

---

Subject: Re: Merged HH and Individual Data

Posted by [Bridgette-DHS](#) on Mon, 05 Nov 2018 12:27:47 GMT

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

---

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

You could do something like the following. Before you do the merge, assign the women in the IR file a random number; then sort the women in the household by this number; then keep the one who is first in the sort. Then do the merge with the household file. in the following, I use the HR file rather than the PR file, and just keep the household-level variables.

```
set more off
set maxvar 10000
use ...IAHR74FL.DTA" , clear
drop *_*
sort hv024 hv001 hv002
save ...IAHRtemp.dta, replace

use ...IAIR74FL.DTA" , clear
gen rn=uniform()
sort v024 v001 v002 rn
egen sequence=seq(), by(v024 v001 v002)
tab sequence
keep if sequence==1
gen hv024=v024
gen hv001=v001
gen hv002=v002
sort hv024 hv001 hv002
merge hv024 hv001 hv002 using e:\DHS\DHS_data\scratch\IAHRtemp.dta
keep if _merge==3
drop sequence _merge
```