Subject: Timing of variables' collection
Posted by gebretsh@gmail.com on Mon, 28 Nov 2022 13:30:23 GMT

Dear Tom,
I really appreciate your prompt response. Thanks
I thought about the problem of adding an extra cluster to the nature of the data. I just left the idea and started working
on a different methodology to my study.

Now, I would kindly like to bring to your kind notice a couple of points I encountered when I prepare (for analysis) the EDHS data in relation with my
research objectives:

1) The occupation variables (both maternal and her partner's) in both IR and KR files were collected in the 12 months preceding the surveys. The ANC and SBA services, on the other hand, were collected five years before the survey, and the PNC service two years before the survey. Given this, is it conceptually appropriate (statistics may produce results) to use occupation as a potential predictor variable for ANC, SBA, and PNC services? Occupation (the presumed cause) comes long after the outcomes have already occurred.
2) In the KR file, I could not find the " place of antenatal care" variable. Is it possible that this variable was not collected in the first place?

Subject: Re: Timing of variables' collection
Posted by Bridgette-DHS on Tue, 29 Nov 2022 13:19:28 GMT

Following is a response from Senior DHS staff member, Tom Pullum:

We often use models in which the temporal sequencing of the variables is inconsistent with causality. I think it's important to be aware of that and say something about it in whatever you write up. The usual justification is that many characteristics don't change much over time, so a value that is too old or too recent can be used as a proxy for the value at the time when you would like to have it. Type of place of residence and occupation are examples. You can think of the difference between the measured value and the "true" value as measurement error that will dilute or attenuate the association. There are some sophisticated methods to adjust for such things, but (personally) I think there are so many other kinds of measurement error in the data that those methods would just give a false sense of confidence in the results. This is a philosophical issue. When the variables are out of sequence, I try to interpret the results in terms of association rather than causation.

I think m57a-x is the set of variables on place of antenatal care that you are looking for. This is a multiple option type of variable--more than one option can be selected. Some options in the standard set are NA.

Subject: Re: Timing of variables' collection
Posted by gebretsh@gmail.com on Tue, 29 Nov 2022 14:22:38 GMT
View Forum Message <> Reply to Message

Thanks so much.
This is a very helpful suggestion.
With regard to the place of antenatal care variables, I did not find them in KR file, but they are available in the IR files.

Thanks