Subject: Strata Computation for Survey Sampling Weighting
Posted by AkiodeN on Mon, 27 Apr 2015 12:01:42 GMT
View Forum Message <> Reply to Message

I am currently running some estimation analysis on 2013 Nigeria demographic and health survey in STATA. But I am confused about computing the strata component of my sampling weighting. Please which is correct:

gen psu=v021
gen strata=v023
gen weight=v005/1000000
svyset psu [pweight=weight], strata(strata)

OR

gen psu=v021
gen strata= group(v024 v025), label
gen weight=v005/1000000
svyset psu [pweight=weight], strata(strata

Subject: Re: Strata Computation for Survey Sampling Weighting
Posted by Liz-DHS on Fri, 01 May 2015 15:52:23 GMT
View Forum Message <> Reply to Message

Dear User,
Here is a response from one of our experts, Dr. Tom Pullum:
Quote:In the Nigeria 2013 survey, the strata in the IR file would be determined by EITHER "gen strata=v023" OR "egen strata=group(sstate v025)". That is, v023 will match exactly with group(sstate v025).
In the PR file, you would replace v023 with hv023 and sstate with shstate.
You are correct that usually v023 will match with group(v024 v025), but in the Nigeria survey v024 gives region and shstate gives state, and the strata are the urban and rural sectors of the states, rather than the regions.
Note that "group" requires "egen" rather than "gen". You omitted the "e".
Otherwise your command is correct. However, you do not need to divide v005 by 1000000. You can do it if you want, but the weights that go with svyset are pweights, and those weights are always normalized internally to have a mean of 1. I suggest that you do a model with "gen weight=v005" and then re-run it with "gen weight=v005/1000000". The results should agree exactly. But note that only pweights are automatically normalized this way; aweights, fweights, and iweights are not normalized.

Thank you for your post.