Understanding Society User Support - Support #80

Stata code for survey weights

09/05/2012 01:28 AM - Daniel Sperber

Status:	Closed		**Start date:**	09/05/2012
Priority:	Normal		**% Done:**	100%
Assignee:				
Category:	Survey design			

Description

Hi,
I hope that this isn't too elementary a question, but I'm running into ambiguity in applying the US weighting variables in STATA code. I'm currently working with cross sectional analysis of the adult self complete section of wave 1 of the innovation panel (a_indscip_xw) , but will likely be using wave 1,3&4 of the innovation panel (d_indscip13_lw), and wave 1 of the main survey.

The other relevant weights (from the user manual) are:
a_psu (primary sampling unit)
a_strata (sample strata)
a_psnenip_xd (cross sectional person design weight)

This is what I have now:
svyset a_psu [pweight=a_indscip_xw], strata(a_strata) vce(linearized) single unit(missing) || a_psnenip_xd

Is this correct? If not, any suggestions on proper implementation of the survey weights in Stata? There is some sample code in the user guide. I think sample code related to this would be of great use.

Thanks in advance,

Dan

History

#1 - 09/10/2012 06:03 PM - Olena Kaminska

Dan,

Thank you for your question. Here is a simple response. I suggest that you use a syntax of the following form:
svyset a_psu [pweight=a_indscip_xw], strata(a_strata)

The above is enough to adjust for complex sample design. Specify other options only if you know what you are doing - otherwise I would rely on Stata default (there is always a good reason for the defaults).

a_indscip_xw is a correct weight for wave 1 self-completion instrument in IP, as well as d_indscip13_lw is a correct weight for waves 1,3,4 longitudinal analysis in IP.

It isn't clear for me whether you are thinking to combine UKHLS mainstage and IP. Note, these datasets are not designed to be combined. If you have a good reason to combine them (which is unlikely, but possible), please describe your analysis in details as well as reasons for combining the datasets. We may provide suggestions on how to best control for complex sample design in this situation. But just a note, from the design point of view it is best if UKHLS mainstage is used on its own for any substantive analysis.

Hope this helps,
Olena

#2 - 09/19/2012 11:11 AM - Redmine Admin

- Status changed from New to Closed

- % Done changed from 0 to 100