UK Data Archive SN 6614 – Understanding Society Waves 1-7: Revision 2018

Date: 3nd July 2018

This document describes revisions to the Understanding Society Waves 1-7 datasets, including Harmonised BHPS datasets, last fully revised in November 2017 with further between-edition adjustments made in January and March 2018. The User Guides and on-line documentation can be consulted for a study-wide perspective and variable level information, respectively.

Further BHPS Harmonisation Changes

There are some revisions to the BHPS data on which the harmonisation is based, including corrections to some household income variables and to the sample status variable (b*w*_sampst) from Wave 11 onwards. This version of the BHPS data includes the net income data currently available separately in SN 5909. These are data added to b*w*_hhresp and b*w*_indall files.

All waves of the BHPS hhsamp file have been newly harmonised for this release.

The suffix _bh has been removed from all waves of bhps file names.

All Special Licence geography datasets now contain geographies for the harmonised BHPS waves where they previously existed for the stand-alone BHPS dataset. They have been partly harmonised in so much as the file and variable names have been amended to reflect the harmonised naming conventions. The methodology used, however, has not been harmonised. For more details see the understanding_society_mainstage_panel_geographical_lookup_tables file available on the UKDS discover website for any Special Licence geography dataset.

The suffix _bh has been removed from the variables in the following files: b*w*_jobhist, b*w*_lifemst, b*w*_cohabit, b*w*_marriag, b*w*_childad, b*w*_childnt, b*w*_jobhstd and b*w*_lifejob.

The suffix _bh for all BHPS only variables has been removed in the xwavedat and xwaveid files.

New harmonised variables have been created in previously harmonised waves. These include:

b*w*_yppkmp (in b*w*_youth)	how much money did you receive last week to spend on yourself?
b*w*_yphlhtcm_dv (in b*w*_youth)	Height in cm, derived
b*w*_yphlwtkg_dv (in b*w*_youth)	Weight in kg, derived
b*w*_ficode (in b*w*_income)	RECODE of br_ficode_bh (income receipt code)
b*w*_finloc (in b*w*_indsamp)	final sample location identification
b*w*_ ivfho (in b*w*_hhresp, b*w*_indsamp and b*w*_indall)	Final household interview outcome
b*w*_maage (in b*w*_indresp)	Age of Mother: not in household
b*w*_paage (in b*w*_indresp)	Age of Father: not in household

b*w*_mafar (in b*w*_indresp)	Distance to where R's mother lives
b*w*_pafar (in b*w*_indresp)	Distance to where R's father lives
b*w*_malone (in b*w*_indresp)	R's mother lives alone
b*w*_palone (in b*w*_indresp)	R's father lives alone
b*w*_masee (in b*w*_indresp)	How often R sees mother
b*w*_pasee (in b*w*_indresp)	How often R's sees father
ba_mvever (in ba_indresp).	lived at address whole life
ba_mvmnth (in ba_indresp)	month moved to current address
ba_mvyr (in ba_indresp)	year moved to current address.

The variable b*w*_iviolw is not harmonisable but was inconsistently named. It has been renamed from b*w*_iviolw to b*w*_iviolw_bh to be consistent with other data files and to reflect the correct naming convention.

Users should note that the 6614_bhps_harmonised_user_guide has not been updated to reflect the harmonisation changes made in this release, however, the online documentation on the Understanding Society website is up-to-date.

Income Changes

In order to support income analysis and to allow the computation of income after housing costs a set of additional housing cost derived variables have been produced on the *w*_hhresp files as follows:

*w*_rent_dv	last rent payment, monthly
*w*_rentgrs_dv	monthly gross rent, including housing benefit
*w*_rentgrs_if	monthly gross rent, imputation flag
*w*_hbadjust_dv	income adjustment for housing benefit included in gross rent
*w*_xpmg_dv	monthly mortgage payment including imputations
*w*_xpmg_if	monthly mortgage payment imputation flag
*w*_xpmgint_dv	estimated interest in monthly mortgage payment
*w*_houscost1_dv	monthly housing cost including mortgage principal repayments
*w*_houscost2_dv	monthly housing cost excluding mortgage principal repayments
*w*_fihhmngrs1_dv	gross household income: month before interview, hb adj
*w*_fihhmnnet4_dv	total household net income – hh adj & individual / household deductions. *Note: this variable is only available in the Special Licence version.*

Income Top Coding Changes

The approach to top coding in the EUL file has been changed. Individual earnings and self-employment profit as well as investment income and rent have been top-coded at £100,000 per annum or its monthly equivalent for gross income. An equivalent less tax and national insurance for net incomes derived total personal income and household income are computed using the top-coded values. There is also a new set of flag variables indicating whether the computed sum variable include top coded amounts. The new flag variables are:

w_fimngrs_tc (in w_indresp)	total monthly personal gross income
w_fimnlabgrs_tc (in w_indresp)	fimnlabgrs_dv includes top coded component
w_fimnlabnet_tc (in w_indresp)	fimnlabnet_dv includes top coded component
w_fiyrinvinc_tc (in w_indresp)	fiyrinvinc_dv top coded
w_fibenothr_tc (in w_indresp)	fibenothr_dv includes top coded component
w_fihhmngrs_tc (in w_hhresp)	fihhmngrs_dv includes top coded component
w_fihhmnlabgrs_tc (in w_hhresp)	fihhmnlabgrs_dv includes top coded component
w_frmnth_tc (in w_income)	frmnth_dv top coded

While these variables are present in the SL version, they are only intended for use in the EUL.

Secure Access dataset (SN 6676) Changes

The Secure Access version (SN6676) previously contained all of the EUL (SN 6614) version files. Due to the secure environment in which the data must be accessed this policy has been changed so that it now contains all of the Special Licence (SN 6931) files instead in order that restricted variables can be more easily accessed.

The Secure Access version (SN 6676) now contains a full date of birth file for the harmonised BHPS waves. This was not previously available and day of birth was not available for any BHPS waves in any version of the data. In addition, the Secure Access version now contains partly harmonised grid reference files for the harmonised BHPS waves.

Childhood Language Changes

The variable f_kidlango_code, other childhood language spoken in the home when a child, in the file f_indresp in both the EUL and Special Licence versions previously did not contain any useful data. This variable has now been correctly populated, however, it is only included in the Special Licence version.

The variable b_kidlang_all, childhood language spoken in the home when a child, in b_indresp, has now been released in the Special Licence version only. This variable contains the full language list. The less restrictive variable b_kidlang, with a reduced language list, continues to be available in both the EUL and Special Licence versions.

Interviewer Corrections

The data file f_callrec for iemb includes interviewer numbers in the range 21680024 to 21680160 that were for 'Transport panel – screener'. These interviewers are now included in the xivdata file.

The variable *w*_int_num, interviewer number, in f_hhsamp, g_hhsamp, f_hhresp and g_hhresp has been renamed to *w*_intnum to be consistent across all files in all waves.

The variable *w*_intnum, interviewer number, in the data files *w*_hhsamp, *w*_hhsresp and *w*_callrec for Waves 6 and 7 has been corrected to include the fieldwork agency prefix allowing matches to be made against the interviewer file xivdata .

Renamed Variables

The variables f_hconda1 to f_hconda9, f_hconds1 to f_hconds9 , g_hconda1 to g_hconda9, and g_hconds1 to g_hconds9, all in *w*_indresp, have been renamed as f_hconda01 etc. to be consistent with Waves 1 to 5.

The variable c_scparoutint, rel with partner: engage in outside interests, in c_indresp, has been renamed c_scparoutintw3.

Variable Corrections

The variable *w*_depchl_dv has been revised so that 16-18 year-old parents of children aged 0 or older are treated as an independent child. This was previously interpreted as 16-18 year-old parents of children aged over 0. The correction also affects *w*_ndepchl_dv and *w*_buno_dv. These variables occur in *w*_indall and *w*_indresp. Additionally, the code for *w*_buno_dv has been re-based from the unedited *w*_adresp17 to the edited *w*_adresp15_dv. This addresses the issue that there were a small number of children who were in their own benefit unit.

The correction of invalid missing value codes [9, 99, 999] in the Wave 7 youth response data, means that Wave 7 derived SDQ variables are now corrected.

The variables g_vote3 / g_vote3_all, party would vote for tomorrow, and g_vote4 / g_vote4_all, which political party closest to, all in g_indresp, were previously incorrectly coded. Values 6 to 11 were affected. In addition, g_vote3_all and g_vote4_all now have value labels.

All non-IEMB sample members had incorrectly been treated as if they were absent household members affecting the variable f_hhresp_dv, household response status, in f_indresp. This has been corrected.

The variables:

fwintdv_dv_bh	first wave interviewed (inc. proxy), (BHPS)
lwintvd_dv_bh	last wave interviewed (inc. proxy), (BHPS)
fwenum_dv_bh	first wave enumerated (BHPS)
lwenum_dv_bh	last wave enumerated (BHPS)

all in xwaveid_bh, only had entries for Waves 1 to 7 and the distribution was incorrect. This has been corrected to cover all 18 BHPS waves and the distribution has consequently been corrected.

The variable f_yposclas8_code, other classes outside school, in f_youth, has been corrected so that it now contains meaningful data.

Various variables in *w*_youth for Waves 6 and 7 previously included values coded as "invalid answer" with the values 9 or 99 which made range checking difficult for users. These have now been recoded to -9. It should be noted that some variables still retain 9, 99 and 999 for out of range values and it is intended to correct these in the next release.

The variables *w*_paedqf, father's educational qualifications, and *w*_maedqf, mother's educational qualifications, in *w*_indresp for Waves 1, 2 and 6, previously contained two "don't know" categories. Those with the value 98 have been recoded to -1.

The variable b_sceverdrnk, ever had an alcoholic drink, in b_indresp, was mislabelled and this highlighted an inconsistency with the coding frame. It has now been corrected with the values 1 "yes" and 2 "no", in line with e_sceverdrnk.

New / Dropped Variables
Two relationship quality derived variables from the Revised Dyadic Adjustment Scale have been added to *w*_indresp: Dyadic satisfaction (*w*_scdassat_dv) and Dyadic cohesion (*w*_scdascoh_dv). They are included in all odd numbered waves with the exception of Wave 3 when only *w*_scdassat_dv is present.

The following coded variables from the iemb module have now been correctly included with the non-iemb data in f_indresp: f_hednto_code, f_kidlango_code, f_mreasono_code, f_otherlang_code, f_scrx_code and f_scux_code.

The variable *w*_hednto_code, other reason for not going to HE, has been added to *w*_indresp for Waves 4 to 7. The variable *w*_hednto_all has been dropped from all waves in which it was previously present.

The variables f_event1 to f_event4 and f_event1s to f_event4s, all in f_indresp, have been dropped as they did not contain any useful information.

The following variables have been dropped from the EUL version and are therefore now only available in the Special Licence (SN 6931) and Secure Access (SN 6676) versions:

lkdobm (in xwlsten)	last known date of birth - month
f_kidlango_code (in f_indresp)	other childhood language
*w*_lacbm (in a_adopt and f_adopt)	month of birth – step/adopted child
f_lchdobm (in f_natchild)	month of birth of non-resident biological child
f_otherlang_code (in f_indresp)	other languages spoken at home
bw_jbsoc90ly (in bw_indresp)	Occupation (SOC901990); job on 1.9.2007

bc_jceiscn (in bc_indresp)	International SOC (numeric): start current job
bc_jcesoc90 (in bc_indresp)	Occupation (SOC): start current job
br_mabrn (in br_indresp)	Country mother born
br_mamabrn (in br_indresp)	Country maternal grandmother born
br_mapabrn (in br_indresp)	Country maternal grandfather born
br_pabrn (in br_indresp)	Country father born
br_pamabrn (in br_indresp)	Country paternal grandmother born
br_papabrn (in br_indresp)	Country paternal grandfather born
bw_ypfsoc (in bw_youth)	office code: job (step)father does?
bw_ypsoc1 (in bw_youth)	main job
bw_ypsoc2 (in bw_youth)	second job
bw_ypsoc00 (in bw_youth)	office code (soc2000): job like when left
bw_ypsoc90 (in bw_youth)	office code: job like when left school?

Miscellaneous Changes

Various minor data inconsistencies have been corrected.

Various labelling issues have been corrected.