

Understanding Society User Support - Support #2160

create a tailored weight

10/10/2024 03:06 PM - Shiyu Yuan

Status:	Feedback	Start date:	10/10/2024
Priority:	Urgent	% Done:	100%
Assignee:	Olena Kaminska		
Category:	Weights		
Description			
Hi there,			
I am currently working on creating a tailored weight for my unbalanced panel dataset, following the guidance provided on Moodle. However, I am encountering some challenges and would greatly appreciate your insights.			
My sample includes data from waves b, d, f, h, j, and l, where I have combined male and female partners into couple-level datasets. My focus is on childbirth, utilizing the newborn variable from the next couple-wave record as the dependent variable to account for the lagged effect. Since this is an unbalanced panel, I excluded observations after childbirth and include couples who exist in any three waves, rather than those present in all waves, which yields a very limited sample size.			
Here are a few specific questions I have:			
1. Base Weight Selection: Which wave should I use as the base weight? Would it be acceptable to choose wave f, given my sampling strategy?			
2. Mortality Adjustment: I understand that we need to account for mortality based on the last wave used. However, I could not find the <code>mortus_tw</code> variable after wave j in the <code>xtwavedat</code> file. Could you guide me on where to find this or how to approach mortality adjustments for my sample?			
3. Non-Response Considerations: When defining non-response, should I consider all variables of interest or focus solely on the dependent variable? What would be best practices for determining which predictors to include in the non-response model if I consider all variables of interest?			
4. General Guidance: As this is my first time creating a tailored weight, I would appreciate any additional tips or resources you might recommend for this process.			
Thank you for your help, and I apologize for any confusion in my questions.			

History

#1 - 10/11/2024 12:54 PM - Olena Kaminska

Shiyu,

Thank you for your question.

1. For base weight it is best to use the first wave in your analysis, wave b in your situation. This will lead to the highest sample size.
2. Mortality adjustment can be ignored, as your population is very young (parents of newborns), and mortality will be very small in this subgroup and won't influence weights much;
3. Predictors related to what you are interested in (likely relationships between variables in your situation), including all variables of interest. It is best to start with as many as possible variables but include only significant variables in the model, e.g. through a stepwise process.
4. Any general information on logistic regression will be helpful.

Hope this helps,
Olena

#2 - 10/12/2024 08:13 AM - Shiyu Yuan

Hi Olena,

Thank you for the help. It is very helpful.

I would like to double-check, in this case, am I correct in choosing wave `b_indinub_xw` as the base weight and then creating a response variable for those missing in my variables of interest and then running the logit regression for considering non-response?

Thank you for your time.

best,
Shiyu

#3 - 10/14/2024 10:49 AM - Olena Kaminska

Shiyu,

Yes, this sounds correct.

Olena

#4 - 10/16/2024 12:11 PM - Shiyu Yuan

Hi Olena,

Thank you very much for your help!

I just wanted to double-check a couple of things after following your advice regarding the definition of the response variable and the assignment of the tailored weight to each sample:

Since my focus is on couples, I multiplied the start weight for both partners. If one partner weights zero while the other has a non-zero weight, I adjusted it so that the couple's weight is set to the non-zero weight. Is this okay?

I aim to include couples who participated in at least two waves in addition to wave 1. To create the response variable, I defined it as couples who answered any two waves after wave 1 and did not have missing data for my key variables in the waves they participated in. After running the logistic regression, I found that the adjusted weight only has values in the first wave, while it's set to zero in all subsequent waves. Am I correct in assuming that the weight from the first wave should be applied across all waves, meaning each couple would have a unique weight for the entire study period? For example, if Couple A has an adjusted weight of 1 in wave 1, I would assign the same weight of 1 for the rest of the five waves. I'm unsure if this is the correct approach.

Could you kindly confirm if my understanding is correct or provide guidance if I'm mistaken?

Thank you again for your assistance!

Best,
Shiyu

#5 - 10/17/2024 12:44 PM - Olena Kaminska

Shiyu,

Couples are a bit complicated entity to study, but here are a few suggestions.

UKHLS represents people longitudinally. At any particular wave, e.g. wave b, some couples have formed, some didn't and some will never be together though they may have a child. You need to define whom you want to study (without thinking which waves they participated in. You could:

1. study by an individual and see when they form the couple. Each individual counts separately, so an already formed couple counts twice (for each separate individual). The analysis would be at an individual level, talking about individuals in couples. This is one of the easiest set up and you can use our weight from wave b as a base weight and make the usual adjustment for nonresponse. Individuals with a weight of 0 should be out of analysis even if their partner has a positive weight.
2. You mention that you are interested in childbirth, so you could centre your analysis around the birth of the first child, for example. You then take the child's weight and study couples per child. You can represent children and the parent situation around them. In this situation use child's enumeration weight as a base weight and model nonresponse from it. You will be representing children.
3. Couples are complicated, if you want them as an entity. This is because a couple made of two OSMs has twice lower selection probability than a couple with only 1 OSM. Couples are not stable longitudinally either, so you need to define exactly what you want to represent, both the starting point of formation of a couple and how definition depends on their longitudinal development of their relationship. Note, in this definition nothing should be related to their response or participation in UKHLS. Definition should depend on things in a population.

I suggest you use one of the first two definitions and see where it takes you.

Thank you,
Olena

#6 - 10/24/2024 12:40 PM - Understanding Society User Support Team

- *Status changed from New to Feedback*

- *% Done changed from 0 to 100*

- *Private changed from Yes to No*