

Understanding Society User Support - Support #1008

Creating longitudinal weights

07/24/2018 11:07 AM - Joseph Woods

Status:	Resolved	Start date:	07/24/2018
Priority:	Normal	% Done:	100%
Assignee:	Joseph Woods		
Category:	Weights		
Description <p>I'm doing analysis using harmonised BHPS and USoc to look at the impact of tax / benefit policy over several years. I would like to include in the final sample people who are in the survey for most, but not necessarily all, waves (e.g. include those who are present for at least 20 out of the 25 waves). For those who attrit, these people don't have longitudinal weights so I need to design my own. I have been trying to replicate the BHPS/USoc longitudinal weights first, and then after that am planning on doing a similar exercise for people in most but not all waves. However, I've been unable to get particularly close to the longitudinal weights in the data. I am using the CHAID process and post stratifying the data.</p> <p>1: Do you have more information on how the longitudinal weights are constructed? E.g. what predictor variables are used? 2: Do you think this is the best way to proceed? An alternative approach – which has been discussed in other forum posts - would be to fit a model (e.g. logit) and adjust the cross-sectional weights accordingly.</p>			

History

#1 - 07/24/2018 01:15 PM - Stephanie Auty

- 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.

Best wishes,
Stephanie Auty - Understanding Society User Support Officer

#2 - 07/30/2018 12:40 PM - Olena Kaminska

Joseph,

If you design your own nonresponse correction I would expect the 'weights' to look different to ours but importantly to lead to largely similar results. If you receive a very different result with your own nonresponse correction I would expect that something went wrong in the process. Please make sure to incorporate wave 1 weight when designing your own weights.

If you are using harmonized UKHLS and BHPS data over 20 years I assume you are not using longitudinal data but using cross-sectional data through pooling. In this situation you won't achieve much gain from using your own weights in comparison to ours. We describe the method of calculating our weights in the User Guide. Unfortunately at the moment we do not have a list of variables that go into each model, but please note that over 25 years there will be at least 51 models that led to the weights that you are using - as we calculate weights using the information from the previous wave.

In post-stratification also be careful to match to the statistics on longitudinal population rather than on cross-sectional population that most government statistics is available for.

Also, it largely doesn't matter which method you use to construct the correction - as long as the method is correct and all the parts of the nonresponse and unequal selection probabilities are taken into account it should be fine and yield largely the same results.

Hope this helps,
Olena

#3 - 07/30/2018 02:30 PM - Stephanie Auty

- Status changed from New to Feedback
- Assignee changed from Olena Kaminska to Joseph Woods
- % Done changed from 0 to 70

#4 - 08/03/2018 10:23 AM - Joseph Woods

Many thanks Olena.

For this analysis I am looking to exploit the longitudinal aspect of the data. Essentially, I'm running each wave through a microsimulation model which allows me to simulate a hypothetical policy reform, and I then follow the impact of this reform on an individual's income over X years to calculate the average impact over this time horizon. (The dataset outputted is long data format, where each row is a unique adult-year combination).

As things currently stand, my sample is limited to adults who appear in BHPS wave 1 and have not attrited by the latest USoc wave - these people have non-zero longitudinal weights in the data. Ideally, I would like more flexibility to include additional observations and boost sample size. For instance, if you first appear in BHPS wave 2 and do not attrit, your longitudinal weight in the latest USoc wave is zero. I would like to create my own non-zero weight for this type of person.

Hopefully this clears up my original post.

Thanks, Joe

#5 - 08/06/2018 01:02 PM - Stephanie Auty

- Assignee changed from Joseph Woods to Olena Kaminska

#6 - 08/15/2018 12:41 PM - Stephanie Auty

Apologies for the delayed response. Our Weighting team are currently all on annual leave and will respond when they return.

Best wishes,
Stephanie Auty - Understanding Society User Support Officer

#7 - 09/12/2018 11:57 AM - Olena Kaminska

Joe,

I see the problem you are facing and I see the reasoning behind it. I will make two further suggestions:

1. If you create your own weights - make sure that you think about the population definition that you want to represent, especially in terms of birth and death. Do you want to represent people you live in 1992 but died since? Keep this in mind when you are describing the dataset and when you are talking about your results.
2. Run the results with the new weights and compare them to the results obtained with our weights. The expectation is that the model values won't change but the confidence intervals /may/ (though don't have to) get smaller - so you may find significant results where the previous results were marginal.

Best of luck,
Olena

#8 - 09/12/2018 12:15 PM - Stephanie Auty

- Assignee changed from Olena Kaminska to Joseph Woods

- % Done changed from 70 to 80

#9 - 10/12/2018 02:34 PM - Stephanie Auty

- Status changed from Feedback to Resolved

- % Done changed from 80 to 100