

Understanding Society User Support - Support #1705

Longitudinal Weighting of UKHLS Data in MLwiN

05/24/2022 02:02 PM - Sue Easton

Status:	Resolved	Start date:	05/24/2022
Priority:	Normal	% Done:	100%
Assignee:	Olena Kaminska		
Category:	Weights		
Description Hi, 1. Are there any reference documents on how/where to enter the UKHLS variables: PSU, strata and k_indinus_lw in order to weight the survey data for multilevel modelling in MLwiN please? A link to a simple "how to" guide with examples would be really useful if possible please (due to limitations of time). Thanks. Also: 2. If the longitudinal weight k_indinus_lw excludes GPS individuals who have missed waves over time in the series a-k (1-11), is there any point trying to "backfill"/impute data for their missed waves? Thanks! Sue Easton			

History

#1 - 05/24/2022 02:13 PM - Understanding Society User Support Team

- Category set to Weights
- 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.

Best wishes,
Understanding Society User Support Team

#2 - 05/25/2022 11:14 AM - Olena Kaminska

Sue,

Thank you for your question.

1. In any multilevel model you should be able to use a weight, sometimes stratification (but if you can't - excluding stratification makes estimates conservative - so not a problem). The cluster variable (PSU) should be one of your levels, usually the highest.
2. If you use k_indinus_lw you don't need to impute anything for missing waves.

Olena

#3 - 05/31/2022 06:46 PM - Understanding Society User Support Team

- Status changed from In Progress to Feedback
- % Done changed from 0 to 80

#4 - 06/07/2022 05:16 PM - Sue Easton

Hi Olena,

—
Thanks very much for your advice.

However, I've done a bit of research on the data, and from a *geographical* multilevel modelling perspective, I'm not sure it would be correct to enter the PSU as the highest level.

This is because the PSU units seem to be very geographically e.g. perhaps postcode districts or sectors.

So say one of my levels is region, then some PSUs are *within* Regions (judging from the clustering and numbers of PSUs by the Region variable for wave 1 participants). Therefore it wouldn't work spatially-statistically to have PSU as the highest level.

So I wonder if it would be better to use a cross-classified multilevel model to allow for the fact that individuals (within households), may be members of both a PSU and a different overlapping spatial category simultaneously. [And also to avoid using the Regional category].

There is also the issue of the complexity/parsimony of the model to allow for because I already have individuals within time:

1. Largest Area - Cross-classified?
2. PSU ?
3. (Household)
4. Individual
5. Year - LOWEST LEVEL

This is still rather a lot of levels for a Multilevel model!

Which may be very difficult to run and complex to interpret.

Thanks for your input.

Sue CE

#5 - 06/08/2022 01:52 PM - Olena Kaminska

Sue,

You have to run a correct model with correct specification, though you don't need to interpret all the levels that you specify. Your analysis is quite unique in the way that you use higher level than PSU, but yes, you PSU needs to be one of the levels, and you of course can have a higher one, like region.

Households would be nested within PSUs. And PSUs in GB are postcode sectors (sometimes old ones as selection was done in 1991, 1999 and 2007) - so some changes could have occurred, but I suspect most or all will be nested within regions. You should check this. If a tiny proportion is out, just recode it - I don't think cross-classified model is needed unless there is a large overlap.

Depending on what you study you can avoid household - it is not a longitudinal concept for example...

Also note, that for NI PSU=household.

Hope this helps,
Olena

#6 - 06/08/2022 02:14 PM - Sue Easton

Hi Olena,

Thanks very much for confirming that PSU is Postcode Sector - I couldn't find this in the User Guides/documentation on Weighting.

Northern Ireland and Scotland are not part of my analysis for various reasons such as differential geographical classification and health service data availability.

Thanks again for your help. Much appreciated.

Sue CE

#7 - 06/08/2022 03:50 PM - Olena Kaminska

To find more information on PSUs you should look into sampling documents for UKHLS and BHPS.

Olena

#8 - 10/24/2023 09:45 AM - Understanding Society User Support Team

- Status changed from Feedback to Resolved

- % Done changed from 80 to 100