Understanding Society User Support - Support #1724

Weights and accounting for individual clustering

07/12/2022 11:58 AM - Catherine Bunting

Status:	Resolved	**Start date:**	07/12/2022
Priority:	Normal	**% Done:**	100%
Assignee:	Olena Kaminska		
Category:	Weights		

Description

Hello,

I am carrying out logistic regression to estimate the association between transitioning from unemployment to employment and health service use. My sample contains all individuals who were unemployed at UKHLS W7. Transition to employment is captured using their employment status at W8 and health outcomes are measured at W9.

I have created an equivalent cohort using individuals who were unemployed at W6, and pooled the two cohorts to increase my sample size. I therefore have some clustering by pidp, as individuals contribute twice to the analysis if they were unemployed at both W6 and W7.

In Stata, I am using svyset to specify the psu, cross-sectional weight (indscui_xw) and the strata. How can I also account for clustering by pidp? Normally I would use logistic regression with option vce(cluster pidp), but this is not possible when using the svy:logistic command.

Thank you!
Catherine

History

#1 - 07/12/2022 01:36 PM - Olena Kaminska

Catherine,

You only need psu as your clustering variable as it is the highest level (pidps are nested within psu), and in a straightforward logistic regression only psu needs to be indicated.
But reading your analysis description, I wonder if you need longitudinal weights. If you are using w6 - w9 information make sure you use w9 lw weight.

Best,
Olena

#2 - 07/12/2022 02:02 PM - Catherine Bunting

Hi Olena - thanks so much for the speedy reply, that's very helpful.

Just to clarify - if I have a group of individuals and am using information about them from waves 7, 8 and 9, I should use the W9 longitudinal weight, not the W7 cross-sectional weight?

Thanks,

Catherine

#3 - 07/12/2022 03:28 PM - Annette Pasotti

- Status changed from New to Feedback

- % Done changed from 0 to 90

#4 - 07/12/2022 04:36 PM - Catherine Bunting

Sorry Olena, one last question - if I use psu as the clustering variable, does that account for clustering by hidp as well as by pidp?

#5 - 07/13/2022 10:52 AM - Olena Kaminska

Kind of yes. The answer is more complicated because households are not a longitudinal concept, but I wouldn't worry about it.
Olena

#6 - 07/13/2022 11:07 AM - Olena Kaminska

And to reply to your earlier question, yes, use longitudinal weight from wave 9. As soon as you use information from 2 or more waves you need

longitudinal weights.
Olena

#7 - 07/13/2022 11:08 AM - Catherine Bunting

Excellent, thanks very much for your help.

Catherine

#8 - 10/20/2022 08:33 AM - Understanding Society User Support Team

- Status changed from Feedback to Resolved

- % Done changed from 90 to 100

#9 - 12/01/2022 06:37 PM - Understanding Society User Support Team

- Private changed from Yes to No