Subject: HIV and Household Data
Posted by joshgooch28 on Mon, 16 Nov 2020 11:37:45 GMT
View Forum Message <> Reply to Message

So for my undergrad economics dissertation I am hoping to look at the impact of HIV on children's education or children's health in South Africa.

Really struggling to wrap my head around how to accomplish this for the household recode. Firstly, trying to merge the HIV dataset with the household recode is obviously difficult because the barcode for HIV tests is listed as multiple variables in the hh recode, but as a single variable in the HIV data. How would I go about merging these?

Furthermore, how would I run a regression on this hh recode to see the impact of a hh with a HIV positive member on the education level/health of children in the household, given that each household (as a single observation) has multiple subset variables for each (as each member of the household)

Thanks in advance!

Subject: Re: HIV and Household Data
Posted by Bridgette-DHS on Wed, 18 Nov 2020 20:39:37 GMT
View Forum Message <> Reply to Message

Following is a response from DHS Research & Data Analysis Director, Tom Pullum:

This is a very ambitious project! The main problem you have is that the data probably do not include enough cases for you to have any chance of statistical significance.

Merging the AR file with other files is done with the household ID code, not the barcode. The merge can be tricky. The household ID code is a string variable that includes hv001 and hv002, which are numeric. Let us know if you still have difficulty with that step. hhid is contained within caseid in the IR and KR files.

You want to compare the child outcomes for children in a household with an HIV positive person with child outcomes for children without an HIV positive person. The health variables for children are in the KR file. School attendance is in the PR file, which has one record for each person in the household. I suppose I would break this into two analyses. One would be for children age 0-4, using the KR file with the HIV status of adults in the household (having the same values of v001 and v002) merged onto the children's records. The other would be for school-age children, using the PR file with the HIV status of adults in the household (having the same values of hv001 and hv002) merged onto the records of the school age children.

Your analysis should control for other characteristics of the household that may be associated with both HIV prevalence and the child outcomes.

Note that you do not know how long the adult has been HIV positive or how long they have been

in the same household, although there are ways to check whether they are the mother or father of the child (using the line numbers given by hv112 and hv114). Good luck.

Subject: Re: HIV and Household Data
Posted by joshgooch28 on Tue, 02 Feb 2021 10:56:16 GMT
View Forum Message <> Reply to Message

Hi

Thank you so much for the response, I am just getting around to the STATA work now and this is proving useful.

My main question now is, given that the AR recode often provides a HIV test/result for more than one member of the same household, meaning their hhid is the same, when trying to merge I am met with:

"variables hv001 hv002 do not uniquely identify observations in the using data" or the same but with the master data.

How would I resolve this?

Thanks

Subject: Re: HIV and Household Data
Posted by Bridgette-DHS on Fri, 05 Feb 2021 14:21:20 GMT
View Forum Message <> Reply to Message

Following is a response from DHS Research & Data Analysis Director, Tom Pullum:

It will be better if you merge the AR file with the PR file, rather than with the HR file. In the PR file, the individual household member's line number is hvidx. That needs to match with hivline in the AR file.