

Understanding Society User Support - Support #22

Matching parent's information to their children (aged >15) in the same household

02/01/2012 09:22 AM - Maren M. Michaelsen

Status:	Closed	Start date:	02/01/2012
Priority:	Normal	% Done:	50%
Assignee:	Redmine Admin		
Category:	Data analysis		
<p>Description</p> <p>Dear Understanding Society survey/BHPS users,</p> <p>I am matching parent's information on age to their children (aged >15) in the same household using STATA.</p> <p>The results show that in 192 (about 5%) of the cases the child is at least as old as the mother (same for some fathers). Also, there are several male mothers and female fathers.</p> <p>Please find below the commands I am using to match household and individual data and the results.</p> <p>Please let me know whether I made a mistake or if there are indeed errors in the data.</p> <p>Maren</p> <pre>/* Merge data sets for cross-section */ /* a_indresp.dta */ use data/a_indresp.dta, clear sort a_hidp pidp save data/indresp_sorted.dta, replace file data/indresp_sorted.dta saved /* a_indall.dta */ use data/a_indall.dta, clear sort a_hidp pidp save data/indall_sorted.dta, replace file data/indall_sorted.dta saved /* merge sorted datasets */ merge 1:1 a_hidp pidp using data/indresp_sorted.dta, keep(3) Result # of obs. ----- not matched 0 matched 50,994 (_merge==3) ----- cap drop _merge /* Merge prents information on children */ bysort a_hidp (pidp): gen m_sex = a_sex[a_mnspno] (45333 missing values generated) bysort a_hidp (pidp): gen m_dvage = a_dvage[a_mnspno] (45333 missing values generated) sum a_dvage m_dvage if a_dvage >= m_dvage Variable Obs Mean Std. Dev. Min Max -----+-----</pre>			

a_dvage	192	23.28125	8.068039	16	69
m_dvage	192	22.38021	7.588654	16	69

History

#1 - 02/01/2012 05:52 PM - Redmine Admin

- Category set to Data analysis
- Status changed from New to In Progress
- Assignee set to Redmine Admin
- % Done changed from 0 to 50

Please examples in User Guide (Example 3)

https://www.understandingsociety.ac.uk/files/data/documentation/wave1/User_manual_Understanding_Society_Wave_1.pdf#page=29

and [BHPS course materials](#)

#2 - 02/02/2012 01:56 PM - Redmine Admin

Dear Maren

bysort a_hidp (pidp): gen m_sex = a_sex[a_mnspno]

generates m_sex as the sex of the person who takes position a_mnspno in a data matrix sorted by a_hidp and pidp. So if the mother's pno=1, m_sex will be generated as the sex of the person with the lowest pidp in the household; if the mother's pno=5, m_sex will be generated as the sex of the person with the fifth lowest pidp in the household. This may or may not be the mother. In fact, since you keep only observations that are in indresp and indall, the mother's record may not even be in the matrix.

Please also note the difference between PIDP and w_PNO. Assignment of PNO within wave is 'random'; the household reference person enumerates all members of the household and whoever is mentioned first gets assigned pno=1, second pno=2 etc. A person's PNO may change over time; pidp is fixed. The person with the highest pidp in the household may not be the person with the highest pno. If it is, this is a first wave co-incidence that may very well change over time.

You could do the following (there may be other ways!):

1. load indresp variables of interest, rename a_pno to a_mnpno (rename all other variables so it is clear it refers to the mother's characteristics, too!),
2. merge file created in (1.) on a_hidp a_mnpno to indall, keeping perfect matches only,
3. merge file created in (2.) on a_hidp a_pno to original indresp to have a record of individual responses plus any of the mother's characteristics reported in indresp.

Hope this helps,
Gundi

#3 - 02/23/2012 01:38 PM - Redmine Admin

- Status changed from In Progress to Closed

#4 - 11/10/2015 02:57 PM - Gundi Knies

- Target version set to X M