Subject: underfive children HW data: why missing a lot in IR files
Posted by bun_2019fall on Fri, 11 Mar 2022 03:05:48 GMT
View Forum Message <> Reply to Message

Hi DHS program,

I recently try to use child nutrition data in IR files: specifically 3 variables, HW70_, HW71_, HW72_

I have been working with data collected from more than 30 countries, and I found that the HW_ data have a lot of missing. That is, I found that for a single survey, not all underfive children have HW_ information. Rather, there is a significant proportion of missing (for many surveys).

Here is a Stata example using NGIR7AFL.DTA:

```
use NGIR7AFL.DTA, clear
*children born during the last five years, the calendar starts 60 months ago
gen calstart=v008-60
*use the birth_01 as an example
gen calbirth1=1 if b3_01 >= calstart & b3_01!=.
tab calbirth1 /*21,855 birth_01*/
tab b8_01 if calbirth1==1 /*20,488 underfive children*/
sum hw70_1 if calbirth1==1 /*only 7,717 observations*/
```

I checked the specific documentation for this survey (and some other surveys) and I could not figure our why there is missing. I assume that HW_ is measured for all underfive children (living births by the time of the survey). But for most of the surveys that I am working at, HW_ is obviously not complete for all children born in the last five years. I wonder if I miss anything here? I also tried to search for any filter variable to see if only certain births are selected for anthropometric measures. But I could not find it. But children's height and weight information (separate variables) are pretty much aligning with the sample size for HW_ variables.

Any directions are much appreciated!
Thank you in advance!

Subject: Re: underfive children HW data: why missing a lot in IR files
Posted by Bridgette-DHS on Fri, 11 Mar 2022 12:56:45 GMT
View Forum Message <> Reply to Message

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

The height and weight of children under five are measured as part of the household survey, for all

children in the household. The measurements and the Z scores appear in the PR file. For children who are alive and living with the mother they also appear in the KR and IR files. The KR file is easier to use than the IR file for characteristics of children.

I opened the PR file for that Nigeria survey and entered the following line in Stata: "tab hc3 if hc1a<.,m". This gives the distribution of height for children under 5 and includes the cases with a dot or ".", which means "Not Applicable". There are 8 such children, out of 12, 867 (unweighted). Another 317 children have codes for not present, refused, other. The 8 children with a dot should have been given one of those three codes, but in analysis they would be lumped with the 317.

The IR or KR files include children who have died (b5=0)and children who are not living with the mother (b9 not equal to 0), as well as some of the 317+8 children described above. Your approach was more complicated than necessary. It's best to work off the PR or KR files for anthropometry.

Subject: Re: underfive children HW data: why missing a lot in IR files
Posted by bun_2019fall on Fri, 11 Mar 2022 17:40:35 GMT
View Forum Message <> Reply to Message

Hi thank you so much for the detailed response. I replicate your way of tabulation in the PR file and what I observed is consistent with your tabulation/explanation. I did a comparison of PR and IR file, and I am still confuse about the gap between IR and PR files, regarding under-five children and their anthropometric data. Below is the Stata scripts that I used to make the comparison (taking Nigeria as an example).

Briefly, I found that in IR files, there are (way) more under-five children as reported by each interviewed women (I use b8_ to tabulate, which indicates children's current age). In the example of Nigeria, there are more than 30K under-five children. Amongst all these children, a subsample has HW measures - which is correspond with the sample size of children having HC measures in PR files. According to this comparison, it seems to me that:
1.not all under-five children reported in IR files have anthropometric measures
2. all under-five children in PR files have anthropometric measures

My question is that why some under-five children sample in #1 (IR files) were not measured for height and weight? This discrepancy not only exist for this Nigeria file; but I observe similar patterns for other country data as well (at least for the 30+ countries that I am working with). I wonder what documentation I should refer to, in order to resolve this puzzle? Thank you again!!!

clear all
set maxvar 20000

```
***Nigeria PR file
use "NGPR7AFL.DTA", clear
*total # of underfive children
gen underfive=1 if hc1 < 60
tab underfive
*12,867 underfive
sum hc7*
*12.5K children have hw info
gen momid=hc60
tab momid
*moms not in IR file
recode momid 993/995=.
drop if momid==.
sum hc7*
*recode the invalid hc values
recode hc70 9996/9999=.
recode hc71 9996/9999=.
recode hc72 9996/9999=.
*all underfive children having valid anthropometric data: 11.4k
sum hc7*

*** Nigeria IR file
use "NGIR7AFL.DTA", clear
forvalues i=1/6{
 gen underfive`i'=1 if b8_0`i' < 5
}
sum underfive*
keep caseid underfive* hw7*
reshape long underfive hw7_ hw70_ hw71_ hw72_ hw73_ , i(caseid) j(birth)
drop if underfive==.
*30,713 underfive: ALL children
sum hw7*
*recode the invalid hw values
recode hw70_ 9996/9999=.
recode hw71_ 9996/9999=.
recode hw72_ 9996/9999=.
recode hw73_ 9996/9999=.
*underfive children having valid anthropometric data: 11.4k
sum hw7*
```

Subject: Re: underfive children HW data: why missing a lot in IR files
Posted by Bridgette-DHS on Fri, 11 Mar 2022 18:08:14 GMT
View Forum Message <> Reply to Message

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

I don't have time to go through your Stata code, but it is simply not efficient to analyze anthropometry for children using the birth histories or the calendar in the IR file. If you want to link children and mothers, do it with the KR file. Children in the KR file have most of the mother's data, including her own anthropometry, on the same record. And if you want to use the IR file to identify births, it's much easier to use the birth histories than to use the calendar. Linking children in the PR file to births in the calendar just doesn't make sense to me because of the availability of the KR file, which already contains all the anthropometry data, except for children who were not in the household on the day of measurement, either because they had died or because they lived elsewhere.

Subject: Re: underfive children HW data: why missing a lot in IR files
Posted by bun_2019fall on Fri, 11 Mar 2022 18:14:57 GMT
View Forum Message <> Reply to Message

Thank you for your suggestion! I will try KR file to link children and mothers then). I guess the discrepancy might be related to what you said "children who were not in the household on the day of measurement, either because they had died or because they lived elsewhere." Thanks again!