Subject: KDHS 2014: Table 2.14 School attendance ratios
Posted by sokiya on Tue, 07 Nov 2023 17:11:46 GMT
View Forum Message <> Reply to Message

I am trying to generate Table 2.14 School attendance ratios using the microdata and the code from the DHS GitHub repo as shown below

```
* open the birth history data to extract date of birth variables needed.
use "KEBR71FL.DTA", clear

* keep only the variables we need
keep v001 v002 v003 b3 b16
* drop if the child in the birth history was not in the household or not alive
drop if b16==0 | b16==.
* rename key variables for matching
rename b16  hvidx
rename v001 hv001
rename v002 hv002
* sort on key variables
sort hv001 hv002 hvidx

* if there are some duplicates of line number in household questionnaire, we need to drop the duplicates
gen dup = (hv001 == hv001[_n-1] & hv002 == hv002[_n-1] & hvidx == hvidx[_n-1])
drop if dup==1
drop dup
* re-sort to make sure still sorted
sort hv001 hv002 hvidx

* save a temporary file for merging
tempfile tempBR
save `tempBR'

* use the PR file for household members for the NAR and GAR indicators
use "KEPR71FL.DTA", clear


* merge in the date of birth from the women's birth history for the household member
merge 1:1 hv001 hv002 hvidx using `tempBR'
* there are a few mismatches of line numbers (typically a small number of cases) coming rom the BR file, so let's drop those
drop if _merge==2

* restrict to de facto household members age 5-24, and drop all others
keep if hv103==1 & inrange(hv105,5,24)

* now we calculate the child's age at the start of the school year
* but first we have to specify the month and year of the start of the school year referred to in the
```

```
survey
* example, for Zimbabwe 2015 survey this was January 2015
global school_start_yr = 2014
global school_start_mo = 1
* also need the age ranges for primary and secondary
global age_prim_min = 6
global age_prim_max = 13
global age_sec_min = 14
global age_sec_max = 17

* produce century month code of start of school year for each state and phase
gen cmcSch = ($school_start_yr - 1900)*12 + $school_start_mo
replace cmcSch = cmcSch+12 if hv008 >= cmcSch+12
* calculate the age at the start of the school year, using the date of birth from the birth history if we have it
gen school_age = int((cmcSch - b3) / 12) if b3 != .
* Impute an age at the beginning of the school year when CMC of birth is unknown
* the random imputation below means that we won't get a perfect match with the report, but it will be close
gen xtemp = hv008 - (hv105 * 12) if b3 == .
gen cmctemp = xtemp - int(uniform()*12) if b3 == .
replace school_age = int((cmcSch - cmctemp) / 12) if b3 == .

* Generate variables for whether the child is in the age group for primary or seconary school
gen prim_age = inrange(school_age,$age_prim_min,$age_prim_max)
gen sec_age  = inrange(school_age,$age_sec_min ,$age_sec_max )

* create the school attendance variables, not restricted by age
gen prim = (hv122 == 1)
gen sec  = (hv122 == 2)

* set sample weight
cap gen wt = hv005/1000000

* For NAR we can use this as just regular variables and can tabulate as follows, but can't do this for GAR as the numerator is not a subset of the denominator
* NAR is just the proportion attending primary/secondary school of children in the correct age range, for de facto children
gen nar_prim = prim if prim_age == 1
gen nar_sec  = sec  if sec_age  == 1
lab var nar_prim "Primary school net attendance ratio (NAR)"
lab var nar_sec "Secondary school net attendance ratio (NAR)"

* tabulate primary school attendance
tab hv104 nar_prim [iw=wt] , row
tab hv025 nar_prim [iw=wt] , row
tab hv270 nar_prim [iw=wt] , row
* tabulate secondary school attendance
```

tab hv104 nar_sec [iw=wt] , row
tab hv025 nar_sec [iw=wt] , row
tab hv270 nar_sec [iw=wt] , row

Any help will be appreciated

Subject: Re: KDHS 2014: Table 2.14 School attendance ratios
Posted by Bridgette-DHS on Wed, 08 Nov 2023 13:08:03 GMT
View Forum Message <> Reply to Message

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

School attendance ratios are complicated and are not central to the purpose of DHS surveys. I will try to get to this but there will be a delay.

Subject: Re: KDHS 2014: Table 2.14 School attendance ratios
Posted by Bridgette-DHS on Fri, 17 Nov 2023 16:27:36 GMT
View Forum Message <> Reply to Message

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

I have added a Stata program that is an improvement on the one you were using and on the GitHub program, although it just calculates the net attendance ratios in table 2.14 and just by the child's sex and place of residence. The other parts of table 2.14 should be easy to add. I had written the Stata program now on GitHub several years ago, basically as a translation of a CSPro program originally written by Trevor Croft long ago. The attached version has some simplifications and more comments.

In the original CSPro program for this table for the Kenya 2014 survey, the year and month for eligibility are 2014 and 2 (February), respectively. You were using month 1 (January). It is quite possible that month 1 would have been more consistent with the Kenyan school system, but the program used month 2. In any application of this program, it is essential to adjust those numbers as well as the age ranges for primary and secondary school.

A peculiarity of the DHS procedure (which probably originates with a UN agency) is that a child who is primary age but attending school is counted as not in school. Similarly, a child who is secondary age but attending post-secondary is counted as not in school. I don't think this is fair to kids who manage to skip grades, but that's the procedure.

As stated in the title, the procedure is limited to children who are de facto residents of the household, i.e. were in the household the previous night. This makes some difference.

Unfortunately, the procedure includes a random component that makes it impossible to match the table. The child's cmc of birth is given by b3 if the child is in the BR file, as well as the PR file.

For children who are in the PR file but not in the BR tile (and there are many of them), the month of birth is imputed with a uniform random distribution that is consistent with the stated age in years (hv105). That means different results will be given by different random number generators and different seeds. My program just uses "uniform()", and the table in the report uses the CSPro generator. There is no way around this issue. It would be possible to produce a version in which all children were assigned to month 6, or to month 7, which would avoid the random component but would not be as accurate. The bottom line is that you can never match the table on school attendance ratios exactly.

File Attachments

1) KE2014_in_school_do_16Nov2023.txt, downloaded 185 times

Subject: Re: KDHS 2014: Table 2.14 School attendance ratios
Posted by sokiya on Fri, 17 Nov 2023 17:46:43 GMT
View Forum Message <> Reply to Message

Thanks so much for writing the do file and the comprehensive feedback. I sincerely appreciate the assistance.

Subject: Re: KDHS 2014: Table 2.14 School attendance ratios
Posted by geoK on Thu, 25 Apr 2024 14:21:08 GMT
View Forum Message <> Reply to Message

Dear DHS Staff,

I am trying to calculate the school attendance ratios for 2022 KDHS. I am using the code published on the GitHub, but I am unable to match the figures on the final report / statcompiler. I get 81.9% for primary NAR (instead of 86.2%) and 42.4% (instead of 48.8%).

I have also came across a few posts regarding possible need to recode a few of the school-related variables for this specific survey. Am I right I should wait for a new release of these variables to be able to correctly estimate the primary and secondary school attendance ratios, or should I try and identify what is causing the mismatch? Thank you!

Subject: Re: KDHS 2014: Table 2.14 School attendance ratios
Posted by Bridgette-DHS on Fri, 26 Apr 2024 11:22:09 GMT
View Forum Message <> Reply to Message

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

Hope you can wait a little longer (a month?) and then check back, if nothing has changed. We have been waiting for approval from Kenya before issuing those changes. It is entirely possible that you are correct.

Subject: Re: KDHS 2014: Table 2.14 School attendance ratios
Posted by geoK on Mon, 03 Jun 2024 15:43:41 GMT
View Forum Message <> Reply to Message

Dear staff,
I was wondering if there was any news about the education-related indicators / data for the 2022 KDHS survey round, and whether changes have been or will be made in the near future.
Many thanks.
Best regards!

Subject: Re: KDHS 2014: Table 2.14 School attendance ratios
Posted by Bridgette-DHS on Tue, 04 Jun 2024 18:25:38 GMT
View Forum Message <> Reply to Message

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

We expect that an updated version of the Kenya 2022 data files will be available on the website within a week or two.

Subject: Re: KDHS 2014: Table 2.14 School attendance ratios
Posted by geoK on Mon, 17 Jun 2024 12:13:45 GMT
View Forum Message <> Reply to Message

Dear Staff, will registered users be notified when changes to the 2022 dataset will be completed and made available online for download?
Thank you!

Subject: Re: KDHS 2014: Table 2.14 School attendance ratios
Posted by Bridgette-DHS on Mon, 17 Jun 2024 12:16:43 GMT
View Forum Message <> Reply to Message

Yes, all users who downloaded the previous version will be notified when we release the new version. The notification/alert will include a description of changes made in the new version.

Subject: Re: KDHS 2014: Table 2.14 School attendance ratios
Posted by geoK on Wed, 31 Jul 2024 15:10:17 GMT
View Forum Message <> Reply to Message

Dear Staff, I was wondering if there is an expected date for the updated primary and secondary school indicators / data for the latest Kenya round to be released. Many thanks!

Subject: Re: KDHS 2014: Table 2.14 School attendance ratios
Posted by Bridgette-DHS on Wed, 31 Jul 2024 17:49:23 GMT
View Forum Message <> Reply to Message

We do not have a release date at this time.

Subject: Re: KDHS 2014: Table 2.14 School attendance ratios
Posted by geoK on Fri, 16 Aug 2024 21:06:59 GMT
View Forum Message <> Reply to Message

Dear Staff,
I would like to pick up this discussion.
I used the revised dataset that was released on August 13th. I'm however still unable to generate Table 2.14.

Can you please help.

Many thanks.

Subject: Re: KDHS 2014: Table 2.14 School attendance ratios
Posted by Bridgette-DHS on Mon, 19 Aug 2024 11:41:23 GMT
View Forum Message <> Reply to Message

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

Your question seems to be about table 2.14 in the Kenya 2014 survey. The updated files were for the Kenya 2022 survey. In the final report on that survey, table 2.14 has nothing to do with schooling. Can you clarity your question? Is it REALLY about table 2.14 in the 2014 survey?

Subject: Re: KDHS 2014: Table 2.14 School attendance ratios
Posted by geoK on Wed, 21 Aug 2024 21:32:27 GMT
View Forum Message <> Reply to Message

Dear DHS staff,
My apologies. I was referring to Table 2.12 KDHS 2022. I have posted a question specific to that dataset.

Thank you so much.

Subject: Re: KDHS 2014: Table 2.14 School attendance ratios
Posted by sokiya on Sun, 29 Sep 2024 18:01:10 GMT
View Forum Message <> Reply to Message

Dear DHS Team,
I am interested in Gross Attendance Ratios in the same table since I want to have the same estimates at county level to explore trends over time. I am using the attached do file which is a modification of what you shared but the results seems to be a bit off. Is there something I am getting wrong? Maybe the CSPro syntax can be of help in this case. There's a section of the report that states Quote:Youth are considered to be attending school currently if they attended a formal academic school at any point during the given school year. Note that the NAR and GAR values reported here are not comparable with those from previous DHS surveys due to an improvement in the precision of calculation.
Thanks in advance!

File Attachments
1) KE2014_in_school_do_16Nov2023-so.do, downloaded 102 times

Subject: Re: KDHS 2014: Table 2.14 School attendance ratios
Posted by Bridgette-DHS on Mon, 30 Sep 2024 15:09:53 GMT
View Forum Message <> Reply to Message

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

Here is a link to the relevant CSPro file for chapter 2 in the final report on the Kenya 2014 survey (attached). Often there is a difference between the table numbers in the CSPro code in the report, and that happens here. The table that appears as 2.14 in the final report is table 2.13 in the attachment. The relevant lines appear at different places in the file. I recommend that you just search for all the appearances of "2.13".

Let us know if you still have a problem.

File Attachments
1) CHAPT2.TXT, downloaded 88 times

Subject: Re: KDHS 2014: Table 2.14 School attendance ratios
Posted by sokiya on Mon, 30 Sep 2024 17:15:04 GMT
View Forum Message <> Reply to Message

Thanks so much for sharing the CSPro syntax. I sincerely appreciate.
I recall you attributed the difference in how CSPro and Stata generate the random number. The Stata do file that you shared for 2022 estimates were really close hence my believe the 2014 can also be close.