Subject: Which data file should I use for nutrition outcomes of children under 5 years of age
Posted by m.hasan1 on Sun, 21 Apr 2013 04:47:42 GMT
View Forum Message <> Reply to Message

Hello,

I would like to have some advice on the following issues:

(1)If I want to run a regression model to examine the effect of different determinants on nutrition outcomes of children under 5 years of age, can I use only the children data file (BDKR file), Or I have to use/merge other files such as births data file?

(2)What is the standard method of dealing the flagged cases? Is it important to include flagged cases during the analysis?

Subject: Re: Which data file should I use for nutrition outcomes of children under 5 years of age
Posted by Reduced-For(u)m on Wed, 01 May 2013 01:12:02 GMT
View Forum Message <> Reply to Message

I think a DHS mod will help here, but I have an ulterior motive...

1) I use the child recode too, but you may have to use the person or household member recode to get the exact numbers the DHS calculates. There is a post here with instructions and code: http://userforum.measuredhs.com/index.php?t=tree&th=137& amp;goto=262&#msg_262

2) Which flags? Age flags? I would consider dropping those, since HAZ is computed including age, but there shouldn't be man. Or do you mean the 9999s or whatever you get when there is an invalid HAZ? Definitely drop those (they aren't Z-scores). Or are there some other flags you are curious about?

Unasked:

3) There are 2 different HAZs coded in the newer DHS rounds - the old CDC standards and the new WHO ones...the new WHO ones are generally better, but maybe not comparable to other studies done before 2007 or so.

4) Which determinants are you trying to estimate? This is totally a selfish question because I'm working on a methodology paper about estimating these, and you can get into some trouble estimating the determinants of HAZ if you use time-varying regressors...

Subject: Re: Which data file should I use for nutrition outcomes of children under 5 years of age
Posted by Liz-DHS on Thu, 02 May 2013 21:51:32 GMT
View Forum Message <> Reply to Message

Dear User,
For guidance on calculating various indicators, please take a look at the Guide to DHS Statistics http://www.measuredhs.com/publications/publication-dhsg1-dhs -questionnaires-and-manuals.cfm and the standard recode manual http://www.measuredhs.com/publications/publication-dhsg4-dhs -questionnaires-and-manuals.cfm

---

Subject: Re: Which data file should I use for nutrition outcomes of children under 5 years of age
Posted by m.hasan1 on Mon, 13 May 2013 06:50:45 GMT
View Forum Message <> Reply to Message

Dear Member,

Thank you very much for your advice. The information avialable in the link is very much useful.

I am still doing literature review to understand which variables should be in the model. But I think I may include maternal factors, household demographic factors and child related factors.

I am not sure whether it is possible to estimate the effects of a time varying regressor when one is dealing with a single DHS data set given that DHS is cross-sectional.....

Subject: Re: Which data file should I use for nutrition outcomes of children under 5 years of age
Posted by Reduced-For(u)m on Mon, 13 May 2013 21:26:56 GMT
View Forum Message <> Reply to Message

Hi,

Glad I could help a little. Estimating cross-sectional determinants of child HAZ or time-invariant ones is certainly easier than trying to estimate cohort-based determinants (what I meant by time-variant), but it still requires, in my opinion, a bit more care than some people give it.

In particular, I worry most that people don't sufficiently worry about the distribution of child age-at-measurement across their explanatory variables of interest. I think we say "this is age adjusted height, so age shouldn't be a big predictor", but if you collapse HAZ by age-in-months, and graph it out, you'll realize how important age-at-measurement actually is in DHS countries...because HAZ is a cumulative measure of health/nutrition up until age-at-measurement, older kids have had a lot more time to "lose" HAZ relative to well-nourished children in the

reference group.

Just a couple of things to keep in mind: 1) if estimating time-invariant factors (say, rural born or maternal age at birth), make sure that the distributions of child age are similar across X (so, if X is "rural born", overlay a histogram or kernel-density plot of ages for rural and urban born children, and see if they match). 2) if you are using "time semi-variant" things like, say, Asset Quintile, you might have a more pronounced problem in that older parents tend to have both more assets and older children (this could bias your estimates of asset effect downward). 3) if you are using "cohort" variables, such as "drought exposure in-utero", you have to be super-duper careful, because some drought year where lots of kids are exposed will be correlated with some age-at-measurement, and thus induce a spurious HAZ-drought association that is driven by a drought/age-at-measurement association.

The gist is that most people include a linear control for age-in-months, and then write "age is a strong predictor of HAZ", which is true, but almost misses the point. Age is THE best predictor of HAZ in a lot of countries, but it is decidedly non-linear, and the model misspecification error (because it is specified erroneously as linear) is often times correlated with age in such a way that any covariates just accidentally associated with child age will pick up the misspecification error and attribute it to the covariate.

I find that in things like estimating effects of maternal age this affects coefficient estimates just a little bit. In things like in-utero/birth-year economic/health environment (cohort stuff), this affects estimates a whole lot. In between...I don't know, depends on the situation.

So... If you feel like it, once you get your list of determinants down, estimate it a few ways, by specifying age as linear, quadratic, a spline with nodes at each age-in-years, and dummy variables for each age in months, and then post the coefficient estimates on a few of the key determinants of interest for each specification. We can see what kind of difference it makes to your estimates.

Sorry...I'm almost done with a paper on this, and so I talk a lot about it.

Best,
j

Subject: Re: Which data file should I use for nutrition outcomes of children under 5 years of age
Posted by ebrofsky on Tue, 02 May 2017 19:57:07 GMT
View Forum Message <> Reply to Message

@Reduced-For(u)m, did you ever publish your paper on this topic? Sounds very interesting and relevant to the analysis I am trying to do.

Thanks,
Emma

Subject: Re: Which data file should I use for nutrition outcomes of children under 5 years of age
Posted by Reduced-For(u)m on Tue, 02 May 2017 20:47:33 GMT
View Forum Message <> Reply to Message

Funny you should ask... no. Every editor seems to think it is right, but that some other field/journal should publish it (some of them have even had it cited in their journals!). That said, I'll email it to you, along with a follow-up paper that extends the thinking to multiple-countries and/or heterogeneous impacts across age.

This is a general offer, so feel free (whoever) to PM me and I'll send it your way.