Understanding Society User Support - Support #2258

Longitudinal versus cross-sectional weighting guidance

06/13/2025 12:56 PM - Michael Francis

Status:	Feedback		**Start date:**	06/13/2025
Priority:	Normal		**% Done:**	50%
Assignee:	Olena Kaminska			
Category:	Weights			

Description

Hi Olena,

I was wondering if you could possibly provide some guidance on weighting in my study in two parts. The first looks at access to training (e.g., trainany where trwho1 == 1, so employers), for different periods in the UKHLS.

To give some context: I am using data from waves f to n (2014-2022). When this is visualised using cross-sectional weights, to give a national annual representation, to compare with the LFS for example, the results are very different to the longitudinal weights (indinub_lw), which is to be expected. My time periods are coded as -1 (2014-2016), 0 (2017-2019) and 1 (2020-2022), as I am doing a Covid-study and testing longer-term trends. I opted for the pooled approach as annual estimates have more noise and may not be representative of a whole period. However, when the longitudinal weights are applied they drop a huge number of cases which are useful for the study as I am trying to test differences between different groups at that period in time as opposed to longitudinally over the whole period.

I used your guidance document: https://iserredex.essex.ac.uk/support/attachments/download/379/Worksheet%20ex6%20R.pdf , however, I don't think it covers my situation. Would it be better for me to consider using separate weights for each of the time-periods to mimic cross-sectional ones? The individual over time here is less important than the integrity of the sample groups (key workers and furlough groups) which are defined with SOC and SIC combinations.

So, in terms of my options - should I use longitudinal weights or cross-sectional weights from the last wave of each of those 3 periods, e.g., indinus? Or do I need to tailored weight based on the pool for each of the 3 periods?

Also - just a final clarification - what is the difference between indinus_lw and indinub_lw, as they have slightly different values for several waves which would be in the study? The description on the website is very similar - is it just that indinub_lw is from wave 3 and takes into account attrition?

Many thanks in advance,

Michael Francis

History

#1 - 06/13/2025 08:12 PM - Olena Kaminska

Michael,

Thank you for your question. The choice of a weight completely depends on the set up of your data and on what / whom you want to represent. You provided long description, but I would still need a few more details to know which weight you need to choose. The best approach for you is to think carefully whom you want to represent with each part of your model. Note, when comparing across time you need to ensure that you treat observations as dependent (they come from the same people) - this is slightly different to repeated cross-section approach.

I think you may find the answer here in question 14 https://www.understandingsociety.ac.uk/wp-content/uploads/working-papers/2024-01.pdf . Read question 6.6 to understand the difference between ub and us weights. Weight ub includes BHPS sample, so has higher sample size and more statistical power. If you start at wave f you should use ui weight.

Hope this helps,
Olena

#2 - 08/04/2025 04:35 PM - Understanding Society User Support Team

- Category set to Weights

- Status changed from New to Feedback

- % Done changed from 0 to 50

- Private changed from Yes to No