

## Understanding Society User Support - Support #947

### the birthyear of all children (in and outside the household) born by a given female respondent

04/03/2018 12:23 PM - Nico Ochmann

<b>Status:</b>	Resolved	<b>Start date:</b>	04/03/2018
<b>Priority:</b>	Normal	<b>% Done:</b>	100%
<b>Assignee:</b>	Nico Ochmann		
<b>Category:</b>	Data management		
<b>Description</b>			
Hi Alita,			
this is my last question for you for a while, promised. I would like to get your general take on it though. It somewhat relates to my other open issue which you are working on.			
I would like to have the number of all children born by a given female respondent (my open issue for which you are currently checking my suggestion) and in addition their birthyear.			
I found this:			
<16 in household: w_birthy from w_child			
out of household: a_lchdoby, f_lchdoby (year of birth of non-resident biological child) from w_natchild			
Now I would like to generate a variable call it birthyear_1_dv for first kid for all female respondents for all seven waves with birthyear_1_dv = 0 if the women is childless.			
And then birthyear_2_dv for the second kid for all female respondents etc. or something like that.			
If you happen to have some general suggestions as to how to go about generating such variables, I would highly appreciate your input.			
Best wishes and thank you very much.			
Nico			

## History

### #1 - 04/03/2018 12:38 PM - Stephanie Auty

- Status changed from New to In Progress
- % Done changed from 0 to 10
- Private changed from Yes to No

Many thanks for your enquiry. The Understanding Society team is looking into it and we will get back to you as soon as we can.

Best wishes,  
Stephanie Auty - Understanding Society User Support Officer

### #2 - 04/10/2018 10:10 AM - Alita Nandi

- Status changed from In Progress to Feedback
- Assignee changed from Alita Nandi to Nico Ochmann
- % Done changed from 10 to 70

In Wave 1, a\_natchild will have birth year information for all children ever born to the respondent. After Wave 1, w\_newborn will have information about all births for female respondents. [For the IEMB sample, f\_natchild, followed by w\_newborn from onwards Wave 7.]

One approach is to gather information of all children's birthdates along with the adult/parent respondent pidp across the waves into one long file and sort by pidp and child's birthdate. keep only female respondents. Each row will represent one child of each female respondent. Create a birthdate variable for each child using Stata's ym() feature. Say, it is called kbdate.

```
bys pidp (kbdate): g chno=_n
g kbdate1=kbdate if chno==1
g kbdate2=kbdate if chno==2
```

etc

The first child will be youngest. Note this will not work for twins. You could tweak the birthdate for one of the twins so that one of them is older by one month before you do the above exercise. Create a flag to identify twins. After you have created kbdate1,... then for correct the kbdate for the twins. Also remember to impute the year and month of birth for those with missing dates. A general rule is to impute missing months to 6. If you cannot impute missing year then recode to system missing. This is a general guide. There may be glitches. Try it and let us know.

Hope this helps. If you have further questions please let me know.

Best wishes,  
Alita

**#3 - 04/12/2018 02:20 PM - Nico Ochmann**

Dear Alita,

my work is under way, I do have one question before I dive into it. You write above that a\_natchild has birthyear info for all children ever born to respondent. Unfortunately, and this might be an error on the website or not I only find the following variable that has birthyear for non-resident or out of household children:

```
fertilityhistory_w1.lchdoby  
Year of birth of non-resident biological child  
Universe  
if ( LNPrt > 1 | LPrnt = 1 ) (Parent of biological child) and if ( LChLv = 2,3 ) (Child non-resident or deceased)
```

As usual, I highly appreciate your help.

Best wishes.

Nico

**#4 - 04/13/2018 11:40 AM - Nico Ochmann**

Dear Alita,

please no hurry, I just want to write this because I want to have it out there. I am trying to find the birthyear of each child because I want to find out the number of children a immigrant women had before she came to the UK. For that purpose, I found for the IEMB f\_lchborn (in f\_natchild) which suits me very well. For the other subsamples, I need to compare the woman's year of arrival to the birthyear of the children (I guess I can leave the newborn out because they were born in the UK). Using a\_natchild and f\_natchild, I have an appended dataset with a\_lchdobm, a\_lchdoby and f\_lchborn as raw data and as you suggested the created variables kbdate1 etc.. Here is my problem, as I understand it the pidp is the identification for the parent in w\_natchild. If that is the case, I am not sure how to link/merge this file with my master dataset that contains all the information on the parents (indresp pidp and year of arrival to the UK). The problem is that the using dataset has the same pidp for each child which makes sense because a parent with three kids has three rows with the same pidp because it is the same parent. But how can I uniquely merge this using set with the master set? I hope I am making sense.

Thank you very much.

Nico

**#5 - 04/18/2018 09:49 AM - Alita Nandi**

- File Linking\_Parents\_Information.do added

Dear Nico,

Sorry, yes you are right.

1. birthyear of co-resident children can be identified using A\_INDALL and A\_MNPID
2. birthyear of non-resident children ever born can be identified using A\_NATCHILD
3. From Wave 2 onwards, date of birth of newborns can be identified using W\_NEWBORN (but also as most newborns live with their mothers, you can also identify them from INDALL)  
FOR IEMB, because it starts in Wave 6, the same rules will apply starting with F\_NATCHILD, F\_INDALL...

Step 1, is the same issue as you have mentioned in the last post: how to link parents information to that of the child. I am attaching a dofile to carry out this linking. Please let me know if there are any problems with this syntax.

Best wishes,  
Alita

**#6 - 04/21/2018 04:04 PM - Nico Ochmann**

Dear Alita,

first of all, thank you very much for your helpful code. I ran it and it works. And I did append the w\_parents files into one doing the following:

```
foreach w in a b c d e f g {  
use pidp mnpid `w'_mbirthy using `w'_parents, clear  
  
gen wave = strpos("abcdefg","`w'")  
renpfix `w'_  
  
save kids`w'wave, replace
```

```
}  
  
use kidsawave, replace  
  
foreach w in b c d e f g {  
  // append all seven waves of parental files  
  append using kids`w`wave  
}
```

Now, looking at the parental files by waves or in the appended form, there is something goofy going on. When I sort by mnpid and wave, and look at the data I observe the following:

The year of birth is the same within mnpid, but the pidp of the kid varies. For instance, the mom has five rows, and each row got the same birthyear although the pidp varies. And this holds true for every individual mom. I take from this, assuming that pidp identifies the kid, and mnpid the mom, that the birthyear is the birthyear of the mom and NOT of the kid. In addition, when I look at the w\_child file, which also contains information on the birth year of the kid(s), and run the same code replace w\_indall with w\_child, I obtain for the year of birth only missings. This might suggest that there is something wrong with the data.

Any comments are more than appreciated.

Thank you very much.

Nico

**#7 - 05/03/2018 12:20 AM - Alita Nandi**

Hi Nico,

In the resulting files w\_paernts there are 5 variables: pidp of the child (pidp), the mother and father pidp (mnpid fnpid) and the mother and father birthyear (w\_mbrthy w\_fbrthy). So for households where someone has more than one child, their birthyear will remain the same for each of the row representing each of their children with different pidp.

Best wishes,  
Alita

**#8 - 05/04/2018 10:25 AM - Nico Ochmann**

Dear Alita,

this is what I thought, that I now have the birthyear of the parents. What I do need, if you are so kind and reread what I wrote previously, is the birthyear of the Children!  
Where would I find that if at all?

Thanks Alita.

Nico

**#9 - 05/09/2018 08:26 PM - Alita Nandi**

Hi Nico,

In the first step when you are extracting the child's pidp and their parents's pidp (mnpid, fnpid) from INDALL, you can also extract the w\_birthy variable which will be the child's birth year.

For step 2, the birth year of the non-resident child associated with each adult is called a\_lchdoby. Note in this file the pidp refers to the parents' pidp.

Best wishes,  
Alita

**#10 - 05/09/2018 09:39 PM - Nico Ochmann**

Dear Alita,

thanks again for your kind reply. As to step 2, that is clear now, thanks a lot. With regard to the INDALL file, you claim that w\_birthy is the child's birth year. This brings me back to my #6 post. It seems to me that there must be something wrong with the data file. I write this because when I look at the w\_birthy in my dataset, I find the variable ranging from the early 20th century to the late 20th century. Note, these must be co-resident children's years of birth.

Sorry for being such a pain.  
Please let me know what you think.

Best wishes,  
Nico

**#11 - 05/11/2018 05:21 PM - Alita Nandi**

- File Linking\_Child\_Information.do added

Hi Nico,

The INDALL file includes everyone in the household. So, birthy can range from early 20th century. The code allows you to extract the children information and link that to their parents. But I have now added a few more lines of code and excluded father information, so it will produce a file containing mother pidp, child pidp, mother birth year and child birthyear. Is this what you were looking for?

Best wishes,  
Alita

**#12 - 05/15/2018 04:15 PM - Nico Ochmann**

Dear Alita,

you have been once again of tremendous help.

This is what I am looking for.

Thank you very much and sorry about the confusion.

Best wishes.

Nico

**#13 - 08/14/2018 04:28 PM - Stephanie Auty**

- *Status changed from Feedback to Resolved*

- *% Done changed from 70 to 100*

**#14 - 08/22/2023 01:01 PM - Understanding Society User Support Team**

- *Category changed from Data analysis to Data management*

**Files**

---

Linking_Parents_Information.do	1.17 KB	04/18/2018	Alita Nandi
Linking_Child_Information.do	1.07 KB	05/11/2018	Alita Nandi