

## Understanding Society User Support - Support #968

### Wording of Annual History Questions

04/25/2018 01:20 PM - Liam Wright

<b>Status:</b>	Resolved	<b>Start date:</b>	04/25/2018
<b>Priority:</b>	Normal	<b>% Done:</b>	100%
<b>Assignee:</b>	Liam Wright		
<b>Category:</b>	Data documentation		
<b>Description</b>			
<p>*Hi,</p> <p>I've also raised this issue in <a href="#">#957</a>, but am not sure how to change assignee and not sure if you will get it.*</p> <p>Hi Stephanie,</p> <p>Thanks for the information. I'd rather not exclude these participants as, though they are small in number, people who state being unemployed or being economically inactive are interesting to study in themselves.</p> <p>I don't think cjob reflects the same job as jbsoc00 because someone could answer "2: No" to cjob and mention other, new jobs later in nextjob.</p> <p>I'm still unclear about the wording of the four questions: empstendd, nxtst, cjob and nextstat. The questionnaires state the wordings to be:</p> <p>empstendd: On what date did you stop being [ff_JBSTAT] {if NotEmpChk = 2} / working in the job you were doing on [ff_IntDate] {if EmpChk = 2} ?</p> <p>nxtst: Immediately following that period of [ff_JBSTAT] {if NotEmpChk = 2} / job {if EmpChk = 2}, did you have a period of paid work or did you do something else?</p> <p>cjob: Was that {if NxtSt = 1} / your {if JbSamR = 2   SameJob = 2} next job your current job?</p> <p>nextstat: Immediately following that period of [NxtStElse] {if NxtSt = 2} / period of [NextStat(i-1)] {if NxtSt = 2 and 2nd or subsequent loop} / job {if NxtSt = 1} , did you have a period of paid employment or did you do something else?</p> <p>(These wordings are all from Wave 6, but all wordings appear to be the same in each wave except for empstendd in Wave 7.)</p> <p>Given the if conditions in each question are not mutually exclusive, which phrasing condition will take precedence? Is it the leftmost condition - e.g. for nxtst "Immediately following that period of [ff_JBSTAT], did you have a period of paid work or did you do something else?" where NotEmpChk = 2 - or is it some other rule?</p> <p>There is a second problem with nextstat. It is possible to be asked that question without being asked nxtst - which the wording depends on - as one can be routed to it if (jbsamr = 2   samejob =2) &amp; cjob = 2. What question would be asked in this case?</p> <p>Thanks,</p> <p>Liam</p>			

#### History

##### #1 - 04/27/2018 03:21 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 - 04/27/2018 04:10 PM - Liam Wright

Thanks, Stephanie.

Just to add to this, I'm interested in how the following fed forward variables are calculated: ff\_ivlowlw, ff\_everint, ff\_jbstat, and ff\_emplw.

Are ff\_jbstat and ff\_emplw fed forward from full interviews only (i.e. where ivfio==1) or can they be fed forward from other types of interview (e.g. proxy)? Which types of interview does ff\_everint refer to - is it again full interviews (ivfio==1) only? Also, are any of these variables calculated from BHPS responses or are they just from UKHLS responses? I'm interested in knowing whether ff\_jbstat may be drawn from the BHPS.

Thanks,

Liam

### #3 - 05/10/2018 11:28 AM - Stephanie Auty

- File example code for annual employment history data.do added
- Category set to Data documentation
- Status changed from In Progress to Feedback
- Assignee set to Liam Wright
- % Done changed from 10 to 70

Dear Liam,

In general, it seems useful to stress that for much of the annual employment history module, there are two parallel streams of information, stored in parallel sets of numbered variables, but where one corresponds to work spells and one corresponds to non-work spells. A lot of confusion could be helped by bearing this in mind and cross-tabbing to see the conditions under which participants have a 'not applicable' for a given variable.

Cjob is a binary classification of currentness, specifically for the first job after the fed-forward job. This is why, if you cross-tab with some other things, you will see that people only have a positive (i.e., not inapplicable) value for it if the fed-forward activity is no longer current. There is a currentness flag for further activities after this, but from the second activity after the fed-forward one the naming convention changes. So for example at wave 4, the remaining currentness flags are variables with the names d\_currjob1 - d\_currjob6, going forward in time if these are employment spells, but d\_currstat1 - d\_currstat6 if these are non-employment spells. So for any pair of w\_currjobn and w\_currstatn, only one should have a positive (i.e., not inapplicable) value. You can check this by doing some cross-tabbing.

In contrast, there is only one variable called w\_jbsoc00, i.e. this is not part of a repeated loop, because this picks up the SOC code only of the activity current on the day of the interview.

Empstendd contains information on the date when the fed-forward activity ended, whether this was an employment or non-employment spell.

Apart from for the small group already discussed in your previous question, either nextst or nextstat should be asked to the participant but not both, such that apart from that small group nobody should have a positive value for both.

As stated above, w\_cjob is a currentness flag for whether the first EMPLOYMENT activity after the fed-forward one is still current at the later interview, so the participant will only have a positive value for this where they have values of nextst or nextstat corresponding to work activities. For where the first activity was a non-employment activity, the same information is contained in w\_cstat. For the second activity following the fed-forward one, the corresponding currentness flag variables change to two numbered sequences of the form w\_currjobn and w\_currstatn.

Regarding the second problem you mention with nextstat, yes, for a handful of cases that happens. Specifically, this can happen when a) the fed-forward activity was an employment spell, and b) the person is doing either the same job for a different employer or a different job for the same employer, i.e. (jbsamr = 2 | samejob = 2), and c) they answered 'not current' to w\_cjob, meaning that in some sense they regard the job as having ended even though it either the same job for a different employer or a different job for the same employer. In this case, there is no information in the slot corresponding to the first activity after the fed-forward one, but there is information in the slot corresponding to the second, third activity etc. after the fed-forward slot. But this is really a vanishingly small proportion of cases, so you can decide what they want to do with this. This can be conceptualised as the fed-forward activity being in effect current until the start of the next activity for which you have a start date.

Whether feed forward variables are calculated from proxy or only full interviews can be very easily checked by merging the relevant files and cross-tabbing to see.

The UKHLS team member who has spent a lot of time working with this data and provided the information for this response would like to share with you the attached code, which is from her exploration of all of the discrepancies and complexities. It was specifically for BHPS participants at wave 4 but you should be able to adapt the code for the waves you are using.

Best wishes,

Stephanie Auty - Understanding Society User Support Officer

### #4 - 05/16/2018 02:31 PM - Liam Wright

Dear Stephanie,

Thank you for taking the time to respond but you haven't answered the questions I asked. I would like to know the wording of the questions empstendd, nextst, cjob and nextstat, that is all. I'm aware of how participants come to be asked these questions or not, which you can see in the routing diagrams I've posted alongside previous questions I've raised about this. I need to know these wordings so I can understand which activity participants are answering about when asked. There must be a document somewhere which provides this information.

You note these problems affect few people. First, they don't: 5,252 individuals are asked notempchk and empchk in the same wave at least once between Waves 2 to 7. Second, these are important participants to keep in the data as these are people who state carrying out multiple activities – they aren't a random selection of participants and they are interesting in themselves. Third, I think it is unethical to throw participant's data away for no good reason when they have agreed to be part of a study in order to contribute to scientific knowledge.

In your response, you note it is easy to find out where the fed forward values come from. It isn't as the data are processed which means ff\_jbstat and jbstat[t-1] may be different (assuming an individual is interviewed in two adjacent waves). I've tried to get match the two up and the best rule I can come up with is that the values are fed forward from full interviews (ivfio==1 | ivfio\_bh==1) or from telephone interviews in the BHPS only (ivfio\_bh==3) (from the xwave datasets). These are very, very similar, but do not match up perfectly and there are some participants with ff\_jbstat who do not meet this condition. So, I would like to know the rule used or please be pointed in the direction of data which gives the wave from which ff\_\* values are taken from. Surely this information must be available somewhere?

Best wishes,

Liam

**#5 - 05/16/2018 02:52 PM - Stephanie Auty**

- Status changed from Feedback to In Progress

- Assignee changed from Liam Wright to Stephanie Auty

**#6 - 05/29/2018 03:57 PM - Stephanie Auty**

- Status changed from In Progress to Feedback

- Assignee changed from Stephanie Auty to Liam Wright

Dear Liam,

You can see the wording of the questions on the questionnaires, here:

<https://www.understandingsociety.ac.uk/documentation/mainstage/questionnaires>

The questions are asked in the order listed in the questionnaire so it is possible to work out the sequence. If the first-listed question was applicable then the respondent would be taken through that loop until completed and then, if the second-listed question was also applicable, the respondent would then be taken through that loop as well until complete. If a respondent was routed into being asked a question that really does not apply to them interviewers tend to quit the loops as quickly as possible.

A quick check looping through the waves shows that at each wave, roughly 5% of people who went through the empchk loop were also sent through the notempchk loop. As a proportion of the whole sample, this is more like 2.5% per wave.

You are correct that the people in this group are not random. The reason why someone gets sent through both loops is if at the previous wave they offered conflicting information in a particular way – if they answered jbhas or jboff in a way which indicated they were working, but gave an answer for jbstat which indicated otherwise. Thus, they are likely to be people in slightly non-standard working situations. Depending on the research question that may be something of particular importance or interest, but it is up to you to decide how to reconcile this discrepancy. The UKHLS team member I mentioned in the previous reply lets jbstat take precedence, since that seemed to pick up a more long-term aspect of identity, rather than reflecting an odd job in the previous week.

If you want to check this, my colleague used the following code:

```
use d_indresp.dta, clear
merge 1:1 pidp using e_indresp.dta
count if e_empchk!=.
count if (e_empchk==1 | e_empchk==2) & (e_notempchk==1 | e_notempchk==2)
fre d_jbhas d_jboff d_jbstat
list d_jbhas d_jboff d_jbstat if (e_empchk==1 | e_empchk==2) & (e_notempchk==1 | e_notempchk==2)
```

Mismatches between ff\_jbstat and previous jbstat can occur because the feed forward file is created on the basis of unchecked and sometimes incomplete interview data: A small number of interviews for the concluding waves are returned from the fieldwork after data collection for the new wave has commenced. E.g. if an individual was employed in Wave 1 (a\_jbstat==1), and b\_jbstat==4 in Wave 2 but their interview came in too late, at Wave 3, c\_ff\_jbstat will be 1 but their b\_jbstat will be 4. It seems that the important thing for you is that ff\_jbstat is the information that guided the routing in the questionnaire. It is not edited post-hoc to match the information that is provided in the last wave's interview.

If anything is not clear it may be easier to discuss by telephone. Please let us know if you would like to talk about it.

Best wishes,

Stephanie Auty - Understanding Society User Support Officer

**#7 - 08/02/2018 11:40 AM - Liam Wright**

Hi Stephanie,

Thanks for this and apologies for delay in replying. I've also replied to Query [#957](#) which restates my initial question.

I think we are getting into a discussion where we are confusing wording with routing. I understand the routing *to* the questions, but not the wording *of* the questions.

Understanding the wording is important because only then I can understand what activity a participant is answering about and what activity this follows from chronologically. This isn't a question of deciding precedence. It's a question of getting the correct start and end dates for spells.

If I could just get the correct wording when participants meet both conditions (e.g. `NxtSt=1 & (JbSamR = 2 | SameJob = 2)` for `cjob`) or no condition (e.g. `NxtSt!=1 & NxtSt!=2` for `nextstat`) for the question above, I'll understand all I need to know.

Thanks,

Liam

**#8 - 08/06/2018 12:52 PM - Stephanie Auty**

Dear Liam,

I'll reply on issue [#957](#) to keep all the answers together.

Best wishes,  
Stephanie Auty - Understanding Society User Support Officer

**#9 - 08/14/2018 05:40 PM - Stephanie Auty**

- *Status changed from Feedback to Resolved*
- *% Done changed from 70 to 100*

**Files**

example code for annual employment history data.do	29.3 KB	05/10/2018	Stephanie Auty
--	---------	------------	----------------