

---

Subject: Clarification on missing variables in PRfile  
Posted by [Lukresha](#) on Thu, 07 Jul 2016 08:48:47 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Hello,

I am using the 2014 KDHS for my dissertation on education. However, I noted that the variables hc61 (mother's highest educational level), hc62 (mother's highest year of education) and sh20a (why not registered) have values for only members of household age 4 and below and missing for the rest (age 5-95).

Is there a reason why the data on the above named variables are available for only household members age 4 and below and missing for the rest?

---

---

Subject: Re: Clarification on missing variables in PRfile  
Posted by [Bridgette-DHS](#) on Thu, 07 Jul 2016 13:00:05 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Following is a response from Senior DHS Stata Specialist, Tom Pullum:

The hc variables are collected in the household survey and are restricted to children under five. You need a different strategy.

For your purpose, you could go to the BR file and keep the children who are alive (b5=1). In that file, b8 is the age of the child. The age of the mother is v012 and her education is given by v106, v107, v133, v149, and some other variables. You could work just with that file, but it does not include any information about the child's schooling. To get that information, you could reduce that file to children who are also in the household file with "keep if b16 > 0 & b16 < .", followed by a merge with the PR file (matching hv001 hv002 hvidx in the PR file with v001 v002 b16 in the BR file), and then connect the mother's educational level with whether the child is currently attending school, for example. If you do that kind of a merge, you will lose the children who are not living in the same household as the mother, perhaps because the mother has died, and those lost cases may be very different from the matched cases. If the mother has died or is not in the household, there is no way to get the mother's educational level. However, this strategy would get at your research question.

The "children" who are living in the same household as the mother may be grown children (adults), in their 20s or even in their 30s. Those adult children, for age ≥ 15, should appear in their own right in the IR or MR file, as well as in the BR file. You could link back to them in the IR or MR file and get a much longer-term perspective on the mother/child correspondence in education.

---

---

Subject: Re: Clarification on missing variables in PRfile

Posted by [Lukresha](#) on Thu, 14 Jul 2016 07:53:33 GMT

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

Does the same reasoning apply to variable sh20a(why not registered in school)?

As I would assume that this question is viable to household members who are more than 5 years as they are the ones attending school.

This is because this question is asked after question on whether a household member attended school or not in the 2014 school year in the PRfile.

---

Subject: Re: Clarification on missing variables in PRfile

Posted by [Bridgette-DHS](#) on Thu, 14 Jul 2016 16:39:03 GMT

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

Following is a response from Senior DHS Stata Specialist, Tom Pullum:

Yes, this is a survey-specific question that only applies to school-age children who are not in school.

---

Subject: Re: Clarification on missing variables in PRfile

Posted by [Lukresha](#) on Fri, 15 Jul 2016 11:58:23 GMT

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

Then if it applies to school age children, shouldn't it be asked to those more than 3 years as they're the ones likely to be of school going age? If I use the command tab sh20a hv105, the results shown are only for those under 4 years as shown below:

tab sh20a hv105

	age of household members					
why not registered	0	1	2	3	4	Total
-----+-----+-----+-----+-----						
too far	130	112	135	118	123	618
no money	156	144	159	196	149	804
not aware	255	256	260	270	242	1,283
not necessary	151	163	174	185	156	829
nomadic life, difficu	110	97	98	113	100	518
other	526	487	488	515	472	2,488
missing	17	11	10	8	14	60
-----+-----+-----+-----+-----						
Total	1,345	1,270	1,324	1,405	1,256	6,600

Yet, my age of household members ranges from 0 to 95. I think this shows that there are no reasons for why not registered for those older than 4 years.

tab hv105

age of   household   members	Freq.	Percent	Cum.
0	4,244	2.76	2.76
1	4,222	2.74	5.50
2	4,398	2.86	8.36
3	4,701	3.06	11.42
4	4,405	2.86	14.28
5	4,283	2.78	17.07
6	5,338	3.47	20.54
7	5,008	3.26	23.79
8	5,184	3.37	27.16
9	4,564	2.97	30.13
10	5,141	3.34	33.47
11	4,028	2.62	36.09
12	4,913	3.19	39.29
13	4,199	2.73	42.01
14	4,412	2.87	44.88
15	3,504	2.28	47.16
16	3,418	2.22	49.38
17	3,137	2.04	51.42
18	3,369	2.19	53.61
19	2,504	1.63	55.24
20	2,892	1.88	57.12
21	2,082	1.35	58.47
22	2,468	1.60	60.08
23	1,888	1.23	61.31
24	2,275	1.48	62.78
25	2,545	1.65	64.44
26	2,305	1.50	65.94
27	2,221	1.44	67.38
28	2,553	1.66	69.04
29	1,868	1.21	70.26
30	2,829	1.84	72.10
31	1,501	0.98	73.07
32	1,994	1.30	74.37
33	1,199	0.78	75.15
34	1,646	1.07	76.22
35	2,157	1.40	77.62
36	1,659	1.08	78.70
37	1,337	0.87	79.57
38	1,552	1.01	80.58
39	1,217	0.79	81.37
40	1,975	1.28	82.65
41	1,107	0.72	83.37

42	1,360	0.88	84.25
43	827	0.54	84.79
44	933	0.61	85.40
45	1,418	0.92	86.32
46	952	0.62	86.94
47	741	0.48	87.42
48	838	0.54	87.97
49	738	0.48	88.45
50	1,126	0.73	89.18
51	941	0.61	89.79
52	1,016	0.66	90.45
53	740	0.48	90.93
54	911	0.59	91.52
55	896	0.58	92.11
56	935	0.61	92.71
57	604	0.39	93.11
58	660	0.43	93.54
59	495	0.32	93.86
60	1,133	0.74	94.59
61	399	0.26	94.85
62	664	0.43	95.28
63	357	0.23	95.52
64	485	0.32	95.83
65	757	0.49	96.32
66	448	0.29	96.62
67	327	0.21	96.83
68	383	0.25	97.08
69	329	0.21	97.29
70	650	0.42	97.71
71	228	0.15	97.86
72	315	0.20	98.07
73	166	0.11	98.17
74	256	0.17	98.34
75	365	0.24	98.58
76	185	0.12	98.70
77	127	0.08	98.78
78	225	0.15	98.93
79	138	0.09	99.02
80	335	0.22	99.23
81	81	0.05	99.29
82	136	0.09	99.38
83	83	0.05	99.43
84	119	0.08	99.51
85	131	0.09	99.59
86	91	0.06	99.65
87	45	0.03	99.68
88	49	0.03	99.71
89	52	0.03	99.75

90	115	0.07	99.82
91	17	0.01	99.83
92	36	0.02	99.86
93	16	0.01	99.87
94	49	0.03	99.90
95+	157	0.10	100.00

---

Total	153,822	100.00
-------	---------	--------

What could be the problem?

I am interested in the reasons for not attending school for members of household age 6-14 years and 15-18 years, but with the above sh20a, I cannot observe any reasons for this age groups due to the above cited reasons.

---

Subject: Re: Clarification on missing variables in PRfile  
 Posted by [Bridgette-DHS](#) on Fri, 15 Jul 2016 13:02:43 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Another response from Tom Pullum:

When you first referred to sh20a, you said that it referred to school registration. I should have checked that. In fact, it refers to birth registration. It is based on a follow-up to the question on whether the child has a birth certificate, which is only asked for children under 5. Whether the child has a birth certificate is given by hv140. If the response is "no", then sh20a is asked. I suggest you do "tab hv140 sh20a, m" in the PR file.

It would be interesting to know whether children who have a birth certificate are more likely to attend school. However, there is no overlap between the age range for hv140 (and sh20a) and the age range for school attendance.