Subject: weighting multiple rounds of Bangladesh DHS
Posted by doris0531 on Tue, 27 Nov 2018 12:54:25 GMT
View Forum Message <> Reply to Message

Dear DHS users,

I am using birth record data from appending Bangladesh DHS 1999/2000, 2004, 2007, 2011 and 2014 to study the determinants of infant/neonatal mortality. I have read some discussion on the user forum about weighting data and setting "svyset" commands and I still have two questions regarding to my specific context:

(1) How to set new "strata" if stratification is different across years? . Bangladesh 2011 and 2014 DHS, the most recent two years have different "strata design" from previous years. Here are two screenshots from BDHS report. Can I still set new strata as "egen strata = group (survey v024 v025) where survey indicates different rounds? And also "v022" in early years are very different from the ones in BDHS 2011 and 2014.

2011

2007

(2) How to re-weight the appended data across survey rounds if I am only interested in the birth records two years prior to the interview time? Are the following ways appropriate for my study?
generate weight = v005/10000000
egen clusters=group(survey v021)
egen strata = group(survey v024 v025)
svyset clusters [pweight=weight], strata(strata)

Any suggestions or comments will be highly appreciated. Thank you!
regards,
Qiao

```
File Attachments
1) 2011.png, downloaded 784 times
2) 2007.jpg, downloaded 832 times
```

Subject: Re: weighting multiple rounds of Bangladesh DHS
Posted by Bridgette-DHS on Wed, 12 Dec 2018 18:28:55 GMT
View Forum Message <> Reply to Message

Following is a response from Senior DHS Specialist, Kerry MacQuarrie:

Here is code to create the correct strata variable in each of the Bangladesh DHS surveys. This code can be run in a loop across all survey datafiles.

egen strata = group(v023 v025) if (v000 =="BD3"& (v007 == 93 | v007==94)) /*Bangladesh 1993-94*/
gen strata = v023 if (v000 =="BD3"& (v007 == 96 | v007==97)) /*Bangladesh 1996-97*/
gen strata = v023 if (v000 =="BD3"& (v007 == 99 | v007==2000 | v007 == 0)) /*Bangladesh 1999-00*/
egen strata = group(v024 v025) if v000 =="BD4" /*Bangladesh 2004*/
gen strata = v023 if v000 =="BD5" /*Bangladesh 2007*/
gen strata = v023 if (v000 =="BD6" & v007==2011) /*Bangladesh 2011*/
gen strata = v023 if (v000 =="BD6" & v007==2014) /*Bangladesh 2014*/

The other variables (weight and cluster) you would need to svyset the data for any of the surveys would be as follows:

gen wt=v005/1000000
gen psu=v001
svyset [pw=wt], psu(psu) strata(strata) singleunit(center)

However, if you are appending different survey years together, you have to be very careful as some of the PSU or strata values may occur in more than one datafile--but they don't refer to the same PSU or stratum. You need to ensure that your PSU values and strata values are unique. There are several ways to do this. One way I often use is to first create a survey indicator variable (e.g. year=1 for Bangladesh 1993-94; year=2 for Bangladesh 1996-97; year=3 for Bangladesh 1999-00, etc). Then:

gen psu=.
replace psu=10000 + v001 if yr==1
replace psu=20000 + v001 if yr==2
replace psu=30000 + v001 if yr==3

replace strata = strata+10000 if yr==1
replace strata = strata+20000 if yr==2
replace strata = strata+30000 if yr==3

Then, you'll end up with harmonized strata and psu variables with unique values and your svyset command with accurately account for the complex sampling design.