and their spouses. SCHL (educational attainment) can be used to measure educational attainment.

Next, create separate data files for each person that will be linked together. For each of these data files, rename the variables so they are unique for each person. Keep only the identifiers and person-specific variables that will be used in the analysis.

Following these steps produces a data file of married couples based only on the householder and his or her spouse. There may be additional married couples within a household—called subfamilies—that can be identified with the subfamily relationship variable. To create a data file that includes all married couples, data users will need to follow similar steps to identify married couples in subfamilies and link spouses together.

Here is a portion of a SAS program that identifies married householders and their spouses.

```
/*Create a person file of householder, which is the
reference person*/

data householder;
 keep serialno reference h_schl h_mar h_pwgtp;
  set psam_pus;
if hht in (1,2,3) and relp='00' and mar='1';

/*define householder and set it equal to 1 for the
householder */;
reference=1;
/*Create variables for the householder/reference
person;
   These variables will start with 'h_'*/;
h_schl=schl;
h_mar=mar;
h_pwgtp=pwgtp;
run;

/*Create a data file of spouses*/
data household_spouse;
 keep serialno spouse schl s_mar s_pwgtp;
 set persons;
if hht in (1,2,3) and relp='01' and mar='1';

spouse=1;
s_schl=schl;
s_mar=mar;
s_pwgtp=pwgtp;
run;

/*Sort the data before merging */;
proc sort data = householder;
 by SERIALNO;
proc sort data = household_spouse;
 by SERIALNO;
run;
```