

Following is a response from Senior DHS staff member, Tom Pullum:

I looked at the data using the HR file (MLHR76A) because households were sampled. As you say, there is no label for hv022 (which has 15 values), but there IS a label for hv023 (which has 12 values). I entered "tab hv022 hv023". The table shows that there is mostly a one-to-one correspondence between these two variables, but there are 3 values of hv022 that identify smaller sampling strata that were repeats of hv022: hv022=2 gives 390 extra households with hv023=8; hv022=8 gives 308 extra households with hv023=4; and hv022=11 gives 73 extra cases with hv023=3. I would say that these 681 households, less than 7% of the sample, are the only problematic ones, but even for them, you have a label from hv023.

If you can't find further documentation on the location of these 3 groups of households, within hv023, you could go to the geographic data and make a map showing where the clusters are located. The displacement of the clusters would not impair your ability to get a general location. However, I personally would be satisfied to use hv023 rather than hv022.

A more serious issue for this Mali survey is the omission of regions 6, 7, and 8 from the sample. (compare "label list HV024" with "tab hv024"). They are sparsely populated but had to be omitted for security reasons. Most of the surveys of Mali, I believe, if not all, have had to omit the northern part of the country.

Hope this helps.