

Understanding Society User Support - Support #2158

Number of biological children ever had

10/02/2024 02:23 PM - Irene Frageri

Status:	Feedback	Start date:	10/02/2024
Priority:	Normal	% Done:	50%
Assignee:	Understanding Society User Support Team		
Category:	Data documentation		
Description			
<p>Hello, I am looking for a variable (or alternatively for some help building it), which expresses the number of biological children a respondent has ever had, at each wave. It does not matter if these children are in the household or not, and/or if they depend on the respondent.</p> <p>I tried to build this variable using n. of children in Wave 1 (built using a_inprnt and a_lprnt) and adding at each sequent wave the number of newborn (x_nnewborn). The issue is: when variable x_nnewborn is missing, I am just keeping the n. of babies they had in the wave before but there is risk of committing an error, if maybe the respondent had a child in the wave in which the variable nnewborn is missing. Is there a more accurate way to have this information about parenthood?</p> <p>Thank you for your help.</p>			

History

#1 - 10/03/2024 06:51 AM - Understanding Society User Support Team

- Status changed from New to In Progress

- % Done changed from 0 to 10

Thank you for your email.

We aim to respond to simple queries within 48 hours and more complex issues within 7 working days.

We are keen to hear about any data issues and experiences that you have as this will help us build the best possible knowledge database for the UKHLS and BHPS data sets.

Best wishes,
Understanding Society User Support Team

#2 - 10/11/2024 11:26 AM - Understanding Society User Support Team

- Category changed from Derived variables to Data documentation

- Status changed from In Progress to Feedback

- % Done changed from 10 to 50

- Private changed from Yes to No

Hello Irene,

There are a couple of approaches you could consider. For example, if you are looking to find the number of biological children in the household, you could use the indall/egoalt files. However, keep in mind that indall/egoalt is reported by one person in the household, so there is a possibility that the information may not always be accurate.

Here is an example of Stata syntax using egoalt:

```
qui {
glo data "" // released data file path
glo wn1 = 13

clear
  g long pidp = .
  tempfile all
  save `all', emptyok

forval i = 1/$wn1 {
  n di as txt "Wave `i'"
  loc w = substr("abcdefghijklmnopqrstuvwxy", `i',1) + "_"
}
```

```
use "${data}/`w'egoalt.dta", clear
contract pidp `w'relationship_dv
keep if `w'relationship_dv==9 // natural parent
drop `w'relationship_dv
ren _freq natpar_w`i'
n merge 1:1 pidp using `all', nogen
order natpar_w`i', last
tempfile all
save `all'
}
use `all', clear
}
```

To increase accuracy, another approach would be:

For participants with missing values in x_newborn, use the egoalt file to check how many times the respondent is listed as a "natural parent." Only use the egoalt value if the total number of biological children you've calculated is lower than the count in egoalt. Keep in mind that egoalt only includes children present in the household during that wave, so it's normal for your count to sometimes exceed this.

Alternatively, you could use the newborn file and count by pidp and c_newchno how many times the participant reported newborns during that wave. You can then use this information to fill in the missing x_newborn value.

I hope this information is helpful.

Best wishes,
Roberto Cavazos
Understanding Society User Support Team