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.