

## Understanding Society User Support - Support #2042

### Survey Weights for Multi-Wave Pooled Analysis

01/29/2024 09:47 AM - Lisa Waddell

<b>Status:</b>	Feedback	<b>Start date:</b>	01/29/2024
<b>Priority:</b>	Normal	<b>% Done:</b>	80%
<b>Assignee:</b>	Olena Kaminska		
<b>Category:</b>	Weights		
<b>Description</b>			
Hello,			
I have constructed an unconventional sample by pooling tab-delineated data files for SN 6614-Understanding Society: Waves 1-13, 2009-2022 and Harmonised BHPS: Waves 1-18, 1991-2009. I request your advice regarding these weights.			
Sample Construction: Using the family matrix, I identify everyone in the sample with both a mother and a father pidp identified. Using all waves of data, I keep participants whose mother and father both responded when the participant was aged 10 or younger. I then filter by participants who responded at the age of 21 or older. These two filtering functions leave me with a sample of around ~2000 people between the ages 21-41, from BHPS and USoc samples. Due to my pooling of BHPS and USoc samples, when I follow the steps for constructing a tailored sample weight, I lose a substantial portion of my sample. For example, if I choose a base weight from Wave 1 of USoc, I lose the entire BHPS sample. If I choose a base weight from Wave 2 of USoc, I lose a substantial portion of the USoc sample.			
Given how I construct my sample, do you have any advice on how I should be applying survey weights?			
All the best, Lisa			

#### History

##### #1 - 01/29/2024 11:05 AM - Understanding Society User Support Team

- Status changed from New to In Progress
- Assignee changed from Understanding Society User Support Team to Olena Kaminska
- % Done changed from 0 to 10
- 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 - 01/30/2024 01:09 PM - Olena Kaminska

Lisa,

Thank you for your question. Yes, you should use tailored weights, accounting for nonresponse of both parents at at 21 or after. Your situation is a bit complex because you pool information from any age of before 10 (unlike just from age of 10). The best weight therefore would be the weight at the time of entrance to the study of the child: it's either at the time of selection (1991, 1999, 2001, UKHLS wave 1 or 6) or at the time of birth. Your starting pool should be all children (regardless of whether they live with parents at the time of observation). This is because some children enter the study at the age of 10, and their history of living with both parents is missed. From this pool of all children (under 10 at any point of time from 1991 to wave 2 (I think)) you should model a chance of them to have lived with both parents, who also responded, and also the person responded after the age of 21 - through one model.

Your challenge would be to find the same predictors from across all waves of BHPS and UKHLS, so the number may be a bit limited. Alternatively, you could run separate models for some sets of children.

I suggest scaling of weights, because some cohorts will be better represented than others. But unlike in other situations, I suggest you scale the weights by year of birth, adjusting it to that of official statistics, such that proportion of children from each year of birth reflect the population proportions.

Hope this helps,  
Olena

**#3 - 02/12/2024 08:42 PM - Understanding Society User Support Team**

- *Status changed from In Progress to Feedback*

- *% Done changed from 10 to 80*