Subject: Districts as cluster-level for multi-level model
Posted by dgodha on Sun, 23 Feb 2020 10:28:13 GMT
View Forum Message <> Reply to Message

Hello,

I will appreciate your expert guidance on my query. We usually use 'psu' as the cluster level in DHS data. In my case, the group size is too small if I use 'psu'.

Group Variable | #Groups Minimum Average Maximum
        psu | 25,063   1   3.1   16

Since NFHS-4 is representative at the district level and we have to anyway create a variable for the cluster-weight, I am wondering if it is possible to use district as the cluster-level. I tried changing my weighting command for psu to district but as you can see in the output, I don't get the p-values and CIs.
*Rescaling of weights
 gen wt=v005/1000000

*Level 1 weights using scaling method 1: New weights sum to district sample size
 gen sqw = wt*wt
 egen sumsqw = sum(sqw), by(sdistri)
 egen sumw = sum(wt), by(sdistri)
 gen pwt11 = wt*sumw/sumsqw

* Survey setting
 gen wt2=1
 svyset sdistri, weight(wt2) strata(v023) , singleunit(centered) || _n, weight(pwt11)

*Output
****
Number of strata =   2,509   Number of obs   = 1,538,126
Number of PSUs   =   2,509   Population size = 1,438,715
Subpop. no. obs =   78,446
Subpop. size   = 73,653.12
Design df   =   0
F(0,   0) =   .
Prob > F   =   .

Linearized
y	Coef.	Std. Err.	t	P>t	[95% Conf. Interval]	
_cons	-1.585093	.0192937	-82.16	.	.	.
sdistri						
var(_cons)	.1527032	.0153514			.	.

Note: 5 strata omitted because they contain no subpopulation members.
Note: Strata with single sampling unit centered at overall mean.

I am not sure what is going wrong and will appreciate any understanding.
Thank you
Deepali