Understanding Society User Support - Support #1797

Child birth years/fertility histories

10/31/2022 06:58 PM - Emily Humphreys

Status:	Resolved	**Start date:**	10/31/2022
Priority:	Normal	**% Done:**	100%
Assignee:			
Category:			

Description

Hello

I am trying to create datasets for each wave of the survey which include the date (year) of birth for each of a respondent's children.

I initially started with the ch1by_dv (DOB of oldest biological child) in the xwavedat file. I found that the number of respondents with oldest children born in each year seems to drop off sharply in the early 2000s, particularly for women. I would expect fewer people to have had their oldest child in very recent years, but not such a sudden drop or a gender difference. Can you explain why this would be the case and/or share the code (in stata) for the derived variable?

(I merged variables from a_indresp and xwavedat to tabulate a_sex_dv and ch1by_dv)

I've also tried using an adapted version of the code from support enquiry #947, intending to identify a years of birth for all resident and nonresident children from the w_indall and a_natchild files, but as my output on oldest children looks very different to the ch1by_dv output, I don't want to go too far with this until I understand where the difference has come from.

Many thanks

Emily

History

#1 - 11/01/2022 08:44 AM - Understanding Society User Support Team

- Status changed from New to In Progress

- % Done changed from 0 to 10

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.

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

Best wishes,
Understanding Society User Support Team

#2 - 11/14/2022 03:34 PM - Understanding Society User Support Team

- Status changed from In Progress to Feedback

- % Done changed from 10 to 80

- Private changed from Yes to No

Hello,

I looked at the age of first child for respondents by UKHLS waves 1-11 and did not find any difference in these distributions across waves. I also estimated the average (weighted) age at first birth and found that it is around 27 across all waves (25-26 for women, 28-29 for men). Does this answer your question? If not, please let us know.

Best wishes,
Understanding Society User Support Team

#3 - 11/14/2022 04:34 PM - Emily Humphreys

Hello

Thanks for your help. That isn't quite what I'm looking for - apologies for not explaining more clearly.

I am not trying to identify how old women were when their first child was born, I'm trying to find how many women in any given sample wave gave

birth to their first child in each calendar year (I'm also, later, hoping to identify the number of sample members giving birth in each calendar year, including second and subsequent births).

The ch1by_dv variable in the xwavedat file appears to be exactly what I am looking for to identify the number of sample members giving birth to their first child in each year. However, if I take a cross section of this variable from any given wave of the survey (say wave 2), I find that the number of women in the unweighted sample whose first child was born in each year of the 1990s is between 300 and 500, but this drops off sharply in 2000, 2001 and 2002 before stabilising at fewer than 100. I can't think of a reason why so may fewer women would have had their first child in the 2000s.

I initially thought it might be because co-resident children were excluded from ch1by_dv. However, that does not appear to be the case - the documentation on this variable says that it uses the indall and newborn data files as well as the natchild data file (https://www.understandingsociety.ac.uk/documentation/mainstage/dataset-documentation/variable/ch1by_dv).

A simplified version of my stata code is:

```
use pidp b_sex_dv using "$inpath\b_indresp", clear
save b_samplesize_test, replace
use pidp ch1by_dv using "$inpath\xwavedat", clear
save xw_samplesize_test, replace
merge 1:1 pidp using b_samplesize_test
keep if _merge == 3
drop _merge
save b_samplesize_test, replace
su
recode ch1by_dv -9/-1 = .
tabulate ch1by_dv b_sex_dv
```

Grateful for your help (if you could possibly share the code for ch1by_dv that would hopefully enable me to work out where I'm going wrong).

Thank you

Emily

#4 - 11/29/2022 10:44 AM - Emily Humphreys

Hello
Just hoping for an update on this query?
Thanks very much
Emily

#5 - 11/29/2022 05:36 PM - Understanding Society User Support Team

Sorry about the delay.

The first child information is based on these variables for UKHLS sample members:
(1) w_ch1bm w_ch1by4 from the w_indresp files
(2) w_lchlv w_lchdoby w_lchdobm from the w_natchild files for non-resident children (born before the start of the survey)
(3) w_birthy w_birthm (and w_m/fnpid) in w_indall files for resident children
(4) w_lchbm w_lchby4 in w_newborn files for new children being born during the survey
(5) If a child is mentioned via w_lprnt or w_nchild_dv but no other information is available then anychild_dv=1 and their dates of birth is set to 0.
For BHPS sample members this information is collected differently. That information is combined with this information collected while they were interviewed as part of the UKHLS to identify the first child for existing BHPS sample members.

As the syntax file is nested with other syntax files, it will take us some time to make transform it to a shareable format. If this information does not help, please let us know and we will look into transforming and sharing the file.

Best wishes,
Understanding Society User Support Team

#6 - 11/30/2022 12:47 PM - Olena Kaminska

Emily,

There may be a few reasons for this, but one of them maybe related to the definition of the population that our study covers. The largest part of the sample started in 1991 (there was a small boost in Scotland and Wales in 1999 and NI sample started in 2001). This means that by the beginning of the 2000s our data represents people who were present in the UK in 1991 (largely) or were born to them.
I am not sure about statistics of birth but according to ONS (https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/livebirths/bulletins/parentscountryofbirthenglandandwales/2021) around a third of all birth was to mothers who were not born in the UK, at least in recent years.
If such trend was similar in the beginning of 2000s, then only children born to those mothers who were either selected in 1991 or were born into a family who was in the UK in 1991 became part of our sample. I imagine that a large majority of mothers who were born outside the UK and gave birth in early 2000s were not in the country in 1991. Therefore our sample will be missing such children. Note, before this the problem is not visible as we may have selected such mothers in 1991 sample (and then maybe they gave birth in 1995 etc.).
Looking into official statistics for birth to immigrant mothers (and mother's time in the UK between immigration and birth), will give you an idea of the subgroup that our samples don't cover.
In addition, children born outside the UK that immigrated since 1991 would not be represented by our sample until UKHLS starts.

In simple words as a longitudinal study we represent people who have been living in the UK since 1991 (and then since each boost at the time we have a boost).
The representation renews at the time of each boost: wave 1 of UKHLS, wave 6 and wave 14.

Hope this helps,
Olena

#7 - 11/30/2022 01:40 PM - Emily Humphreys

Thanks so much Olena, really appreciate it.
I don't think this can be the cause of the issue, though, because it is in the UKHLS dataset, not only the BHPS dataset. Presumably people who had migrated since 1991 would have had a chance of selection in 2009, so they should appear in the sample from that point - is that correct?
I'll try to compare the variables you've given above with what I had come up with myself and let you know about whether I still think I'll need the syntax file.
Emily

#8 - 12/19/2022 10:02 AM - Emily Humphreys

Dear Olena
I've had a look into this and still haven't been able to get similar numbers so I think I do need the syntax file for ch1by_dv unfortunately. Will that be possible, and if it is, could you give me an idea of how long it might take?
Many thanks
Emily

#9 - 01/13/2023 10:48 AM - Emily Humphreys

Hello
Please could I have an update on whether this is possible and how long it might take?
Thank you
Emily

#10 - 10/24/2023 09:43 AM - Understanding Society User Support Team

- Status changed from Feedback to Resolved

- % Done changed from 80 to 100

Conversation continued via email