Documentation for *MarriageVariables*

Contact

For all questions and errata associated with *MarriageVariables*, please email cmurray@aei.org.

Source Datafile

The ACS 1-year surveys from 2013 through 2020. Downloaded from IPUMS USA

Nomenclature

- *Male* and *female* denote the sexes.
- *Straight* denotes heterosexuality.
- *Same-sex* denotes homosexuality in general,
    - *Gay* denotes male homosexuality.
    - *Lesbian* denotes female homosexuality.

The dataset is named *MarriageVariables.* Two versions are posted: A *.csv* file in fixed comma-separated format with variable names in the first line, and a *.dta* file for analysis in Stata. The advantage of the Stata version is that codes for all the variables are labeled.

Both datasets have the same variables. The file has 26 variables, 15 of which are available in IPUMS USA's ACS files and 11 of which were created for this datafile. The universe for the variables consists of all persons ages 15 and older in the ACS surveys from 2013 through 2020 (15 is the youngest age eligible for a code of "married"). Three general notes:

1. The 15 variables downloaded from IPUMS USA use the numeric codes assigned by IPUMS USA. These may or may not correspond with the Census Bureau's numeric codes. Some of those variables also use labels for codes that differ from the ones assigned by the Census Bureau.
2. The Census Bureau made several adjustments for the 2020 ACS dataset to cope with the effects of the COVID-19 pandemic on data collection. This is especially relevant to the calculation of household and person weights, HHWT and PERWT. See the Census Bureau's guidance for details [ED embed link to https://usa.ipums.org/usa/acspumscovid19.shtml] and note its injunction to proceed with caution when using samples with 2020 data.
3. The word *partner* always indicates an intimate relationship with another person and can apply both to legal marriage and cohabitation.

Codebook for the IPUMS USA Variables

- **year**: The year in which the data were collected.
- **serial**: The ID number for the household (which may be duplicated across years).
- **pernum**: The person number within each household. The first-listed person (*pernum*=1) is called the householder, who may be anyone in the household who is owner or renter of the residence. *pernum* for the householder's married partner, if any, is 2.

- **hhwt**: Household weight. A number indicating how many households in the U.S. population are represented by a given household in the ACS sample.
- **perwt**: Person weight. A number indicating how many people in the U.S. population are represented by a given person in the ACS sample.
- **sex**: The person's sex.

  1 Male
  2 Female
- **age**: The person's age in years.
- **marst**: Marital status.

  1 Married, spouse present
  2 Married, spouse absent
  3 Separated
  4 Divorced
  5 Widowed
  6 Never married.
- **marrno**: Number of marriages.

  0 Not applicable
  1 Married once
  2 Married twice
  3 Married 3+
- **related**: IPUMS USA's detailed codes for identifying the relationship of each person in a household to the householder (a smaller set of codes is used in a variable named *relate*).

  101 Householder
  201 Spouse
  301 Child
  302 Adopted child
  303 Stepchild
  401 Child-in-law
  501 Parent
  601 Parent-in-law
  701 Sibling
  801 Sibling-in-law
  901 Grandchild
  1001 Other relative
  1114 Unmarried partner
  1115 Housemate/roommate
  1241 Roomers/boarders
  1242 Foster child
  1260 Other non-relatives
  1270 Group quarters non-inmate
  1301 Institutional inmate
- **sploc**: A variable constructed by IPUMS USA giving *pernum* for the respondent's partner (married or cohabiting). If no partner has been identified, *sploc*=0.
- **fertyr**: Was a child born to the person in within the last year?

  0 N/A
  1 No
  2 Yes

- **educd**: Educational attainment, detailed version.

  1 N/A
  2 No schooling completed
  11 Nursery/preschool
  12 Kindergarten
  14 Grade 1
  15 Grade 2
  16 Grade 3
  17 Grade 4
  22 Grade 5
  23 Grade 6
  25 Grade 7
  26 Grade 8
  30 Grade 9
  40 Grade 10
  50 Grade 11
  61 12th grade, no diploma
  63 High school diploma
  64 GED or alternative
  65 Some college, < 1 yr
  71 1+ yrs coll, no degree
  81 Associate's degree
  101 Bachelor's degree
  114 Master's degree
  115 Professional degree
  116 Doctoral degree

- **empstatd**: Employment status, detailed version

  0 N/A
  10 At work
  12 Has job, not working
  14 Armed forces—at work
  15 Armed forces—not at work
  20 Unemployed
  30 Not in Labor Force

- **occ2010**: The person's occupation, 2010 classification. An Excel file, *OCC2010 List.xlsx*, with 443 occupational codes and their labels, is included in the package of OSRC files for *MarriageVariables*. The labels are also embedded in the Stata version of the dataset.

The Variables Created for *MarriageVariables*

- **race5**: A person's assignment to one of five categories of race/ethnicity. "NL" denotes "non-Latino."

  1 NL White
  2 NL Black
  3 Latino
  4 NL Asian
  5 NL Other

- **altmarst**: Augmented marital status

  10 Married, spouse present, straight
  11 Married, spouse present, lesbian
  12 Married, spouse present, gay
  13 Married, spouse absent

20 Cohabiting, straight
21 Cohabiting, lesbian
22 Cohabiting, gay
30 Single, separated
31 Single, divorced
32 Single, widowed
33 Single, never married

- **psex**: Partner's sex. Same codes as *sex*
- **page**: Partner's age in years.
- **prace5**: Partner's race/ethnicity. Same codes as *race5*.
- **pmarst**: Partner's marital status. Same codes as *marst*.
- **pmarrno**: Number of marriages. Same codes as *marrno*.
- **pfertyr**: Has the partner had a baby within the last year? Same codes as *fertyr*.
- **peducd**: Partner's educational attainment. Same codes as *educd*.
- **pempstatd**: Partner's employment status. Same codes as *empstatd*
- **pocc2010**: Partner's occupation. Same codes as *occ2010*.

Additional Information

The Construction of *sploc*

- The IPUMS USA variables need no additional description except *sploc*, which stands for "spouse location." It represents the person number (*pernum*) of a partner (married or cohabiting) in the same household.
- In the 2013–2020 combined ACS surveys, 97.9 percent of all persons who said they were married with spouse present (*marst*=1), were either the householder or the partner of the householder. For the other 2.1 percent of people for whom *marst*=1, *sploc* for both partners was determined by a set of decision rules developed by IPUMS USA staff. Those rules are classified in a separate variable, *sprule*, that is not part of *MarriageVariables*.
- The nature of the rules can best be conveyed an example. Recall that *related* indicates the relationship of each member of a household to the householder. If the value of *related* for one person represents "child" and the value of *related* for *sploc* represents "child-in-law" and *marst* for both is "married, spouse present," the inference is that they are a couple living in the household of a parent. If *sex* is coded "female" for both of them, they are probably a lesbian couple. This is simplified example. To get a sense of the complications associated with constructing *sploc* and the care taken to deal with them, see two working papers from the Minnesota Population Center:
- Matthew Sobek and Sheela Kennedy, "The Development of Family Interrelationship Variables for International Census Data," Working Paper No. 2009-02, Minnesota Population Center, 18 November 2009, available online at

https://international.ipums.org/international/resources/misc_docs/pointer_working_paper_2009.pdf.

- Marina M. Gorsuch and Kari C.W. Williams, "Family Matters: Development of New Family Interrelationship Variables for US IPUMS Data Projects," Working Paper No. 2016-5, October 2016, available online at https://assets.ipums.org/_files/mpc/wp2016-05.pdf.

The Difference Between *altmarst* and *coupletype*

In 2019, IPUMS USA included a new Census Bureau variable in its ACS files, *coupletype*, that discriminates among four types of households in which the householder has a partner. Using my nomenclature, the codes for *coupletype* are:

- 0 N/A
- 1 Straight married
- 2 Same-sex married
- 3 Straight cohabiting
- 4 Same-sex cohabiting

For years from 2019 onward, and combined with *marst*, *sex*, and *psex*, *coupletype* can provide the same detailed codings as *altmarst* for couples that involve the householder. But *coupletype*'s codes for the householder and his or her partner apply to all the members of the household. For a researcher analyzing marital status of individuals, this requires identifying and coding as "missing" all persons in the household who are neither the householder nor the householder's partner, but who have a *marst* code of "married, spouse present." That amounts to 162,378 persons in the 2019 survey. *altmarst* provide codes that apply to those 162,378 people as well as to couples involving the householder.

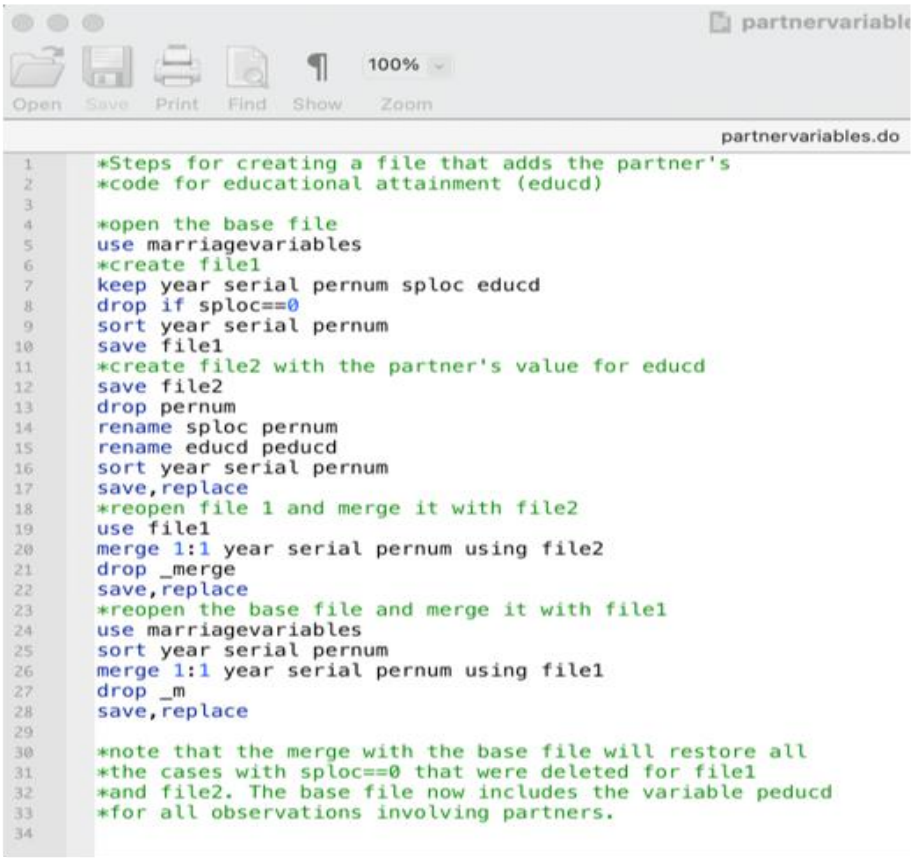
A Quick and Easy Method for Bringing Partner Variables onto the Same Line

The nine "partner" variables (*psex*, *prace5*, etc.) in *MarriageVariables* represent only some of the variables for characterizing a partner that a researcher might want to use. If you want to add additional partner variables, here is the method that I used:

1. Open the base file with which you are working
2. Create a subfile with *year*, *serial*, *pernum*, *sploc*, and the partner variables you wish to define (File 1)
3. Delete all cases for which there is no partner (*sploc*=0) and sort by YEAR, *serial*, and *pernum*. Save.
4. Create a duplicate file (File 2).
5. In File 2, delete the variable *pernum* and rename *sploc* as *pernum*.
6. In File 2, rename all the partner variables with a prefix indicating that they refer to the partner (in *MarriageVariables*, I used the prefix "P")
7. Sort File 2 by *year*, *serial*, and *pernum*. Save.
8. Open File 1. Merge with File 2 using *year*, *serial*, and *pernum*. Save.

File 1 now contains *year*, *serial*, *pernum*, *sploc*, the original variables, and the partner values for those variables. It is now ready to merge with your base file. That merge will restore all the cases for which *sploc*=0 and which were dropped in File1.

For Stata users, here is the .do file for creating a partner value for *educd*. It can be generalized to any set of ACS variables.




```
*Steps for creating a file that adds the partner's
*code for educational attainment (educd)

*open the base file
use marriagevariables
*create file1
keep year serial pernum sploc educd
drop if sploc==0
sort year serial pernum
save file1
*create file2 with the partner's value for educd
save file2
drop pernum
rename sploc pernum
rename educd peducd
sort year serial pernum
save,replace
*reopen file 1 and merge it with file2
use file1
merge 1:1 year serial pernum using file2
drop _merge
save,replace
*reopen the base file and merge it with file1
use marriagevariables
sort year serial pernum
merge 1:1 year serial pernum using file1
drop _m
save,replace

*note that the merge with the base file will restore all
*the cases with sploc==0 that were deleted for file1
*and file2. The base file now includes the variable peducd
*for all observations involving partners.
```

Caveats

altmarst's coding of same-sex married couples applies only to ACS surveys conducted from 2013 onward. Before 2013, the Census Bureau recoded *relate* and *marst* values to exclude same-sex married couples.

The ACS may be used only to count gay or lesbian *couples*, inferred from the sexes of the two partners. It permits no inferences about the sexual orientation of married persons if the spouse is absent, nor about the sexual orientation of never-married, separated, divorced, or widowed persons, which also means that the ACS may not be used to estimate the overall size of the lesbian and gay populations.

Using MarriageVariables

MarriageVariables can be used as a stand-alone file for some analyses, but for most topics it will need to be merged with additional ACS variables downloaded from IPUMS USA. Sorting on *year*, *serial*, and *pernum* provides the basis for a one-to-one match with the observations in files to be merged.