

Understanding Society User Support - Support #1590

Creating a weight variable for fixed effects analysis

10/08/2021 02:22 PM - Karen Mak

| | | | |
|--|----------|--------------------|------------|
| Status: | Resolved | Start date: | 10/08/2021 |
| Priority: | High | % Done: | 100% |
| Assignee: | | | |
| Category: | Weights | | |
| Description | | | |
| Dear UKHLS support, | | | |
| Hope you are well. | | | |
| May I ask for your advice on whether you think weighting should be applied in a fixed-effects analysis? | | | |
| I would like to use data from Waves 2, 4, 6, 8, and 10. From my understanding, if we use the longitudinal weight provided in the Wave 10 data (<code>j_indscub_lw</code>), it will only retain the sample who completed waves from Waves 2 to 10, which could lead to a substantial drop in the sample size. | | | |
| I have read the paper "Weighting and Sample Representation: Frequently Asked Questions", which was really useful and clear, and learnt that we can derive our own weight variable. However, I am not sure how to do so with a longitudinal dataset (in a long Stata format). | | | |
| I would be very grateful to have your advice on (i) whether or not weighting is necessary for a fixed-effects model, and (ii) if so, what would be the best way to derive a weight variable for this study. | | | |
| Best wishes, Karen | | | |

History

#1 - 10/08/2021 03:21 PM - Understanding Society User Support Team

- Status changed from New to Feedback

- % Done changed from 0 to 50

- Private changed from Yes to No

(1) This is an analysis question. This is a good discussion about why and when we should weight. <http://jhr.uwpress.org/content/50/2/301.refs>

(2) Please email us at usersupport@understandingsociety.ac.uk and we will send you the guidance for producing your own weights.

#2 - 10/08/2021 03:45 PM - Olena Kaminska

Karen,

Yes, weighting is necessary in fixed effect models.

Look for similar questions about pooled analysis on this forum - they may answer your question.

If you need more help, please give us details on what you want to estimate (e.g. people or events etc.)

Thanks,
Olena

#3 - 10/08/2021 04:11 PM - Karen Mak

Hi Olena,

Thanks so much for your prompt reply and for your help.

I tried to look for similar questions on this forum, but couldn't seem to find any solutions.

I am interested in how changes in volunteering behaviours are associated with changes in wellbeing. Questions on volunteering were asked in alternative waves (i.e. Waves 2, 4, 6, 8 & 10). I have now merged/append all relevant waves into one dataset in a long format.

(1) I'd like to ask whether it would be more appropriate to use the wave 10 longitudinal weight (`j_indscub_lw`), or to create a specific weight for the

analysis?

(2) If creating a new weight is more appropriate here, what would be the steps to create one in a long data format? Do I need to reshape it to wide format -> then create a binary variable "response" (1=completed all 5 waves, 0=only completed Wave 2) -> run a logistic regression predicting "response" using predictors (e.g. age, gender, marital status) on a condition that participants not known to have died/emigrated and that participants have a Wave 2 weight value greater than 0 -> then generate a new weight = $\text{gen weightW25} = (1/p)*b_indscus_lw$?

Thank you and best wishes,
Karen

#4 - 10/11/2021 03:03 PM - Olena Kaminska

Karen,

There is no particular reason to create your own weight.
To help you select a weight, please explain how you analyse the data: are you looking at a change between 2 waves, or 3 waves at a time?

Thanks,
Olena

#5 - 10/11/2021 03:45 PM - Karen Mak

Hi Olena,

Thank you for your reply.

I am looking at a change between 5 waves (Waves 2, 4, 6, 8, and 10) at a time using fixed effects modelling. Is it correct to choose `j_indscub_lw` as the weight for my analysis?

Best wishes,
Karen

#6 - 10/12/2021 09:59 AM - Olena Kaminska

Karen,

Do you mean you study change between wave 2 and wave 10?
Or do you mean you study change between wave 2 and 4, and separately between 4 and 6 etc.
What is your outcome variable?

Thanks,
Olena

#7 - 10/12/2021 10:02 AM - Karen Mak

Hi Olena,

So sorry for the confusion. We are interested in studying changes between waves 2 and 4, and separately between 4 and 6 etc. The outcome variables are GHQ12 and SF12.

Best wishes,
Karen

#8 - 10/13/2021 10:24 AM - Olena Kaminska

Karen,

In this situation use longitudinal weight for wave 4 for the set of waves 2 to 4 outcome, lw weight for wave 6 for the set of waves 4 to 6 outcome etc. In a long format you will create a new weight variable and give it respective values.

Hope this helps,
Olena

#9 - 10/13/2021 10:38 AM - Karen Mak

Hi Olena,

Thanks so much for your reply and incredible help!

I think I might have misunderstood your question earlier. Just wanted to confirm with you that if I wanted to explore the changes from Wave 2 to 10, I should be using Wave 10 lw weight?

Best wishes,
Karen

#10 - 11/04/2021 12:33 PM - Olena Kaminska

Karen,

Yes, that's correct.

Olena

#11 - 11/05/2021 09:17 AM - Understanding Society User Support Team

- *% Done changed from 50 to 90*

#12 - 12/14/2021 01:34 PM - Understanding Society User Support Team

- *Status changed from Feedback to Resolved*

- *Assignee deleted (Olena Kaminska)*

- *% Done changed from 90 to 100*