Subject: Merging Birth and Household Recodes
Posted by cbdolan on Tue, 05 Apr 2016 13:47:07 GMT
View Forum Message <> Reply to Message

I am using the 2007 and 2013/14 Birth and Household Recode files for the DRC.

According to the matching variable chart http://www.dhsprogram.com/data/Merging-Datasets.cfm, I can merge the birth and household recode file by constructing an equivalent match variables of v001+v002. I have done this using the Stata code below.

However, it isn't a perfect match. There are records in the using (household recode) that don't match the master (birth recode). This could be plausible. There could be households with no kids and therefore the output would make sense.

When searching the forum I see some people using additional variables to match birth and household recodes (b16, b6, and sometimes bidx or hidx). Therefore, I wanted to clarify that I am interpreting the match variable chart correctly and verify that I am correctly doing the match on v002+v002 (renamed in HH file).

Thanks!

```
*2013/2014
set more off
use "Y:\Carrie\REU\Malaria\Data\4_DHS_HouseholdRecode\CDHR61FL.DTA"
keep hhid hv000 hv001 hv002 hv003 hv004 hv005 hv006 hv007 hv008 hv009 hv010 hv011 hv012
hv013 hv014 hv015 hv016 hv017 hv018 hv019 hv020 hv021 hv022 hv023 hv024 hv025 hv026
hv027 hv028 hv030 hv031 hv032 hv035 hv040 hv041 hv042 hv044 hv201 hv202 hv204 hv205
hv206 hv207 hv208 hv209 hv210 hv211 hv212 hv213 hv214 hv215 hv216 hv217 hv218 hv219
hv220 hv221 hv225 hv226 hv227 hv228 hv230a hv230b hv232 hv232b hv232c hv232d hv232e
hv232y hv234 hv234a hv235 hv236 hv237 hv237a hv237b hv237c hv237d hv237e hv237f hv237g
hv237h hv237i hv237j hv237k hv237x hv237z hv238 hv239 hv240 hv241 hv242 hv243a hv243b
hv243c hv243d hv244 hv245 hv246 hv246a hv246b hv246c hv246d hv246e hv246f hv246g
hv246h hv246i hv246j hv246k hv247 hv252 hv253 hv253a hv253b hv253c hv253d hv253e hv253f
hv253g hv253h hv253x hv253z hv270 hv271 hml1 hml1a hml2
rename hv001 v001
rename hv002 v002
sort v001 v002
save "Y:\Carrie\REU\Malaria\Data\4_DHS_HouseholdRecode\HO2_RECH2_13.dta",replace
clear

use "Y:\Carrie\REU\Malaria\Data\4_DHS_BirthRecode\CDBR61FL.dta"
sort v001 v002
merge m:1 v001 v002 using
"Y:\Carrie\REU\Malaria\Data\4_DHS_HouseholdRecode\HO2_RECH2_13.dta"
```

```
keep if _merge==3
drop _merge
*N=59,276(matched)
save "Y:\Carrie\REU\Malaria\Data\4_DHS_HouseholdRecode\HO2_RECH2_BR_13.dta", replace
clear

*2007
use "Y:\Carrie\REU\Malaria\Data\4_DHS_HouseholdRecode\cdhr50fl.DTA"
keep hhid hv000 hv001 hv002 hv003 hv004 hv005 hv006 hv007 hv008 hv009 hv010 hv011 hv012
hv013 hv014 hv015 hv016 hv017 hv018 hv019 hv020 hv021 hv022 hv023 hv024 hv025 hv026
hv027 hv028 hv030 hv031 hv032 hv033 hv035 hv040 hv041 hv042 hv043 hv044 hv201 hv202
hv204 hv205 hv206 hv207 hv208 hv209 hv210 hv211 hv212 hv213 hv214 hv215 hv216 hv217
hv218 hv219 hv220 hv221 hv225 hv226 hv227 hv228 hv234 hv235 hv236 hv237 hv237a hv237b
hv237c hv237 hv237e hv237f hv237g hv237h hv237i hv237j hv237k hv237x hv237z hv238 hv239
hv240 hv241 hv242 hv243a hv243b hv243c hv243d hv244 hv245 hv246 hv246a hv246b hv246c
hv246d hv246e hv246f hv246g hv246h hv246i hv246j hv246k hv247 hv248 hv249 hv250 hv251
hv270 hv271 hml1 hml1a hml2
rename hv001 v001
rename hv002 v002
sort v001 v002
save "Y:\Carrie\REU\Malaria\Data\4_DHS_HouseholdRecode\HO2_RECH2_07.dta", replace
clear

use "Y:\Carrie\REU\Malaria\Data\4_DHS_BirthRecode\CDBR50FL.dta"
sort v001 v002
merge m:1 v001 v002 using
"Y:\Carrie\REU\Malaria\Data\4_DHS_HouseholdRecode\HO2_RECH2_07.dta"
keep if _merge==3
drop _merge
*N=29,548 (matched)
save "Y:\Carrie\REU\Malaria\Data\4_DHS_HouseholdRecode\HO2_RECH2_BR_07.dta",replace
append using "Y:\Carrie\REU\Malaria\Data\4_DHS_HouseholdRecode\HO2_RECH2_BR_13.dta"
save "Y:\Carrie\REU\Malaria\Data\4_DHS_HouseholdRecode\HO2_RECH2.dta",replace
```

Subject: Re: Merging Birth and Household Recodes
Posted by Bridgette-DHS on Thu, 07 Apr 2016 18:23:14 GMT
View Forum Message <> Reply to Message

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

To merge the BR (or KR) file with the household file, you need to use the cluster id, household id, and line number. In the BR (or KR) file, the line number of the child, if in the household, is b16.

The following lines will work for either the DRC 2007 survey or the DRC 2013-14 survey. Just change the paths and add in the other variables you want to have in the merged file.

There are a few surveys from other countries where there is a sub-household id that must also be matched on. The best way to find that is to see whether hhid (a string) corresponds with hv001 and hv002. If there is something else embedded in hhid, you must extract it with the substring command, and also extract it from caseid in the IR/BE/KR file, and include it in the sort and merge lines. However, these two DRC surveys do not include sub-households.

* MERGE THE BR AND PR FILES FROM THE 2007 DRC SURVEY

```
use e:\DHS\DHS_data\PR_files\CDPR50FL.dta, clear
keep hhid hv001 hv002 hvidx
sort hv001 hv002 hvidx
save e:\DHS\DHS_data\scratch\temp.dta, replace

use e:\DHS\DHS_data\BR_files\CDBR50FL.dta, clear
list caseid v001 v002 v003 b16 if _n<=20, table clean
keep caseid v001 v002 v003 b16
rename v001 hv001
rename v002 hv002
rename b16 hvidx
sort hv001 hv002 hvidx
merge hv001 hv002 hvidx using e:\DHS\DHS_data\scratch\temp.dta
```

* This match works fine

* MERGE THE BR AND PR FILES FROM THE 2013-14 DRC SURVEY

```
use e:\DHS\DHS_data\PR_files\CDPR61FL.dta, clear
keep hhid hv001 hv002 hvidx
sort hv001 hv002 hvidx
save e:\DHS\DHS_data\scratch\temp.dta, replace

use e:\DHS\DHS_data\BR_files\CDBR61FL.dta, clear
list caseid v001 v002 v003 b16 if _n<=20, table clean
keep caseid v001 v002 v003 b16
rename v001 hv001
rename v002 hv002
rename b16 hvidx
sort hv001 hv002 hvidx
merge hv001 hv002 hvidx using e:\DHS\DHS_data\scratch\temp.dta
```

* This match works fine

Thanks for this. This code makes sense and I ran the merge.

When you say "*this match works fine". What do you mean by that?

I ran the match using the exact code below and only changed the paths. See tab output below.

I understand how some PR (n=54,799) might not have kids, but not how some BR (18,126) might not have households. Only 36% of the data is matching perfectly.

Does that make sense to you? What am I missing?

tab _merge

```
    _merge |   Freq.    Percent     Cum.
------------+-----------------------------------
         1 |   18,126     15.89     15.89
         2 |   54,799     48.04     63.93
         3 |   41,150     36.07    100.00
------------+-----------------------------------
     Total |  114,075    100.00
```

The same for 2007

tab _merge

```
    _merge |   Freq.    Percent     Cum.
------------+-----------------------------------
         1 |    9,716     16.75     16.75
         2 |   28,459     49.06     65.81
         3 |   19,832     34.19    100.00
------------+-----------------------------------
     Total |   58,007    100.00
```

Subject: Re: Merging Birth and Household Recodes
Posted by Bridgette-DHS on Fri, 08 Apr 2016 17:27:22 GMT
View Forum Message <> Reply to Message

A response from Senior DHS Stata Specialist, Tom Pullum:

Hi Carrie--

I'm glad it worked. That's all I meant by "*this match works fine". There were no error messages. I should have removed the comments. I checked one merge and then the other and just inserted the comments in case I had to switch to something else....

The BR file includes all "children" in the birth histories, many of whom are now adults and living in their own households. I'm sure that almost all of the of the very youngest children will be living in the same household as their mother. For children of increasing age, the probability that the mother has died or lives in a different household will increase steadily. It's not that the older children don't have households--it's just that they are not in the same household as the mother.

You may actually want this merge only for children under five, which you would get by using the KR file rather than the BR file.

Let me know if you have other questions.

Subject: Re: Merging Birth and Household Recodes
Posted by Bridgette-DHS on Fri, 15 Apr 2016 11:29:59 GMT
View Forum Message <> Reply to Message

Hi Tom,

I'm still working on this merge. This should be relatively intuitive, but I must not understand something about the file construction.

I've used the attach code which reflects your previous suggestions. However, when I keep the match==3 (meaning in both the BR and PR files) no one dies. B7 is '.'. for all _merge==3.

What am I missing? I'm trying to assign household level characteristics to the kids in the BR file. For the purpose of the Filmer Pritchett index I need the H02 RECH2 variables that are in the PR recode to be matched to the kids.

Thanks. I'm sure I'm missing something obvious.

Carrie

Subject: Re: Merging Birth and Household Recodes

Posted by Bridgette-DHS on Fri, 15 Apr 2016 11:34:00 GMT

Another response from Tom Pullum:

* If the child is not in the household, which happens if the child has died or is alive but elsewhere,
*   use the mother's line number as a proxy
* 90 in the next line is arbitrary
replace hvidx=v003 if hvidx==0 | hvidx>90

Then you will retain the children who have died.

Please let me know if this works.

Subject: Re: Merging Birth and Household Recodes
Posted by cbdolan on Fri, 15 Apr 2016 15:16:02 GMT

Works perfectly. Thanks!