Subject: Having problem in merging household members data using birth data
Posted by ddd2332 on Sat, 08 Nov 2014 00:51:55 GMT

View Forum Message <> Reply to Message

Hello.

I am trying to merge the birth order and twin variables from the birth dataset to the householdmember dataset for my research analysis (for India dataset).
I know that the identifier variables are v001 and v002, but I can see that merging by using v001 and v002 will mess up since different individuals have same v001 and v002 within house and that makes the identifiers not unique..
I am trying to figure out how to make a unique identifier so that I can merge these two data.

I would really really appreciate if someone could help me out on this..

Thank you so much!

Jung.

Subject: Re: Having problem in merging household members data using birth data
Posted by Trevor-DHS on Sat, 08 Nov 2014 01:51:36 GMT

View Forum Message <> Reply to Message

Unfortunately you can not get a complete match, even for children as the information about birth order and whether the child is a twin are only available if the mother of the child also lives in the household, is aged between 15 and 49 and is successfully interviewed. the following code will match as much as is possible:

```
* Use the births recode file
use "IABR52FL.dta"
* Keep only the variables of interest
* Cluster, household, birth order, whether twin, line number from household schedule
keep v001 v002 bord b0 b16
* rename the cluster, household and line number for merging
rename v001 hv001
rename v002 hv002
rename b16 hvidx
* drop cases that don't have a household line number - either dead or live elsewhere
drop if hvidx==. | hvidx==0
* sort the data on cluster, household and line number
sort hv001 hv002 hvidx
* save a temporary working file
save "iabr52_subset.dta", replace

* Use the household members recode file
use "IAPR52FL.dta", clear
```

* Sort on the cluster, household and line number
sort hv001 hv002 hvidx
* Merge the data to add the birth order and whether a twin
merge 1:1 hv001 hv002 hvidx using "iabr52_subset.dta"
* Tabulate by age to see the match
tab hv105 _merge, row

Tabulate _merge by age to see which cases do not match (where _merge==1) and we get about an 85% match for children up to about age 8, and then it drops off to only about 70% by age 15. It is not perfect, but hopefully will be useful for your needs.

Subject: Re: Having problem in merging household members data using birth data
Posted by ddd2332 on Sat, 08 Nov 2014 02:56:25 GMT
View Forum Message <> Reply to Message

Thank you so much for your help! I really appreciate it!