

---

Subject: Re: merging wealth index from Ethiopia DHS 2000 to women's file  
Posted by [acp49](#) on Tue, 17 Oct 2017 19:58:22 GMT

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

---

I have recently used the "Merge\_WI\_onto\_IR\_do.txt" attachment posted on the "Re: merging wealth index from Ethiopia DHS 2000 to women's file" chain on 22 March 2013 to merge children's and wealth index data sets for certain countries. However, for the Burkina Faso 1998 data set, the cluster and household ID numbers in the children's data set do not line up with the numbers being reported in the "whhid" variable in the wealth index data set.

For the children's data set, the range for cluster ID (v001) is [1,210] and the range for household ID (v002) is [1,40]. For reference, the first part of the code listed in the provided attachment:

- \* whhid has 12 characters but we need to figure out how to use them
- \* which part is cluster id and which part is household id?

```
gen str1 c1=substr(whhid,1,1)
gen str1 c2=substr(whhid,2,1)
gen str1 c3=substr(whhid,3,1)
gen str1 c4=substr(whhid,4,1)
gen str1 c5=substr(whhid,5,1)
gen str1 c6=substr(whhid,6,1)
gen str1 c7=substr(whhid,7,1)
gen str1 c8=substr(whhid,8,1)
gen str1 c9=substr(whhid,9,1)
gen str1 c10=substr(whhid,10,1)
gen str1 c11=substr(whhid,11,1)
gen str1 c12=substr(whhid,12,1)
```

However, the returned values from this code do not match up with these cluster and household ID ranges. I have tried parsing in several different ways. Strings c1,c2,c3,c4 are empty. Then the ranges for the remaining strings are: c5 [1,2]; c6 [0,9]; c7 [0,9]; c8 [1,7]; c9 [0,9]; c10 [0,9]; c11 [1,2]; c12 [0,9]. Therefore, c5,c6,c7 is clearly the cluster ID variable, but the remaining numbers do not match the household ID or other types of ID variables.

Would it be possible to receive further assistance with determining the relevant ID variable(s) to properly merge the children's and wealth index data sets for Burkina Faso 1998?

---