Understanding Society User Support - Support #1344

discrepancy in W_age_dv .

05/12/2020 05:41 PM - Vernon Gayle

Status:	Resolved		**Start date:**	05/12/2020
Priority:	Normal		**% Done:**	100%
Assignee:	Understanding Society User Support Team			
Category:	Derived variables			

Description

We notice a potential discrepancy in a_age_dv .

Dataset SN6614 downloaded 21:16 07/05/20.

University of Essex, Institute for Social and Economic Research. (2020). Understanding Society: Waves 1-9, 2009-2018 and Harmonised BHPS: Waves 1-18, 1991-2009. [data collection]. 12th Edition. UK Data Service. SN: 6614, http://doi.org/10.5255/UKDA-SN-6614-13

We have attached a Jupyter notebook with our Stata code and a pdf of the Jupyter notebook.

We have selected a sub-sample of children under age 11.

We selected on variable a_age_dv (keep if a_age_dv<11) Sample member's age at time of the interview, derived.

Number of children (kids) under 11 in wave a = 12,331; After removing missing values (drop if a_age_dv==-9).

Potential Inconsistencies Children Under 11

1.The salient examples are.'.

pidp 885684379 ; 886142027 and 68029939 appear to have errors in a_age_dv given their date of birth a_birthy
These respondents are all much older than 11 years old.

Potential Inconsistencies (all wave a)

There are 77,309 observations in a_indall.dta.

2,055 cases have different values for a_age_dv and a_dvage (age at interview and last birthday).

1. 205 cases identified if a_dvage>(a_age_dv+2) ie. age at last birthday is greater than age at interview by +2 years.

For example
pidp 68029939
a_hidp 68029923
a_pno 4
a_birthy 1993
a_intdaty_dv 2009
a_age_dv 5
a_dvage 15

2. 218 cases identified if a_dvage<(a_age_dv-2) ie. age at last birthday is lower than age at interview by -2 years.

For example
pidp 68372651
a_hidp 68372643
a_pno 2
a_birthy 1957
a_intdaty_dv 2009
a_age_dv 57
a_dvage 51

Potential Inconsistencies (i_inall)

1. In i_indall.dta 1,242 where i_age_dv is not equal to i_dvage.

pidp 68595023
i_hidp 70148816
i_pno 3
i_birthy 2003
i_intdaty_dv 2017
i_age_dv 16
i_dvage 13

Potential Inconsistencies (b_youth file)

1. In b_youth.dta 34 cases where b_age_dv is not equal to b_dvage

For example
pidp 204400535
b_hidp 206176002
b_pno 3
b_ypdoby 1999
b_intdaty_dv 2010
b_age_dv 15
b_dvag 11

Potential Inconsistencies (b_inresp file)

1. In b_indresp.dta 470 where b_age_dv is not equal to b_dvage.

For example

pidp 69059445
b_hidp 102700402
b_pno   1
b_birthy   1959
b_intdaty_dv   2010
b_age_dv   56
b_dvage 50

History

#1 - 05/13/2020 02:09 PM - Annette Pasotti

- Assignee set to Gundi Knies

- % 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. While we will aim to keep to this response times due to the current coronavirus (COVID-19) related situation it may take us longer to respond.

Best wishes,

Annette

#2 - 05/13/2020 02:13 PM - Annette Pasotti

- Status changed from New to In Progress

#3 - 05/14/2020 08:22 AM - Gundi Knies

- Category set to Derived variables

- Assignee changed from Gundi Knies to Vernon Gayle

- Target version set to X M

- % Done changed from 10 to 50

Dear Vernon,

do you also get inconsistencies when you use the origin variables used in creating w_age_dv, specifically w_doby_dv instead of w_birthy?

w_birty is collected in the enumeration grid or fed forward from the sampling file, and then used to compute w_dvage during the interview. There sometimes are data entry errors - such as a number not registering correctly making a 15 year old a 5 year old, or the DOB of another household member being entered in the household grid. This is what the derived w_doby_dv w_dobm_dv w_age_dv variables try to address drawing on a wider range of information available than is the case during the interview, including comparison across earlier and later waves.

Please let us know if this solves the issue.

Best wishes,
Gundi

#4 - 05/29/2020 04:24 PM - Alita Nandi

- Status changed from In Progress to Feedback

#5 - 10/13/2021 11:45 AM - Understanding Society User Support Team

- Status changed from Feedback to Resolved

- Assignee deleted (Vernon Gayle)

- % Done changed from 50 to 100

#6 - 02/19/2026 06:18 PM - Understanding Society User Support Team

- Assignee set to Understanding Society User Support Team

- Private changed from Yes to No

Files

error_age_20200512_vg_v3.pdf	584 KB	05/12/2020	Vernon Gayle
error_age_20200512_vg_v3.ipynb	134 KB	05/12/2020	Vernon Gayle