

---

**Subject: Re: Merging Individual and HH Member Recode Files**

Posted by [amw289](#) on Sun, 31 Jul 2016 16:02:45 GMT

[View Forum Message](#) <> [Reply to Message](#)

---

Thank you for your response. I understand that cluster numbers are repeated across waves, which is why I added v007 to the end of the list of merging variables. v007 captures the year the survey data were collected (and years were not repeated across waves, as far as I know). I have tried your suggestion of merging individual and household member files separately for each wave, but I still run into the same two problems.

1. In 2012, there are 288 duplicate observations in terms of v001 v002 v003 in the individual file (PEIR6IFL). Yet there are no duplicates in terms of caseid. Isn't the caseid is based on v002 v002 and v003? If so, shouldn't there be no duplicates in terms of v001 v002 and v003? Please advise.

2. When I try merging the individual file to the household member file, separately by wave, I run into the same problem I described earlier, where there are many observations from the household member file that remain unmatched. For example, just using the data files from 2011, when I attempt a 1:m merge for (individual to household member), I get the following results in stata:

Result	# of obs.	
-----		
not matched	76,145	
from master	0	( <code>_merge==1</code> )
from using	76,145	( <code>_merge==2</code> ) (these are all from the household member file)
matched	22,517	( <code>_merge==3</code> )

When I go to check if there are eligible women among the unmatched observations from the household members file (`_merge==2`) this is what I get:

Number of  
eligible  
women in HH Freq. Percent Cum.

0	19,335	25.39	25.39
1	41,003	53.85	79.24
2	12,080	15.86	95.11
3	2,926	3.84	98.95
4	699	0.92	99.87
5	85	0.11	99.98
6	9	0.01	99.99
7	8	0.01	100.00

Total 76,145 100.00

In other words, of the 76,145 unmatched household members from 2011, 79% reside in a household with at least one woman who was eligible to participate. Is this just telling me about

non-response rates? Do I not need to worry about these households not being merged to observations in the individual file?

Thank you again for your time and assistance!