

Understanding Society User Support - Support #881

weighting values of zero

11/21/2017 11:51 AM - Andrew Brown

Status:	Closed	Start date:	11/21/2017
Priority:	High	% Done:	100%
Assignee:	Andrew Brown		
Category:	Weights		
Description			
Hi			
Could I ask a question related to issue #877 ?			
For those cases not assigned a weight because 'a person in a household where there is no person who has been enumerated at every wave up to wave w will get a weight of zero. Such people should not be given a weight, as the weights for all other sample members are calculated in a way that compensates for these "missing" people'			
How could/ should they be included in any analysis as SPSS 'makes them invisible' - could they be assigned the mean weight of 1? or should they be excluded from any analysis as the weighting for the other cases takes account of this?			
Many thanks			
Andrew			

History

#1 - 11/21/2017 01:34 PM - Stephanie Auty

- Assignee changed from Peter Lynn 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.

Best wishes,
Stephanie Auty - Understanding Society User Support Office

#2 - 11/21/2017 04:40 PM - Olena Kaminska

Andrew,

They should not be included in the analysis, as weighted analysis will compensate for them being missing.
You should not assign them any other value than zero - otherwise your analysis will not represent the population any more.

Best wishes,
Olena

#3 - 11/21/2017 10:30 PM - Andrew Brown

Dear Olena

Many thanks for your swift reply. I wonder if you would be able to provide further clarification regarding correct weights to use before I plough on with any analysis.

Essentially my project hopes to compare SDQ responses in the Youth Self-completion survey between adopted and non-adopted children. I have identified 37 adopted children over waves 1, 3 and 5 (when the SDQ was included in the questionnaire). If children have responded to the SDQ in multiple waves I have taken responses from their earliest wave. I have used Propensity Score Matching to create a matched comparison group (5 per adopted child, random selection) from those not identified as adopted and who have completed the SDQ. Note that I am not comparing changes in individual scores over waves.

My thinking is that I should use the 'a_ythscus_xw' weight for those who responded in wave 1; 'c_ythscub_xw' for those who responded in wave 3; and 'e_ythscub_xw' for wave 5 responders. Cases that have a zero value in the weight will be automatically excluded by SPSS and the sample size that I report is a weighted *n* as given in the output? (Because the weight compensates for them being missing?)

I hope this makes sense and eagerly look forward to your reply

Kind regards

Andrew

#4 - 11/22/2017 11:59 AM - Olena Kaminska

Andrew,

Your situation is one of the complex ones for which we don't have ready weights. You could use your approach if your selection of people was not based on response / nonresponse (you are not included responses from wave 3 if people already responded in wave 1). For this reason you can't use the cross-sectional youth weights as you suggested. Instead you could use a suboptimal weight which would be a longitudinal enumeration weight from wave 5 for all people. Ideally you would then want to additionally correct for nonresponse to youth questionnaire conditional on being enumerated (having non-zero `e_psnenus_lw`). You may even gain some numbers in this way as well.

Hope this helps,
Olena

#5 - 11/22/2017 04:33 PM - Andrew Brown

Dear Olena

Thanks again for such a quick reply. So just to double check, I shouldn't use the cross wave weights as I suggested as my selection is based on cases responding or not in earlier waves.

Presumably the `e_psnenus_lw` is the wave 5 longitudinal weight to use.

What is the best way to additionally correct for non response, could you provide some guidance here?

Many thanks again

Kind regards

Andrew

#6 - 11/23/2017 12:32 PM - Andrew Brown

P.S. After looking through the weighting section of the user guide I'm wondering if an appropriate weight to use would be:

For wave 1 cases: `e_psnenus_lw * a_ythscus_xw`

For wave 3 and wave 5 cases: `e_psnenus_lw * n_ythscub_xw`

I suspect it's not as straightforward as this but would greatly appreciate your continuing advice and expertise on this.

Kind regards

AB

#7 - 11/27/2017 03:18 PM - Olena Kaminska

Andrew,

No, you can't just multiply weights - it does not work like that.
Yes, the enumeration weight to use is `e_psnenus_lw`.

Considering your numbers I wonder how much you will lose if you simply use wave 1 information - in this situation you can just use the youth weight. In any case I recommend that you run simple wave 1 analysis (with the youth weight) and the complex one with more people and compare the results - the results shouldn't be too different (so run just as a check).

To correct for nonresponse I suggest you run logistic regression to predict response (versus not) - make sure to condition on `e_psnenus_lw` - and multiply the predicted values by the enumeration weight `e_psnenus_lw`. But there are many other options available.

Hope this helps,
Olena

#8 - 11/27/2017 10:30 PM - Andrew Brown

Dear Olena

Yes, this helps greatly!

I was thinking that a simple analysis of just the wave 1 data might be a suitable solution - that would give an unweighted *n* of 22 adopted children. My original thinking was that 22 was rather small and so looked to the other waves to boost the number of adopted children.

When correcting for nonresponse in the more complex analysis involving responses in multiple waves, I'm wondering how best to construct the dataset in order to run the regression (is the regression run in each wave separately, or are the waves merged?).

I would also be interested to hear about other possible options.

Many thanks for your time (and patience!)

Kind regards

Andrew

#9 - 11/28/2017 12:21 PM - Stephanie Auty

- *Status changed from New to In Progress*

- *% Done changed from 0 to 60*

#10 - 01/02/2018 12:50 PM - Olena Kaminska

Andrew,

This is how Peter Lynn explained adjusting the weights in his reply to issue num 703.

"... This would involve fitting a model (e.g. logit) based on all year X cases in your sample, in which the dep var is a 0/1 indicator of whether they also responded at year Y (and removing from the base any known to have died or emigrated before Y). Predictor variables can be anything relevant observed at X. This will give you a predicted probability for every year X respondent of responding at year Y. Call this P. You then need to adjust the year X weight by multiplying it by 1/P for all the cases that can be included in your analysis (those who responded in both years)."

What Peter talks about as year X weight is e_psnenus_lw in your situation.

Hope this helps,
Olena

#11 - 01/03/2018 10:50 AM - Stephanie Auty

- *Assignee changed from Olena Kaminska to Andrew Brown*

#12 - 02/12/2018 12:34 PM - Alita Nandi

- *Status changed from In Progress to Resolved*

- *% Done changed from 60 to 90*

#13 - 03/07/2018 05:44 PM - Stephanie Auty

- *Status changed from Resolved to Closed*

- *% Done changed from 90 to 100*