

Understanding Society User Support - Support #1374

Weights

07/07/2020 10:50 AM - Maria Cotofan

Status:	Resolved	Start date:	07/07/2020
Priority:	High	% Done:	100%
Assignee:			
Category:			
Description			
Dear Sir/Madame,			
I have a question regarding the use of weights when pooling the 9 waves of Understanding Society. At the moment, I simply want to pool all the 9 waves and perform a cross-sectional analysis. Later on, I would like to also use the panel dimension of the data.			
I then have two questions about the use of weights:			
(1) For the cross-sectional analysis, since I am only using the 9 waves of Understanding Society, can I simply append all the waves and use the associated <code>_xw</code> weights for each wave (namely <code>indinui_xw</code> for waves 6 to 9, <code>indinub_xw</code> for waves 2 to 5, and <code>indinus_xw</code> for wave 1) to create a new weight, say <code>weight_xw</code> ? Or would I have to perform any additional transformations? And lastly, if this is true, do I understand correctly that the data should be used in combination with the <code>psu</code> (for example, in Stata: <code>svyset psu [pweight=weight_xw]</code>), and that using the strata information is less important in this case?			
(2) When using the panel dimension of the data to perform, for example, a fixed effects analysis, is only the <code>_lw</code> weight in the most recent wave relevant? In other words, I would merge all 9 waves of understanding society using the respondent id, transform the data into long format, and in all analyses use the <code>i_indinui_lw</code> .			
Would you say that this approach is correct? Thanks you in advance for you advice!			
Best, Maria			

History

#1 - 07/08/2020 06:05 AM - Alita Nandi

- Assignee set to Maria Cotofan
- % Done changed from 0 to 80

Hello,

- (1) Take a look at item 12 in the weighting FAQ which discusses pooling waves for cross-sectional analysis.
https://www.understandingsociety.ac.uk/sites/default/files/downloads/documentation/user-guides/mainstage/weighting_faqs.pdf
- (2) Yes

Best wishes,
Alita
On behalf of Understanding Society User Support Team

#2 - 07/08/2020 06:05 AM - Alita Nandi

- Private changed from Yes to No

#3 - 07/08/2020 11:43 AM - Maria Cotofan

Dear Anita,

Thank you for your quick reply!

Regarding point (1), I was already familiar with item 12 in the FAQ, but I'm not sure it fully answers my question. If I understand correctly, the example discusses how to pool data from one single calendar year using multiple waves ("For example, for a financial year (April to March), months 4 to 15 from wave n can be combined with months 16 to 24 from wave n-1 and months 1-3 from wave n+1. And equivalently for any other period that is a multiple of 12-months.").

However, I would like to be able to pool the data from all the 1 to 8 waves and look at it as a pooled cross-section. I thought this previous discussion was relevant to my question: <https://iserswww.essex.ac.uk/support/issues/1257>. The first answer states that "for example observations from wave 1 would have wave 1 weight, and observations from wave 2 would have wave 2 weight and so on. It would be good to scale the weights [...] but this is

less important if you are using only UKHLS."

So to sum up, I don't want to pool different waves within the same calendar year, but rather use all available waves/years in the UKHLS and treat them as a pooled cross-section. In that case:

- (1) Can I simply use the weight in each wave and create a new weight as suggested here: <https://userswww.essex.ac.uk/support/issues/1257> ?
- (2) Is scaling still an issue if I only use UKHLS data and do not include the BHPS?
- (3) Would it be problematic if each wave has all the 24 months of data in the pooled cross-section?

I hope I was able to clarify my question and thank you again in advance!

Best,
Maria

#4 - 07/09/2020 10:27 AM - Alita Nandi

- Assignee changed from Maria Cotofan to Olena Kaminska

ok. I am now assigning this to our Survey Statistician, Olena.

#5 - 07/09/2020 02:47 PM - Olena Kaminska

Maria,

Thank you for your questions. I am replying in order:

In your original message:

- (1) Yes, your set up is correct - use the most relevant xw weight per wave. Ideally it would be good to scale, but it is less important if you start at wave 1 than if you start in 1991. Also, yes, always correct for clustering (psu) and weighting. Stratification is indeed optional - it is to your advantage and without it your results are more conservative. Your results and conclusions will be correct if you omit stratification though.
- (2) The weight you choose depends on what you study rather than the set up of your data. If for example you study between wave change (always using 2 waves for your analysis), you should use weight b_ for waves a-b combination, weight c_ for waves b-c combination etc. This logic extends to any number of wave-combinations you use - always use the lw weight from the last wave for each wave combination.

In your follow up message:

You are right about the relevance of FAQ and [#1257](#) - the latter is more relevant to you.

- (1) Yes;
- (2) Don't exclude BHPS when looking at cross-sectional weights in UKHLS - they include BHPS. Just use the data as you described in (1) earlier. Scaling is more important if your analysis starts before wave 1 of UKHLS. Scaling is to compensate for the difference in sample size across waves.
- (3) Our data is designed to be used in 24 months chunks - so this makes perfect sense.

Hope this helps,
Olena

#6 - 07/10/2020 10:31 AM - Maria Cotofan

Dear Olena,

This perfectly answers my question. Thanks so much for the clarification!

Best,
Maria

#7 - 07/15/2020 03:35 AM - Alita Nandi

- Status changed from New to Resolved

#8 - 10/13/2021 11:33 AM - Understanding Society User Support Team

- Assignee deleted (Olena Kaminska)

- % Done changed from 80 to 100