

Understanding Society User Support - Support #2241

Link COVID-19 Youth postal survey waves (4, 6, 8) to COVID-19 Adult surveys

04/17/2025 08:51 AM - James Laurence

Status:	New	Start date:	04/17/2025
Priority:	Normal	% Done:	80%
Assignee:			
Category:	Data management		
<p>Description</p> <p>Hi there,</p> <p>I hope this message finds you well. I am trying to link respondents in the COVID-19 Youth postal surveys (cd_youth_p, cf_youth_p, ch_youth_p) to the files on their 'parents' in the adult individual response files (ca_indresp_w, cb_indresp_w ... ci_indresp_w). I basically want to be able to link in information on the parents and households of respondents in the COVID-19 youth surveys. However, I am struggling to match all respondents in the youth files to their parents in the adult files, but may be going about it all wrong.</p> <p>Taking the wave 4 (cd_youth_p) youth file as an example, as far as I can see, the youth file only contains two identifier variables: pidp_c and cd_aid_dv. However, the adult file (e.g., cd_indresp_w) only contains pidp and cd_aid_dv.</p> <p>However, I found that the CHILD files (e.g., cd_schild_w) contain a pidp_c. They also contain pidp_m and pidp_f, the identifiers for a child's 'mother' and 'father'. So, using this I created a new pidp variable:</p> <pre>clonevar pidp = pidp_m replace pidp = pidp_f if pidp_m==--8</pre> <p>Which hopefully gives me a complete pidp identifier for a child's parent.</p> <p>Using this new identifier, I have found I am able to perfectly merge the child file (e.g., cd_schild_w) to the adult file (e.g., cd_indresp_w) i.e., all children in the child file are matched to an adult in the adult file, using the pidp identifier.</p> <p>So, now I have a file that contains both pidp_c for the child and pidp for the adult. My plan was to use this to merge with the YOUTH file (cd_youth_p). Therefore, I would match the pidp_c from the YOUTH file with the pidp_c in the new merged adult/child dataset, to get the pidp for the respondents in the YOUTH file.</p> <p>However, the problem I'm having is that there appears to be respondents in the YOUTH file (cd_youth_p) that are not present in the child file. So, if I merge the youth file (cd_youth_p) with the child file (cd_schild_w) based on pidp_c, there are n=416 cases in the YOUTH file that are not merged with the child file.</p> <p>Even if I merge the YOUTH file (cd_youth_p) with all available child files (cx_schild_w, where x=a, d, e, f, g) I am still left with n=237 cases from the youth file that don't appear to have equivalent records in the child files (as a note, I'm focusing on the web surveys - not telephone). As such, there are around n=237 cases in the YOUTH file for whom I cannot get their pidp, and thus cannot link them to the adult files. If I take the pidp_c of one of these unmatched cases and search for it through all the CHILD files I can't find equivalent records.</p> <p>I may be going about this all wrong. As I mentioned at the start, my primary aim is link records in the YOUTH files to information in the ADULT files on their 'parents' and 'households'.</p> <p>Any help or guidance would be hugely appreciated.</p> <p>Thanks so much in advance,</p> <p>James</p>			

History

#1 - 04/17/2025 06:48 PM - Understanding Society User Support Team

- File youth_covid_pub.do added
- Category set to Data management
- Assignee deleted (Understanding Society User Support Team)
- % Done changed from 0 to 80

Hi James,

The way you retrieved the responsible adults' identifiers for the children in the youth file is correct.

However, the wave 4 paper youth questionnaire was posted to the 4,139 responsible adults in households that were issued the July 2020 COVID-19 web survey and that contained children in the relevant age band.

As a result, some households returned completed youth questionnaires even though no adult interview was completed at the same address in wave 4, so these addresses do not appear in cd_indresp_w. You can confirm this with the attached code.

The resulting non-matches are children who cannot be linked to an adult respondent in wave 4, although you may be able to match them to adults in other COVID-19 waves or in the main-stage panel if that is helpful for your project.

I hope this helps.

Best wishes,
Piotr Marzec
UKHLS User Support

#2 - 04/17/2025 06:48 PM - Understanding Society User Support Team

- Private changed from Yes to No

Files

youth_covid_pub.do	773 Bytes	04/17/2025	Understanding Society User Support Team
--------------------	-----------	------------	---