

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

I assume you have already looked at chapter 17, and the relevant part of the household questionnaire, in the main report. In this survey, of all the children age 2-14 in the household, one was randomly selected. The child discipline questions were asked with reference to that one child. However, in the PR household file, which has one record for everyone in the household, the responses for that child were attached to every person in the household, regardless of whether they are the reference child or are even age 2-14. The variables are sh148-sh161.

For most purposes, the units of analysis are the children. If you want to get an estimate of, say, what proportion of children age 2-14 are "Yes" on sh150, I would use the PR file, restrict to children age 2-14, and run "tab sh150 [iweight=hv005/1000000]". (You can also use svyset and svy.) The response for the reference child would be imputed to all the other children in the household. But if you wanted this proportion by characteristics of the children, such as their age, you would need to restrict to the reference children, with "if hvidx==sh148". I would recommend that when you select with hvidx=sh148, you also weight by k\*hv005, where k is the number of children age 2-14 in the household. The sampling probability within the household is 1/k, and you should compensate for that.

For the selected child (the one with hvidx=sh148) you have up to 3 possible relevant adults in the household: the mother, the father, and the caregiver. The line numbers for these adults are hv112, hv114, and sh149, respectively. Usually the caregiver will be the mother (that is, sh149==sh112), but not always. The mother and father are not always alive and in the household. I see that some children are missing a code for the caregiver.

If you look on the forum you can find ways to attach characteristics of the mother (in the IR file) to the child in the PR file, IF the child and the mother are in the same household. The same approach can be used to link the characteristics of the father (in the MR file) to the child. You may be able to link the occupation (for example) of the caregiver to the child, if the caregiver is in the IR or PR file. The caregiver may be a grandparent and outside the age range of the IR or MR file. If the caregiver is the mother or father, and in the same household as the child, you will be ok.