Understanding Society User Support - Support #2136

Weights for multiple imputation in Mplus

08/06/2024 04:24 PM - Ivelina Hristova

Status:	Feedback	**Start date:**	08/06/2024
Priority:	Normal	**% Done:**	80%
Assignee:	Olena Kaminska		
Category:	Weights		

Description

Dear UKHLS team,

I am estimating a latent model in Mplus. I would like to do multiple imputation to fix the pattern of missingness.

Unfortunately, accounting for complex survey design is not currently available with multiple imputation in Mplus. In their forum, they recommend the following approach: 'If you have to do MI and the above alternative is not an option you can use 3-level imputation where the weight variable or a log of it is added as a variable (so information from that variable can be used) during the imputation process, in addition to any other variables used for constructing the weights.' [https://www.statmodel.com/discussion/messages/12/26948.html?1556842127]

I use the h_indscui_lw. What are the variables used for constructing this weight?

I look forward to your reply.

Kind regards,

Ivelina

History

#1 - 08/07/2024 09:47 AM - Understanding Society User Support Team

- Category set to Weights

- Assignee changed from Understanding Society User Support Team 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. 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 - 08/13/2024 12:29 PM - Understanding Society User Support Team

- % Done changed from 0 to 10

#3 - 08/19/2024 11:01 AM - Understanding Society User Support Team

- Status changed from New to In Progress

#4 - 09/06/2024 03:30 PM - Olena Kaminska

Ivelina,

It's likely that you won't be able to include all the variables used in our weighting model as there are too many of them (see below). But more importantly, if you want to correct for missingness using multiple imputation you would need to start at wave 1 (where non-contacted households are not interviewed but still need correction). Between 30% and 60% of total nonresponse comes from wave 1. Predictors from interviews won't give you information to correct for this missingness.
Another problem is a need to adjust for unequal selection probabilities. The difference in these probabilities is relatively large in UKHLS, and without account of them you have too many minorities, too much NI, and too many households selected recently, too few recent immigrants etc. You can't correct this with imputation, because this is not missingness.

My suggestion is to look whether you can start from cov-var matrices rather than data for your analysis. These could be obtained from UKHLS weighted and should be possible to feed into Mplus, though not sure how it works now. Long time ago I used LEM and could feed in matrices directly.

On variables used. These are the sets of variables used in h_indscui_lw:
- Census, other geographically linked variables, and sampling frame variables from 1991, separately from 2001, and from 2011 (note, there are 3

Censuses in UK, so separate variables for each);
- Variables from indall, hhresp and inresp from all waves of BHPS and all waves of UKHLS up until and including wave g. Our correction is wave on wave, so variables come from each wave.
- There is also post-stratification to external geographical variables at some points in time, and mortality correction (the latter has its own modelling, though covariates are similar to the above).

I worked on UKHLS weights, and not on BHPS weights. For UKHLS if you open the data and look at any question asked to everyone, it's in the model (unless excluded for collinearity). Only significant variables are included in each nonresponse model.

Hope this helps,
Olena

#5 - 09/17/2024 05:39 PM - Understanding Society User Support Team

- Status changed from In Progress to Feedback

- % Done changed from 10 to 80