

## Understanding Society User Support - Support #369

### missing paju maju

04/30/2015 12:22 PM - Carolina Zuccotti

<b>Status:</b>	Closed	<b>Start date:</b>	04/30/2015
<b>Priority:</b>	High	<b>% Done:</b>	100%
<b>Assignee:</b>			
<b>Category:</b>	Data documentation		
<b>Description</b> Hello, From what I understood in one of the replies you gave, paju (and maju) were asked only in Wave 1. But then, what are b_paju c_paju and d_paju? From the question/universe it seems as if it was asked in all waves to all respondents:  Universe if (ff_ivlowlw = 2 3 MIS) & ff_everint <> 1 //proxy last wave, non-interviewed adult or new entrant never interviewed, excluding rising 16 year olds  However, the number of missing for Waves 2, 3 and 4 is much higher (some new entrants are actually responding to the question, while others aren't). Why is this the case?  Looking forward to your response. Thanks in advance. Best regards, Carolina			

### History

#### #1 - 05/04/2015 03:50 PM - Redmine Admin

- % Done changed from 0 to 50

Only asked of adults not interviewed before, why the frequencies look sparse for the later waves. The cross-wave file, XWAVEDAT, contains for convenience a consolidated version of these and other stable characteristics;

<https://www.understandingsociety.ac.uk/documentation/mainstage/dataset-documentation/wave/xwave/datafile/xwavedat/variable/maju>

Jakob

#### #2 - 05/05/2015 10:08 AM - Carolina Zuccotti

Hi Jacob,

Thanks for your message. I cannot find the file. Where should it be?

I think, however, that my question is a different one. I have created "paju" and "maju" variables myself, with information from all 4 waves (I know that this is background information and, hence, stable information, and that is why it is b\_paju c\_maju etc. are sparse in Waves 2, 3 and 4).

However, the number of missing "paju" and "maju" is much higher in waves 2, 3 and 4 (20% or more) than in wave one (less than 2%). The only way to explain this is that new entrants are not being asked the question. See also: <https://www.understandingsociety.ac.uk/support/issues/159>. Here I understand that parental information is collected in Wave 1; and then in Wave 8, for entrants between Wave 2 and Wave 8.

Am I missing something?

Thanks,

Carolina

This is the syntax:

I merge all 4 waves and then create the following variables:

```
gen a_flag=1 if a_psu!=.
```

```
gen b_flag=1 if b_psu!=.
```

```
gen c_flag=1 if c_psu!=.
```

```
gen d_flag=1 if d_psu!=.
```

```
. gen paju=.
```

```
(73119 missing values generated)
```

```
. replace paju=a_paju
```

```
(50994 real changes made)
```

```
. replace paju=b_paju if paju<0 | paju==.
```

```
(18599 real changes made, 1452 to missing)
```

```
. replace paju=c_paju if paju<0 | paju==.
```

(7181 real changes made, 2804 to missing)

. replace paju=d\_paju if paju<0 | paju==.  
(6224 real changes made, 2327 to missing)

. gen maju=.  
(73119 missing values generated)

. replace maju=a\_maju  
(50994 real changes made)

. replace maju=b\_maju if maju<0 | maju==.  
(18523 real changes made, 1417 to missing)

. replace maju=c\_maju if maju<0 | maju==.  
(7150 real changes made, 2795 to missing)

. replace maju=d\_maju if maju<0 | maju==.  
(6187 real changes made, 2291 to missing)

. lab values paju a\_paju

. lab values maju a\_maju

end of do-file

. tab paju

paju	Freq.	Percent	Cum.
missing	5	0.01	0.01
inapplicable	11,028	16.34	16.34
proxy respondent	2,412	3.57	19.92
refused	1	0.00	19.92
don't know	15	0.02	19.94
father working	45,322	67.14	87.08
father not working	3,844	5.69	92.77
father deceased	2,410	3.57	96.34
father not living with respondent so do	2,469	3.66	100.00
Total	67,506	100.00	

. tab paju if a\_flag==1

paju	Freq.	Percent	Cum.
inapplicable	143	0.29	0.29
proxy respondent	557	1.13	1.42
refused	1	0.00	1.42
don't know	1	0.00	1.42
father working	40,833	82.79	84.22
father not working	3,378	6.85	91.07
father deceased	2,237	4.54	95.60
father not living with respondent so do	2,169	4.40	100.00
Total	49,319	100.00	

. tab paju if b\_flag==1

paju	Freq.	Percent	Cum.
inapplicable	9,359	18.45	18.45
proxy respondent	1,163	2.29	20.74
don't know	1	0.00	20.74
father working	33,991	67.00	87.74
father not working	2,693	5.31	93.04
father deceased	1,791	3.53	96.57
father not living with respondent so do	1,738	3.43	100.00
Total	50,736	100.00	

. tab paju if c\_flag==1

paju	Freq.	Percent	Cum.
	inapplicable	9,938	20.96
	proxy respondent	1,421	3.00
	refused	1	0.00
	don't know	3	0.01
	father working	30,555	64.45
	father not working	2,367	4.99
	father deceased	1,564	3.30
father not living with respondent so do	1,563	3.30	100.00
	Total	47,412	100.00

. tab paju if d\_flag==1

paju	Freq.	Percent	Cum.
	missing	5	0.01
	inapplicable	11,028	23.39
	proxy respondent	2,412	5.11
	refused	1	0.00
	don't know	15	0.03
	father working	28,645	60.74
	father not working	2,188	4.64
	father deceased	1,440	3.05
father not living with respondent so do	1,423	3.02	100.00
	Total	47,157	100.00

### #3 - 05/05/2015 11:24 AM - Redmine Admin

The XWAVEDAT file will be in the same folder as the other Understanding Society data files downloaded from UKDS. The older issue refers to the BHPS study.  
Jakob

### #4 - 05/05/2015 11:51 AM - Carolina Zuccotti

Hi Jakob,  
Sorry but I am a bit confused. From what I see the file Xwavedat has more information on parental background (as compared to the paju/maju I created by adding a\_paju + b\_paju + c\_paju + d\_paju).  
Why is that? Why I do get fewer cases with parental background if I do not include the Xwavedat? Does this mean that in order to work with background variables we always have to use the Xwavedat?  
I would appreciate very much if you could explain this to me with some detail (or refer to a document), since I am really puzzled at the moment.  
Thanks in advance,  
Carolina

### #5 - 05/05/2015 11:59 AM - Redmine Admin

XWAVEDAT also consolidates data from the surviving BHPS sample in the study, see e.g.  
<https://www.understandingsociety.ac.uk/documentation/mainstage/dataset-documentation/wave/xwave/datafile/xwavedat/variable/maju>  
Jakob

### #6 - 05/05/2015 12:10 PM - Carolina Zuccotti

Ok, got it. Thanks.

### #7 - 05/05/2015 12:49 PM - Redmine Admin

- Status changed from New to Closed  
- % Done changed from 50 to 100