
Subject: Inconsistencies between recode number of deaths and birth history

Posted by [cportner](#) on Thu, 15 Dec 2022 04:23:12 GMT

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

Hello,

It looks like there are 142 women in NFHS-4 for whom v206+v207 (sons and daughters who have died) does not match the count of deaths across b7_*. Is there a reason for this mismatch, and which one should I trust more?

All the best,

Claus Pörtner

Subject: Re: Inconsistencies between recode number of deaths and birth history

Posted by [Bridgette-DHS](#) on Thu, 15 Dec 2022 16:38:38 GMT

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

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

I used the following Stata code to check whether there are any discrepancies between the v20* variables and the b4* and b5* variables in terms of the numbers of sons and daughters who are alive or dead. I get perfect agreement. I used b5 rather than b7, but b5 and b7 should agree, in that b7 is <. If the child died (b5=0). However, b7=. could mean either that the child survived (b5=0) or there was no birth for the specified value of bidx (b5=.) Can you check again?

```
use v20* b4* b5* using "...IAIR52FL.dta", clear
rename *_0* *_*
```

```
gen v20sons_alive=v202+v204
gen v20dtrs_alive=v203+v205
```

```
gen sons_alive=0
gen dtrs_alive=0
gen sons_died=0
gen dtrs_died=0
```

```
forvalues li=1/16 {
```

```
scalar si=`li'
scalar list si
```

```
replace sons_alive=sons_alive+1 if b4_`li'==1 & b5_`li'==1
replace dtrs_alive=dtrs_alive+1 if b4_`li'==2 & b5_`li'==1
replace sons_died=sons_died+1 if b4_`li'==1 & b5_`li'==0
replace dtrs_died=dtrs_died+1 if b4_`li'==2 & b5_`li'==0
```

```
}
```

```
tab v20sons_alive sons_alive  
tab v20dtrs_alive dtrs_alive  
tab v206 sons_died  
tab v207 dtrs_died
```

```
gen d1=v20sons_alive-sons_alive  
gen d2=v20dtrs_alive-dtrs_alive  
gen d3=v206-sons_died  
gen d4=v207-dtrs_died
```

```
tab1 d1 d2 d3 d4
```

Subject: Re: Inconsistencies between recode number of deaths and birth history
Posted by [cportner](#) on Thu, 15 Dec 2022 18:22:03 GMT
[View Forum Message](#) <> [Reply to Message](#)

The data file that I used was IAIR74FL.DTA. The 52 one is NFHS-3.

If I run Tom's code on IAIR74FL.DTA I get:

```
. tab1 d1 d2 d3 d4
```

-> tabulation of d1

d1	Freq.	Percent	Cum.
-3	1	0.00	0.00
-2	4	0.00	0.00
-1	7	0.00	0.00
0	699,600	99.99	99.99
1	66	0.01	100.00
2	5	0.00	100.00
3	2	0.00	100.00
4	1	0.00	100.00
Total	699,686	100.00	

-> tabulation of d2

d2	Freq.	Percent	Cum.
-4	1	0.00	0.00
-2	1	0.00	0.00

-1	10	0.00	0.00
0	699,601	99.99	99.99
1	67	0.01	100.00
2	6	0.00	100.00

Total	699,686	100.00
-------	---------	--------

-> tabulation of d3

d3	Freq.	Percent	Cum.
-4	1	0.00	0.00
-3	2	0.00	0.00
-2	5	0.00	0.00
-1	66	0.01	0.01
0	699,599	99.99	100.00
1	8	0.00	100.00
2	4	0.00	100.00
3	1	0.00	100.00

Total	699,686	100.00
-------	---------	--------

-> tabulation of d4

d4	Freq.	Percent	Cum.
-2	6	0.00	0.00
-1	66	0.01	0.01
0	699,603	99.99	100.00
1	9	0.00	100.00
2	1	0.00	100.00
4	1	0.00	100.00

Total	699,686	100.00
-------	---------	--------

Subject: Re: Inconsistencies between recode number of deaths and birth history
 Posted by [Bridgette-DHS](#) on Fri, 16 Dec 2022 13:19:39 GMT

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

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

Apologies for checking the wrong survey! I have revised that Stata program to list the inconsistent cases and loop through the IR files from the five India surveys. I will paste it below. It's pretty general and could be used on any DHS surveys. I use a prefix "v" for subtotals coming from

v202-v207 and a prefix "b" for subtotals coming from the b4_* and b5_* variables. The number of children "living with the mother" could be checked against the number with b9_* equal to 0.

There are no inconsistent cases in the NFHS-2, -3, and -5 surveys. There is one inconsistent case in NFHS-1, and it is easily found, because it is the only case with v201=. . In NFHS-4, I find 128 cases with 2 inconsistencies and 21 cases with 4 inconsistencies. I suspect that the inconsistencies occur in pairs, for example with a difference of +1 for sons alive and -1 for sons dead. I did not take the time to check that.

Such inconsistencies should not occur. If you need to give priority to one over the other, the b variables should be more reliable. In the interview, there is probing during the collection of the birth histories. The totals and subtotals in v201-v207 should be forced to agree with the b variables. Thank you for identifying this issue with the NFHS-4.

* this program checks v202-v207 against the birth histories

```
program define check_subtotals
```

```
keep v024 v001 v002 v003 v20* b4_* b5_*
rename *_0* *_*
```

```
gen vsons_alive=v202+v204
gen vdtrs_alive=v203+v205
gen vsons_died=v206
gen vdtrs_died=v207
```

```
gen bsons_alive=0
gen bdtrs_alive=0
gen bsons_died=0
gen bdtrs_died=0
```

```
* Compare the maximum observed value of v201 with the maximum allowed value of the birth
history index.
summarize v201
```

```
scalar sbstop=0
forvalues li=15/25 {
  if sbstop==0 {
    capture confirm numeric variable b4_`li', exact
    if _rc>0 {
      scalar sbstop=1
      scalar sbmax=`li'-1
    }
  }
}
```

```
scalar list sbmax
local lbmax=sbmax
```

```
quietly forvalues li=1/'lbmax' {
  scalar si=`li'
  replace bsons_alive=bsons_alive+1 if b4_`li'==1 & b5_`li'==1
  replace bdtrs_alive=bdtrs_alive+1 if b4_`li'==2 & b5_`li'==1
  replace bsons_died=bsons_died+1 if b4_`li'==1 & b5_`li'==0
  replace bdtrs_died=bdtrs_died+1 if b4_`li'==2 & b5_`li'==0
}
```

```
*tab vsons_alive bsons_alive
*tab vdtrs_alive bdtrs_alive
*tab vsons_died bsons_died
*tab vdtrs_died bdtrs_died
```

```
gen d1=vsons_alive-bsons_alive
gen d2=vdtrs_alive-bdtrs_alive
gen d3=vsons_died -bsons_died
gen d4=vdtrs_died -bdtrs_died
```

```
tab1 d1 d2 d3 d4
```

```
gen D=0
label variable D "case has D inconsistencies"
replace D=D+1 if d1~=0
replace D=D+1 if d2~=0
replace D=D+1 if d3~=0
replace D=D+1 if d4~=0
```

```
sort v024 v001 v002 v003
```

```
end
```

```
*****
*****
*****
*****
*****
```

```
* Execution begins here
```

```
local lpvs "23 42 52 74 7D"
```

```
foreach ipv of local lpvs {
  scalar spv="`ipv'"
  use "...IAIR`ipv'FL.dta", clear
  quietly check_subtotals
}
```

```
scalar list spv
tab D if D>0
list v024 v001 v002 v003 v201 *alive *died if D>0, table clean abbrev(11)
}
```
