

UK Data Archive SN 6614 – Understanding Society Waves 1-12: Revision November 2023

Date: 1st November 2023

This document describes revisions to the Understanding Society Waves 1-12 datasets, including Harmonised BHPS datasets, since the Waves 1-12 release in November 2022. The User Guides and on-line documentation can be consulted for a study-wide perspective and variable level information, respectively.

New variables

The following crosswave design weights, where the mother's design weight is also applied to OSM babies, have been created and added to xwavedat:

- psnenu_xd – UKHLS design weight
- psnengp_xd – GPS design weight
- psnenub_xd – UKHLS+BHPS design weight

The following contact flag variables, which were previously missing from Wave 12, are now included in l_indresp_ip:

- l_ctadd1_code – stable contact address line 1 given
- l_ctadd2_code – stable contact address line 2 given
- l_cttown_code – stable contact town given
- l_ctcnty_code – stable contact county given
- l_ctpcode_code – stable contact postcode given
- l_cttel1_code – stable contact phone 1 given
- l_cttel2_code – stable contact phone 2 given
- l_ctemail_code – stable contact email given
- l_remail_code – participant email given
- l_remail2_code – participant email second entry given
- l_rhland_code – participant landline given
- l_rphmob_code – participant mobile given
- l_rphwrk_code – participant work number given

The variable l_cvverniov, COVID-19 questionnaire version used at interview, was not previously released but has now been included in l_indsamp.

Variable corrections

The variable w_kidlang_all, childhood language, which previously was only released in Wave 2, has been made consistent by incorporating the coded variable w_kidlango_code, other childhood language, and released in Waves 6 and 8 as well. w_kidlango_code has now been dropped as it is superfluous. w_kidlang_all is available in the Special Licence version only. The EUL version of the variable w_kidlang, childhood language, is retained in the same waves.

The variable w_intnum, interviewer number, in w_hhresp, w_callrec and w_hhsamp for Waves 9 to 13, has been re-populated and may have changed from the previous release due to the data provided by the fieldwork agency being re-supplied from a slightly different source.

The coding frames of the following variables, all in `l_indresp`, have been corrected to be consistent with other vote-related variables. Values 14-41 have been recoded to 30-57 respectively:

- `l_vote3_all`, Party would vote for tomorrow
- `l_vote4_all`, Which political party closest to
- `l_vote8_all`, Party voted for in last general election
- `l_voteeuroparl_all`, EU Parliament elections: other party

Renamed variables

The variables `w_reasendoth_code_1`, empl spell no. 1: Job finished for other reasons, to `w_reasendoth_code_9`, empl spell no. 9: Job finished for other reasons, in `w_indresp`, have been renamed to match the correct standard naming convention, i.e. `w_reasendoth1_code` etc. Waves 8 to 12 are affected.

Dropped variables

The variable `l_xpcnto_code` in `l_indresp` was mistakenly released and has been dropped from the current new version. It had no labels or frequencies. The correct variables to use are `l_xpcnt` and `l_xpcnt_all`, both of which are unchanged.

The variable `vote8_all`, Party voted for in last general election, has been removed from `l_indresp`. This variable has no wave prefix. The same variable, with the wave prefix, is the one that should be used and is unchanged.

The variable `w_kidlango_code`, other childhood language, has been dropped from `w_indresp` for waves 6 and 8 to be consistent with the approach taken in Wave 2 which is for it to be incorporated into `w_kidlag_all`.

The variable `bw_fihhmn`, Household income: month before interview, has been dropped from `bw_indresp` as household income variables should not appear in `bw_indresp` or `w_indresp` files.

Special Licence geographical datasets

The variable `w_oscty`, county, has been included in the `w_oslaua` files for all waves in the Local Authority Districts dataset (SN 6666) to complement the pre-existing `w_oslaua` (Local Authority District) variable. These have been taken from the relevant ONSPD utilising the same methodology that already exists. The variable contains the upper tier code (county) where applicable, i.e. for two tier authorities. Pseudo codes are used for English Unitary Authorities, and for Wales and Scotland where the second tier is not applicable.

Scottish Intermediate Zones (IZ) have been included for Waves 1 to 5 of the MSOA 2011 dataset (SN 7249) where they were previously missing. They were not included before because they were not made available in the ONS Postcode Directory until the May 2015 version, and an earlier version was previously used for those waves.

New data files for all waves have been provided in new datasets for LSOA 2021 (SN 9169) and MSOA 2021 (SN 9170). The May 2023 version of the ONS Postcode Directory (ONSPD) has been used to generate the data for all waves. Scottish DZs and IZs are not included yet as they are not available in ONSPD's. Please note that Output Area Codes and Urban / Rural Indicators for census 2021 are not yet available for any country and therefore no new datasets for these have been produced.

Miscellaneous changes

Various minor data corrections have been made.

Various missing and incorrect variable and value labels have been added or corrected as appropriate.