Posted by UserUkbasedresearcher on Wed, 19 Jul 2017 13:38:53 GMT
View Forum Message <> Reply to Message

Hello,
I am using the Yemen 2013 DHS data. I am trying to recreate the tables in the DHS report for child nutrition. Using the PR file I can recreate all estimates correctly. However, when I try to merge KR and PR files I get results for the variables which use KR variables that are out by by between .2 and .7 percent. For instance for WHZ and HAZ by birthsize and birth interval etc. (these use the KR dataset).

I am using weights and, as stated, I can recreate the variables that call only for variables in the PR file perfectly. Below is an example of my code. I feel there is something I am missing here in the merging or in the observations I am using for the estimates. The discrepancies are small but big enough to really matter (up to .7%).
Here is an example of my code:
To merge:

```
set maxvar 30000
use "C:\Users\Documents\Yemen DHS 2013\yepr61dt\YEPR61FL.DTA", clear
rename hv001 v001
rename hv002 v002
gen in_PR=1
sort v001 v002
save "C:\Users\Documents\Yemen DHS 2013\temp.dta", replace
use "C:\Users\Documents\Yemen DHS 2013\yekr61dt\YEKR61FL.DTA", clear
gen in_IR=1
sort v001 v002
merge v001 v002 using "C:\Users\Documents\Yemen DHS 2013\temp.dta"
tab1 _merge in_PR in_IR
--
```

To calculate % WHZ -2 SD by birthsize I open the file :

```
set maxvar 30000

use "C:\Users\Documents\Yemen DHS 2013\temp.dta"

gen wgt = hv005/1000000

svyset [pw=wgt], psu(hv021) strata(hv022)

gen eligch = 0
replace eligch = 1 if (hv103==1 & hc70 < 9996 & hc71 < 9996 & hc72 < 9996)

codebook hc72
tab hc72 if hc72>9990,m
```

```
tab hc72 if hc72>9990,m nolabel
gen WHZ=hc72
replace WHZ=. if WHZ>=9996
gen wasted=.
replace wasted=0 if WHZ ~=.
replace wasted=1 if WHZ<-200
tab wasted

gen birthsize1 = .
replace birthsize1 = 1 if m18 ==4  & eligch==1
replace birthsize1 = 2 if m18 ==1  & eligch==1
replace birthsize1 = 2 if m18 ==2  & eligch==1
replace birthsize1 = 2 if m18 ==3  & eligch==1
replace birthsize1 = 3 if m18 ==8. & eligch==1
replace birthsize1 = 3 if m18 ==. & eligch==1
label define birthsize1 1"smaller than average" 2"Average or larger" 3"missing"
label values birthsize1 birthsize1
 --
svy: tab wasted birthsize1 if eligch==1, col
```

The estimates are out by small amounts. What am I doing doing wrong here?
Any help would be much appreciated!!
Thanks,