

---

Subject: Re: Merged HH and Individual Data  
Posted by [Bridgette-DHS](#) on Fri, 02 Nov 2018 14:40:49 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Following is a response from Senior DHS Stata Specialist, Tom Pullum:

I'm not sure what your question is. In order to do a merge of the IR and PR files for this survey, you must match hv024, hv001, hv002, hvidx in the PR file with v024, v001, v002, and v003, respectively, in the IR file. When there are multiple women respondents in the same household, this will give you all of them. Are you asking how to randomly select just one of these women, in households that had more than one woman? Why would you want to do that? If you do that, you will lose representativeness of the population of women (although you can restore representativeness by multiplying v005 by the number of women in the household) and you will reduce the effective sample size.

If you want to look at how the women's characteristics are related to household characteristics, you do not need to do any such subsampling. You would keep all the women, and just interpret the household's characteristics as characteristics of the woman. For some purposes, you could think of doing multi-level analysis and treating the household as the level 2 unit (and the cluster as level 3), but there is little gain from doing this because most households contain only one woman.