
Subject: Discrepancy between household and household members datasets
Posted by [mic321](#) on Thu, 05 May 2016 12:06:45 GMT

[View Forum Message](#) <> [Reply to Message](#)

Hello,

I think I have found inconsistency in datasets. Specifically the data on maternal education(variable hc62) are inconsistent within household surveys and household member surveys. Let me explain it with example of standard DHS VI from Benin. For instance third member of household 22, cluster 1 in household survey has reported value of maternal education (hc62_3) of 6 years. In household members dataset the value value is the dot. Line number for this observation is 148. According to other variables it is the same person.

This is common within whole dataset. Another example might be third member of household 24 (cluster 49) (in household members survey - line 6964) I might be missing something,however same pattern seems to appear for another countries as well. I would be very thankful if anyone could explain to me.

Thank you in advance for reading and possible even answering.

Subject: Re: Discrepancy between household and household members datasets
Posted by [Bridgette-DHS](#) on Thu, 12 May 2016 12:57:30 GMT

[View Forum Message](#) <> [Reply to Message](#)

Following is a response from Senior DHS Stata Specialist, Tom Pullum:

I think there are two problems here. The first one is that you may think the line number in the household file is given by hv003. Actually the line number is hvidx. hv003 is the line number of the household respondent. Second, education in years is given in two places. One is for everyone (above a cutoff age) in the household file and it is given as hv107. However, it is also asked in the survey of women. That response is carried over to the PR file as ha67. That variable, not hv107, is used for coding the education level of the mother. Open BJ61IRFL.dta and enter this line: "list hv001 hv002 hv003 hvidx hv101 hv104 hv105 hc62 ha67 hv107 if hv001==1 & hv002==22, table clean". You will get the following, and then you will see that hc62 is correct.

File Attachments

1) [discrepancy.jpg](#), downloaded 679 times
