

Understanding Society User Support - Support #1664

IEMB sample when combining data to financial years

03/02/2022 06:15 PM - Dorothee Schneider

Status:	Resolved	Start date:	03/02/2022
Priority:	Normal	% Done:	100%
Assignee:	Olena Kaminska		
Category:	Weights		
Description			
Dear UKHLS team,			
I am trying to understand, based on Q11 and Q12 of the Weighting FAQ document, whether I can include the IEMB sample when pooling the data into financial years for cross-sectional analysis. Our population of interest are adults aged 18+ in England that have a common mental disorder (proxied by GHQ-12 score). I hope we can include the IEMB sample to help with numbers of observation when analysing by ethnicity.			
I've pooled the data following Q12: months 4 to 15 from wave n, months 16 to 24 from wave n-1 and months 1-3 from wave n+1. Q12 says "data ... can be combined for cross-sectional analysis, provided that each of the 24 monthly samples is included in the analysis base an equal number of times". In the case of financial years, each has 12 sample months with IEMB (months 13 to 15 from wave n, and months 16 to 24 from wave n-1) and 12 sample months without IEMB (months 4 to 12 from wave n, and months 1 to 3 from wave n+1), so is equivalent to the original waves. My assumption was that (after adjusting the weights following the code on p10) I can use all subsamples including IEMB.			
However in Q11, which is about calendar years or months, the advice is to exclude the IEMB because it is only part of months 13 to 24. For the BHPS and IEMB samples the advice is to use longitudinal weights to exclude them, but it seems one can use the Northern Ireland sample. Why is this, I don't understand the difference between these samples?			
I also don't understand the example given for Northern Ireland weight adjustment: "please note that if you use months 13-24 you are excluding Northern Ireland from your analysis. If you use months 1-12 Northern Ireland will be over-represented without an additional adjustment to the weight. Here is the Stata syntax for adjustment if you use month 1-12: (...)." This sounds like using only sample months 1-12 (i.e. year 1) without months from year 2, which I thought I understood from Q12 shouldn't be done? Else, if it means using months 1-12 as part of a dataset pooling year 1 + year 2 sample months from different waves, then why do the Northern Ireland cases need extra adjustment? I seem to be missing something which might also help me understand if I can include the IEMB in my analysis sample.			
Best wishes Dorothee			

History

#1 - 03/03/2022 08:50 PM - Understanding Society User Support Team

- Status changed from New to In Progress
- Assignee set to Olena Kaminska
- Private changed from Yes to No

Many thanks for your enquiry. The Understanding Society team is looking into it and we will get back to you as soon as we can.

We aim to respond to simple queries within 48 hours and more complex issues within 7 working days. While we will aim to keep to this response times due to the current coronavirus (COVID-19) related situation it may take us longer to respond.

Best wishes,
Understanding Society User Support Team

#2 - 03/14/2022 01:31 PM - Olena Kaminska

Dorothee,

If you include all sample months 1 through 24 in your analysis you don't need to worry about samples / subsamples. You should be fine. Taking all months 1-24 will represent the population fine, including NI and ethnic minorities. So, please continue with our ui_xw or ui_lw weights. Make sure not to subsample just IEMB or exclude it. Use full dataset and select subgroups based on population characteristics (not sample characteristics) in your analysis.

Hope this helps,

Olena

#3 - 03/18/2022 01:05 PM - Understanding Society User Support Team

- *Category set to Weights*
- *Status changed from In Progress to Feedback*
- *% Done changed from 0 to 50*

#4 - 06/06/2022 08:56 PM - Understanding Society User Support Team

- *Status changed from Feedback to Resolved*
- *% Done changed from 50 to 100*