# **Understanding Society User Support - Support #2260**

## Request for explaining the different missing rates of variables across all waves

06/14/2025 03:25 PM - Evan Zhang

Status: Feedback Start date: 06/14/2025

Priority: Urgent % Done: 50%

Assignee: Understanding Society User Support Team

Category: Data documentation

## Description

Dear Understanding Society Team,

I am currently working with data from the UKHLS and have some questions regarding response patterns across different waves. Specifically:

- 1. For the GHQ12 variable (scghq2\_dv), there appear to be large differences in the rates of missing, inapplicable, and proxy responses across waves. For example, the total rate of these responses in Wave 1 is about 20%, but the total rate in Wave 14 is about 5%.
- 2. Similarly, for the variables jbsoc00, jbsoc10, and jbsoc20, we observed significant differences in the rates of missing responses across waves. For example, the missing rate of jbsoc20 in Wave 14 is about 27%, the missing rate of jbsoc10 in Wave 14 is about 6%, and the missing rate of jbsoc00 in Wave 14 is about 1%.

We have reviewed the "Main Survey User Guide" but could not find specific explanations for these discrepancies. We would appreciate it if you could clarify the reasons for these differences. Thank you for your time and assistance.

Best regards, Evan Zhang

#### History

## #1 - 06/18/2025 04:03 PM - Understanding Society User Support Team

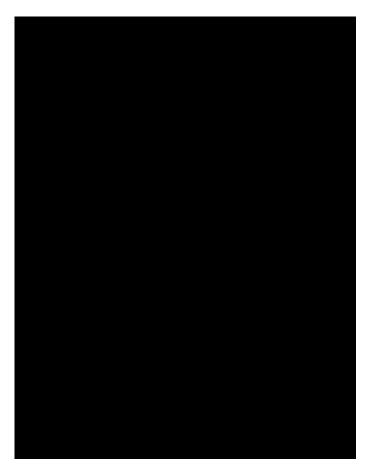
- File clipboard-202506181601-dnhew.png added
- Category set to Data documentation
- Status changed from New to Feedback
- % Done changed from 0 to 50
- Private changed from Yes to No

### Hello Evan

In Waves 1 and 2, the self-completion questionnaire, including the GHQ questions, was administered on paper. We've observed a relatively high level of missing responses for certain GHQ items in these early waves. However, the proportion of missing data decreases steadily in later waves, as shown in the table below:

Variable: scghq2\_dv

06/20/2025



To identify valid responses, it's important to review the Question Universe, which specifies eligibility for each question. In this case, scghq2\_dv is a derived variable created using the following items: scghqa to scghql.

You can find information on the universe for each question using the Mainstage Variable Search. For example, here is the link for scghqa, and if you scroll the page, under "Question asked in the latest wave," you will see the universe specification: <a href="https://www.understandingsociety.ac.uk/documentation/mainstage/variables/scghqa/">https://www.understandingsociety.ac.uk/documentation/mainstage/variables/scghqa/</a>

If you're interested in the syntax used to construct the scghq2\_dv variable, it's available here: https://www.understandingsociety.ac.uk/wp-content/uploads/documentation/main-survey/syntax/stata/qhg\_dv.do

Regarding occupation variables like jbsocXX, these are provided by the fieldwork agency and are already coded. At the moment, there is no official process for mapping SOC00 codes to SOC10 or SOC20, but it is something that could be explored further. If you'd like to carry out the conversion yourself, you can use the coding index provided by the ONS here:

https://www.ons.gov.uk/methodology/classificationsandstandards/standardoccupationalclassificationsoc/soc2020/soc2020volume2codingrulesandconventions

Lastly, if you're interested in survey response rates, further details can be found in the following resources:

- Main survey user guide: https://www.understandingsociety.ac.uk/documentation/mainstage/user-guides/main-survey-user-guide/response-rates/
- · Response tables:

https://understandingsociety.ac.uk/wp-content/uploads/documentation/user-guides/6614 main survey user guide response tables.pdf

I hope this information is helpful.

Best wishes,

Roberto Cavazos Understanding Society User Support Team

## **Files**

clipboard-202506181601-dnhew.png 14.2 KB 06/18/2025 Understanding Society User Support Team

06/20/2025