Understanding Society

The UK Household Longitudinal Study

Harmonised British Household Panel Survey (BHPS)

User Guide

Laura Fumagalli

Gundi Knies

Nick Buck

Institute for Social and Economic Research

University of Essex

Colchester

Essex

Version 1, November 2017

University of Essex

ISER INSTITUTE FOR SOCIAL & ECONOMIC RESEARCH

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

CONTENTS

Table of Tables 3
Table of Figures 3
Table of Appendixes 3
List of Abbreviations 4
1. Introduction 5
1.1. How to Navigate this User Guide 6
1.2. User guidance for different data usage routes 7
1.2.1. Using BHPS Wave 1-18 only (Usage 1) 9
1.2.2. Using Understanding Society Wave 1-7 only (Usage 2) 9
1.2.3. Using the BHPS sample in both BHPS and Understanding Society data (Usage 3) 9
1.2.4. Using BHPS Wave 1-18 and Understanding Society Wave 1-7 cross-sectionally (Usage 4) 10
2. Harmonsing the BHPS with *Understanding Society* 10
2.1. Principles 10
2.2. Generic Naming Conventions 10
2.3. Harmonisation of data files 11
2.3.1. Cross-wave files 11
2.3.2. Wave-specific files 11
2.3. Harmonisation of variables 15
2.3.1. General procedure 15
2.3.2. Procedures for specific types of variables 17
2.3.3. Value and variable labels 21
3. Documentation 21
4. Data release 22
5. Example Stata code 23
6. The *Understanding Society* harmonised BHPS project: Next Steps 25
7. Citations and Acknowledgements 25
7.1. Citation of the Data 26
7.2. Citation of the User Guide 26
7.3. Acknowledgments 26
8. References 27
9. Appendix 28

TABLE OF TABLES

Table 1: Signposting of User Guide Sections 6
Table 2: Cross-wave data files in the two studies and their treatment 12
Table 3: Wave-specific data files available in both studies and their treatment 13
Table 4: Wave-specific data files included only in the BHPS and their treatment 14
Table 5: Overview of identifiers and pointers in the harmonised BHPS 17
Table 6: Series wnvest and its treatment in the harmonised BHPS 18
Table 7: Income variables and their harmonisation 19
Table 8: Harmonised coding frame for the sample origin indicator (hhorig) 20
Table 9: Missing value codes 21
Table 10: BHPS data products available through the UKDS 22

TABLE OF FIGURES

Figure 1: Schematic overview of samples in the BHPS and *Understanding Society* .. 8
Figure 2: Flow chart of harmonisation of variables 16
Figure 3: Example Stata Code: Merging individual files from harmonised BHPS and UKHLS in long format 23
Figure 4: Example Stata code: Weighted longitudinal analysis of wellbeing in Britain (Usage 3) 24
Figure 5: Example Stata code: Weighted cross-sectional time series analysis of wellbeing in the United Kingdom (Usage 4) 25

TABLE OF APPENDIXES

Appendix 1: Variables included in both the harmonised BHPS and *Understanding Society*: individual enumeration files (indall) 28
Appendix 2: Variables included in both the harmonised BHPS and *Understanding Society*: household response files (hhresp) 28
Appendix 3: Variables included in both the harmonised BHPS and *Understanding Society*: adult response files (indresp) 28
Appendix 4: Variables included in both the harmonised BHPS and *Understanding Society*: individual youth response files (youth) 30

LIST OF ABBREVIATIONS

BHPS	British Household Panel Survey
ECHP	European Community Household Panel
EMB	Ethnic Minority Boost
ESRC	Economic and Social Research Council
EUL	End-User Licence
GOR	General Office Region
IP	Innovation Panel (UKHLS component for methodological research)
IEMB	Immigrant and Ethnic Minority Boost
ISER	Institute for Social and Economic Research
NatCen	National Centre for Social Research (now NatCen Social Research)
NI	Northern Ireland
ONS	Office for National Statistics
SL	Special Licence
UKDS	UK Data Service (also known as: UK Data Archive/ UKDA)
UKHLS	UK Household Longitudinal Study (official acronym for *Understanding Society*)

1. INTRODUCTION

Understanding Society: the UK Household Longitudinal Study (UKHLS) draws heavily on the success of its predecessor, the British Household Panel Survey (BHPS). The BHPS is a longitudinal social survey of households and individuals living in the UK. It started in 1991 with 5,000 households selected at random within Great Britain. In 1999, an additional 1,500 households were added in each of Scotland and Wales and in 2000, an additional 2,000 households were added in Northern Ireland. Data collection under the umbrella of the BHPS study title stopped in the year *Understanding Society* started (2009). Many design features, instruments, and questions from the BHPS live on in *Understanding Society* and data collection from eligible BHPS sample members continues as part of *Understanding Society*, offering opportunities to exploit data from the two studies jointly to create a long panel of data. There are, however, some structural differences between the two studies, such as separate study documentations, not identical data structures, naming conventions and identifiers. These differences present complexities for researchers wishing to use the two studies together.

The '*Understanding Society* harmonised BHPS' project (henceforward: harmonised BHPS) started in 2016, led by Laura Fumagalli and Nick Buck. The overall objective is to facilitate and encourage the combined use of *Understanding Society* and BHPS by overcoming actual and perceived barriers. The project aim is to produce a fully combined version of BHPS data from Wave 1 to 18 with *Understanding Society* data from Wave 1 onwards. The combined data will be updated and released jointly as part of *Understanding Society* in the future. With the *Understanding Society* Wave 1-7 data we present the result of the first phase of a more comprehensive project.

The project makes it possible to use the BHPS in conjuction with *Understanding Society* (and *vice versa*). This was also possible in the past, but it was an error-prone process requiring significant effort. We removed four main hurdles. First, we harmonised naming conventions, such that files pertaining to BHPS and *Understanding Society* can now be appended, merged and used together with very minimal recoding required. Second, we identified "spurious matches", i.e. cases of variables with the same name in the two studies, but different content, and we made sure users do not erroneously treat them as the same variables. Third, we identified cases where the same information was present in the two studies under different forms (e.g., under different variable names). In this case, we renamed, recoded and combined existing BHPS variables to match the *Understanding Society* equivalent. Fourth, we harmonised the cross-wave information in xwavedat.

As a result of this effort around a fifth of the variables in the individual response data file (**indresp**) of the harmonised BHPS are harmonised, i.e., they have a non-spurious match in *Understanding Society*. The respective shares of harmonised variables in other files are 23 per cent in the young person response data file (**youth**), around 27 per cent in the case of the individual enumeration data file (**indall**) and 31 per cent in the case of the household response data file (**hhresp**). Note that none of these matches are spurious, and thus variables with the same name now measure the same construct. Finally, users can now find all the cross-wave information for all BHPS and *Understanding Society* sample members in a single file (**xwavedat**).

The documentation has also been harmonised. The *Understanding Society* online data documentation now covers the harmonised BHPS, as well as the *Understanding Society* data. The stand-alone BHPS questionnaires have been included, but these documents have not been updated to reflect the new variable names in the harmonised BHPS. Users will need to consult the online documentation for information about the treatment of specific variables.

Data harmonisation is a challenging process. Although significant efforts have been made to achieve the highest data quality, it can still be possible that variables with the same name, within or across studies, are derived through slightly different question wording or code frame. For example, the code frame for the employment status variable jbstat changed over time in both studies. Analysts should always check the questionnaires, frequencies and labels of any data they want to analyse.

This project wouldn't have been possible without the support of data users. We are grateful to a range of beta users who tested and fed back on an earlier version of this file. Future phases will expand the range of variables harmonised. We would welcome suggestions and code from users for areas that could be harmonised further.

1.1. How to Navigate this User Guide

The *Understanding Society* harmonised BHPS User Guide accompanies the first edition of the *Understanding Society*-harmonised BHPS, released as part of the *Understanding Society* data series. It supplements the *Understanding Society* Wave 1-7 User Guide (Knies 2017) and the stand-alone BHPS Wave 1-18 User Guide Volume A (Taylor 2010). The focus is on the harmonisation process for the BHPS data. We do not cover any structural differences in scope, fieldwork practices, questionnaire design and content of the two studies. Analysists should refer to the respective Study's stand-alone user guide for this information and evaluate how the differences may affect the results of their analysis. Table 1 lists specific sections of the stand-alone user guides.

Table 1: Signposting of User Guide Sections

	BHPS Volume A documentation	UKHLS User Guide	Harmonised BHPS
Introduction	A2-2 – A2-3	1	1
Data structure	A2-11	3.1	2.3
Unique identifiers	A3-9	3.2.4	2.3.2.1
Variable naming conventions	A3-13	3.2.2	2.2 and 2.3.2
Missing values	A3-15	3.2.3	2.3.3.1
Sampling	A4-1 – A4-5	2.2	See BHPS and UKHLS guides.
Data collection methods	A4-6 – A4-13, A4-20 – A4-23	2.3	See BHPS and UKHLS guides
Weighting	A5-1 – A5-12	3.9	2.3.2.5
Imputation	A5-22 – A5-24	3.10	See BHPS and UKHLS guides.
Accessing the data	A6-1	4	2

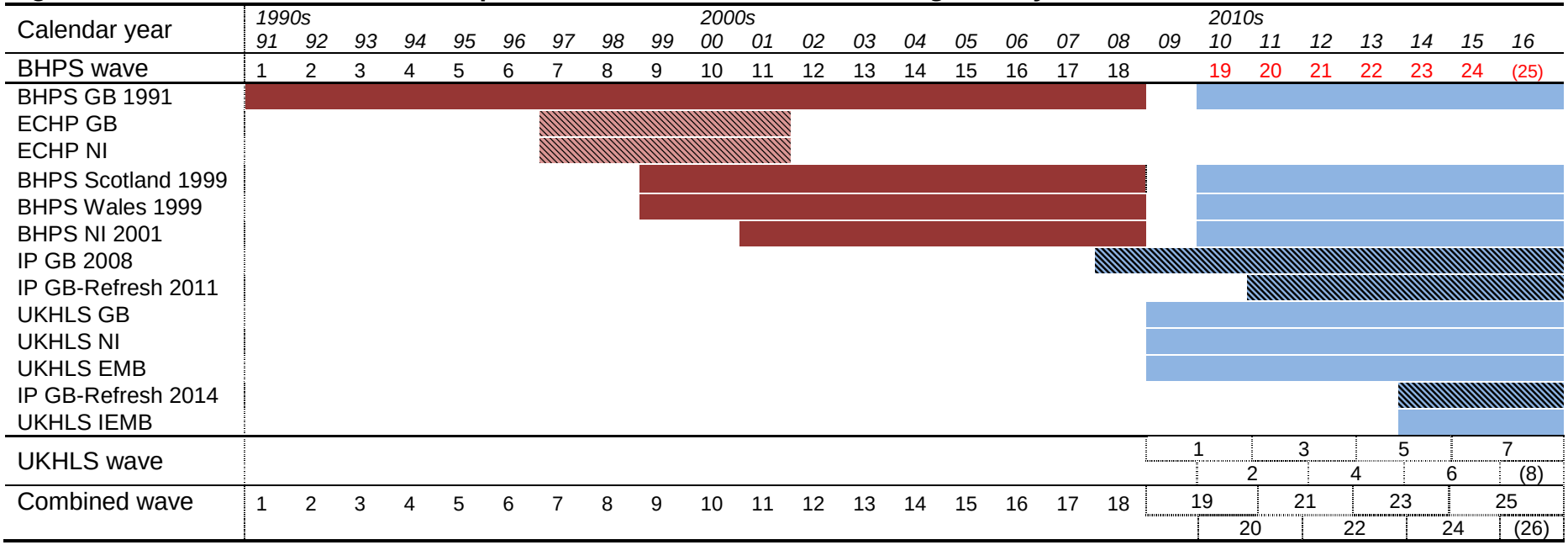
This guide is organised as follows. Section 1.2 provides a short overview of different usage of the data sets. Section 2 presents the harmonised BHPS. It outlines the principles that governed the production of the harmonised BHPS (Section 2.1) and how the harmonisation was implemented for data files (Section 2.3) and variables (Section 2.3). Section 3 describes the online documentation of the harmonised BHPS data, and Section 4 provides references on how to access the data. Section 5 gives analysts worked examples (in the Stata programming language) of how to exploit the harmonised BHPS and *Understanding Society* data together. Section 6 focuses on next steps in the '*Understanding Society* harmonised BHPS' project and Section 7 provides users with information about how to cite data and documentation.

1.2. USER GUIDANCE FOR DIFFERENT DATA USAGE ROUTES

Together, BHPS and *Understanding Society* provide rich data for longitudinal samples of individuals living in the UK going back to 1991. Figure 1 provides a schematic overview of the different samples and when they were introduced in the studies. BHPS samples are shaded maroon and *Understanding Society* samples are shaded light blue. Samples shaded with diagonal lines are not part of the UKHLS harmonised BHPS; they are included for completeness only. The top row lists the calendar years, followed by the BHPS wave count (continues in red once part of *Understanding Society*). The next rows list the BHPS and *Understanding Society* samples in chronological order. The last row indicates the *Understanding Society* wave count. Note that each wave in the *Understanding Society* stretches over two calendar years; data for the BHPS samples participating in *Understanding Society* is collected in the first year of each wave, starting from *Understanding Society* Wave 2.

There are four different data usage routes for the two studies: (1) using only BHPS Wave 1-18, (2) using only *Understanding Society* Wave 1-7, (3) combining the harmonised BHPS and *Understanding Society* data, but only using longitudinally the continuing BHPS sample, and (4) combining the harmonised BHPS and *Understanding Society* data cross-sectionally, without restricting the analysis to the continuing BHPS sample. For example, usage (3) is for researchers who want to carry out a longitudinal analysis at the individual level for a time span covering BHPS and *Understanding Society* years, while usage (4) is for researchers interested in studying an event (e.g., national wellbeing) over a long period of time (e.g., over 25 years).

Figure 1: Schematic overview of samples in the BHPS and *Understanding Society*



Notes: Fieldwork for each wave typically continues into the following calendar year. (#) Data for BHPS Wave 25, combined Wave 26, are being collected as part of UKHLS Wave 8 which will be released in November 2018.

1.2.1. USING BHPS WAVE 1-18 ONLY (USAGE 1)

Users who want to analyse BHPS Wave 1-18 only are advised to use the stand-alone BHPS and associated documentation. The BHPS documentation is available on the ISER webpage: http://www.iser.essex.ac.uk/bhps.

This online documentation covers:

- Information on how to acquire the data (http://www.iser.essex.ac.uk/bhps/acquiring-the-data)
- Information on the sample (http://www.iser.essex.ac.uk/bhps/about/sample)
- Information on content of the questionnaire (http://www.iser.essex.ac.uk/bhps/about/questionnaire-content)
- Frequently Asked Questions (http://www.iser.essex.ac.uk/bhps/faqs)

Volume A of the stand-alone BHPS User Guide (Taylor 2010) is a pdf file (http://www.iser.essex.ac.uk/bhps/documentation/vola/vola.html) which includes descriptive information about the survey and data. Appendix 2 provides descriptions of derived variables and Appendix 3 provides the variable labels for long code frames such as for the National Statistics Socio-economic classification (NSSEC). A slightly amended PDF copy of Taylor (2010) is also provided with the *Understanding Society* Wave 1-7 and harmonised BHPS Wave 1-18 data release.

Volume B of the BHPS User Guide is online-only and helps users identify which variables are available, where, for which waves, and so on. Volume B can be searched either by subject thesaurus, index terms, record types or waves. The URL is: https://www.iser.essex.ac.uk/bhps/documentation/volb/index.html.

We provide worked examples on using the BHPS in our online Moodle course (https://www.understandingsociety.ac.uk/documentation/training/online/) and additional user support is provided through our *Understanding Society* Online User Support Forum at https://www.understandingsociety.ac.uk/support/projects/support.

The BHPS Quality Profile (Lynn, Buck et al. 2006) provides a comprehensive overview of all aspects of the design and implementation of the BHPS and on sources of survey error.

In producing the *Understanding Society* harmonised BHPS data a small number of issues have been identified that needed correcting in the stand-alone BHPS. These issues have been addressed in the harmonised BHPS and will be addressed in a new release of the stand-alone BHPS and documentation in early 2018.

1.2.2. USING UNDERSTANDING SOCIETY WAVE 1-7 ONLY (USAGE 2)

The *Understanding Society* User Guide (Knies 2017) provides all necessary information for users who want to use only *Understanding Society* data. See Table 1 for key sections of the guide.

1.2.3. USING THE BHPS SAMPLE IN BOTH BHPS AND UNDERSTANDING SOCIETY DATA (USAGE 3)

Users who want to analyse information from up to 24 waves of harmonised BHPS and *Understanding Society* data for the continuing BHPS sample are advised to read this guide as an introduction. More detailed information about specific aspects of the

two studies is available in the stand-alone user guides and in the online documentation.

The appropriate weights for this usage route are the longitudinal weights provided in *Understanding Society* and described in Section 3.9 of Knies (2017). Figure 4, in this user guide, provides example code (in the Stata programming language).

1.2.4. USING BHPS WAVE 1-18 AND UNDERSTANDING SOCIETY WAVE 1-7 CROSS-SECTIONALLY (USAGE 4)

Users who want to use BHPS Waves 1-18 and *Understanding Society* Wave 1-7 cross-sectionally are advised to read this guide, as well as the stand-alone user guides and the on-line documentation.

The appropriate weights for this usage route are the cross-sectional weights provided in the two studies. For the BHPS, these are described in Section A5-1 – A5-12 of Taylor (2010). For *Understanding Society*, these are described in Section 3.9 of Knies (2017). Figure 5, in this user guide, provides example code (in the Stata programming language).

2. HARMONSING THE BHPS WITH *UNDERSTANDING SOCIETY*

2.1. PRINCIPLES

Four general principles guide the harmonised BHPS project. First, the aim of the project is revising the BHPS data, so they can be analysed more easily in conjunction with *Understanding Society* data. Therefore, most changes were made to the BHPS, and only minor changes were made to *Understanding Society*. Second, the harmonisation criteria are quite strict (see Section 3.4). This implies the quality of the harmonisation process across variables is roughly consistent, and data recoding and variable renaming is minimal. Third, the BHPS is released in full as part of *Understanding Society* Wave 7. In orther words, content that cannot be or has not yet been harmonised is released alongside content that is equivalent to UKHLS. Fourth, the project is dynamic. The number of harmonised variables is likely to increase over time. This is because new matches can be found or made possible. Value added datasets (e.g., geographical identifiers) will be provided with future releases.

2.2. GENERIC NAMING CONVENTIONS

The BHPS and *Understanding Society* have many similarities. There is one set of files for each wave, there are stem names for files and variables and the wave is identified by a wave prefix. The content of files with the same name across the two studies is principally the same. For example, **indresp** files always contain information from adult interviews; **hhresp** files always contain information from household interviews and so on. However, there are also differences, such as different wave prefixes and variable naming conventions.

The harmonised BHPS uses the *Understanding Society* conventions. In particular:

Prefixes. Like *Understanding Society*, the harmonised BHPS uses a wave prefix followed by an underscore for variable and file names. To distinguish harmonised BHPS files and variables from *Understanding Society* ones, the harmonised BHPS wave prefix is preceded by the letter "b". For example, the harmonised BHPS Wave 1 uses the prefix ba_, and the harmonised BHPS Wave 18 uses the prefix br_.

Identifiers. The harmonised BHPS includes the cross-wave person identifiers of both studies (**pidp** and **pid)** and both identifiers are applicable to all sample members. Likewise, it includes the within-wave household identifiers in the BHPS and *Understanding Society* formats (***bw***_**hid** and ***bw***_**hidp**).

Suffixes. In line with the U*nderstanding Society* rules, the suffix "_dv" indicates derived variables, the suffix _if indicates imputation flags, and the suffix "_cc" indicates condensed versions of socio-economic classifications.

2.3. HARMONISATION OF DATA FILES

The BHPS and *Understanding Society* differ in the number and type of data files they contain. Both studies contain cross-waves and wave-specific files. In the following we summarize how files have been treated in the harmonised BHPS.

2.3.1. CROSS-WAVE FILES

Cross-wave files (**xwavedat, xwaveid, xivdata, xwlsten)** store information about all sample members ever enumerated in the respective study. The data files xwavedat, xwaveid, xivdata exist in both studies. The BHPS and the *Understanding Society* xwavedat files have been combined into one. The files **xwaveid** and **xivdata** have not been harmonised yet. **xwlsten** only exists in BHPS and thus did not require harmonisation.

Table 2 below gives details on how these files have been treated in the harmonised BHPS.

2.3.2. WAVE-SPECIFIC FILES

Both the BHPS and *Understanding Society* contain a number of wave-specific files. For the purpose of harmonisation, they can be divided in two groups. The first group is composed of data files which exist in both studies. Table 3 shows how these data files have been treated. Most data files (i.e., **indall**, **egoalt**, **indsamp**, **hhresp**, **indresp**, **youth**, and **income**) have been fully harmonised. The data file storing information about all issued households (**hhsamp**) has not been harmonised yet. The second group is composed of data files which exist only in the stand-alone BHPS (see Table 4). This group includes the **child** data file, which exists in both studies, but stores different information. No attempt has been made to harmonise data files with different names, or to harmonise the **child** data file. Note that there are also a number of files which only exist in *Understanding Society*. They are of no concern for the harmonisation project.

Table 2: Cross-wave data files in the two studies and their treatment

	BHPS		UKHLS	Description	Treatment received
	Stand-alone	**Harmonised**			
Harmonised	xwavedat	xwavedat		Contains stable characteristics of all individuals ever enumerated	Both studies' files have been fully combined in one data file containing cases from all sample members ever enumerated in either study. The variable **xwdat_dv** can be used to identify sample members enumerated in one or both the surveys.
Not yet harmonised	xwaveid	xwaveid_bh	xwaveid	Contains information for enumerated matching individuals between waves	The data file has largely different content in BHPS and *Understanding Society*. The data files have not (yet) been harmonised. The BHPS data file **xwaveid** was renamed to **xwaveid_bh** and all variables in the data file received the prefix bw_ and the suffix _bh
	xivdata	xivdata_bh	xivdata	Contains information about all interviewers ever enumerated	The data file has largely different content in BHPS and *Understanding Society*. The data files have not (yet) been harmonised. The BHPS data file **xivdata** was renamed to **xwaveid_bh** and all variables in the data file received the prefix bw_ and the suffix _bh
BHPS only	xwlsten	xwlsten	-	Contains information on the latest known sample status of all individuals ever enumerated	There is no equivalent to the BHPS data file **xwlsten** in *Understanding Society*. The data file was not renamed **xwlsten** and all variables in the dataset received the suffix _bh

Table 3: Wave-specific data files available in both studies and their treatment

BHPS		**UKHLS**	**Description**	**Treatment received**
Stand-alone	**Harmonised**			
*w*indsamp	*bw*_indsamp	*w*_indsamp	Individual-level data for issued households	***w*indsamp** was renamed to ***bw*_indsamp** and all variables in the data file received the prefix *bw*_.
*w*hhsamp	*bw*_hhsamp_bh	*w*_hhsamp	Sample and Household level data for issued households	***w*hhsamp** was renamed to ***bw*_hhsamp_bh** and all variables in the data file received the prefix *bw*_ and the suffix _bh.
*w*indall	*bw*_indall	*w*_indall	Household grid data for all enumerated persons in household, including children and non-respondents	***w*indall** was renamed to ***bw*_ indall** and all variables in the data file received the prefix *bw*_.
*w*egoalt	*bw*_egoalt	*w*_egoalt	Kin and other relationships between enumerated pairs of individuals in the household	***w*egoalt** was renamed to ***bw*_egoalt** and all variables in the data file received the prefix *bw*_.
*w*hhresp	*bw*_hhresp	*w*_hhresp	Substantive data from responding households	***w*hhresp** was renamed to **bw*_ hhresp** and all variables in the data file received the prefix *bw*_.
*w*indresp	*bw*_indresp	*w*_indresp	Substantive data for responding adults (16+) including proxies and telephone interviews from individual questionnaires including self-completion	***w*indresp** was renamed to ***bw*_indresp** and all variables in the data file received the prefix *bw*_.
*w*income	*bw*_income	*w*_income	Income and payment information for respondents	***w*income** was renamed to ***bw*_income** and all variables in the data file received the prefix *bw*_.
*w*youth	*bw_youth*	*w*_youth	Substantive data from respondents to the youth questionnaire (age 11-15 for BHPS and 10-15 for UKHLS)	***w*youth** was renamed to ***bw*_youth** and all variables in the data file received the prefix *bw*_.

Table 4: Wave-specific data files included only in the BHPS and their treatment

BHPS		**Description**	**Treatment received**
Stand-alone	**harmonised**		
*w*jobhstd	*bw*_jobhstd_bh	Information from the employment history for responding adults, based on dependent interviewing	***w*jobhstd** was renamed to ***bw*_jobhstd_bh** and all variables in the data file received the prefix *bw_* and the suffix _bh.
*w*jobhist	*bw*_jobhist_bh	Employment history for responding adults	***w*jobhist** was renamed to ***bw*_jobhist_bh** and all variables in the data file received the prefix *bw_* and the suffix _bh.
*w*lifejob	*bw*_lifejob_bh	Information about jobs; held in employment spells for responding adults	***w*lifejob** was renamed to ***bw*_lifejob_bh** and all variables in the data file received the prefix *bw_* and the suffix _bh.
*w*lifemst	*bw*_lifemst_bh	Information about employment status spells; for responding adults	***w*lifemst** was renamed to ***bw*_lifemst_bh** and all variables in the data file received the prefix *bw_* and the suffix _bh.
*w*child	*bw*_child_bh	Information about respondent's children	***w*child** was renamed to ***bw*_child_bh** and all variables in the data file received the prefix *bw_* and the suffix _bh.
*w*cohabit	*bw*_cohabit_bh	Data about each cohabitation spell outside legal marriage; for responding adults	***w*cohabit** was renamed to ***bw*_cohabit_bh** and all variables in the data file received the prefix *bw_* and the suffix _bh.
*w*marriag	*bw*_marriag_bh	Information about previous marriages; for responding adults	***w*marriage** was renamed to ***bw*_marriage_bh** and all variables in the data file received the prefix *bw_* and the suffix _bh.
*w*childad	*bw*_childad_bh	Information about adopted and/or step-children for responding adults	***w*childdad** was renamed to ***bw*_childdad_bh** and all variables in the data file received the prefix *bw_* and the suffix _bh.
*w*childnt	*bw*_childnt_bh	Information about natural children for responding adults	***w*childnt** was renamed to ***bw*_childnt_bh** and all variables in the data file received the prefix *bw_* and the suffix _bh.

2.3. HARMONISATION OF VARIABLES

2.3.1. GENERAL PROCEDURE

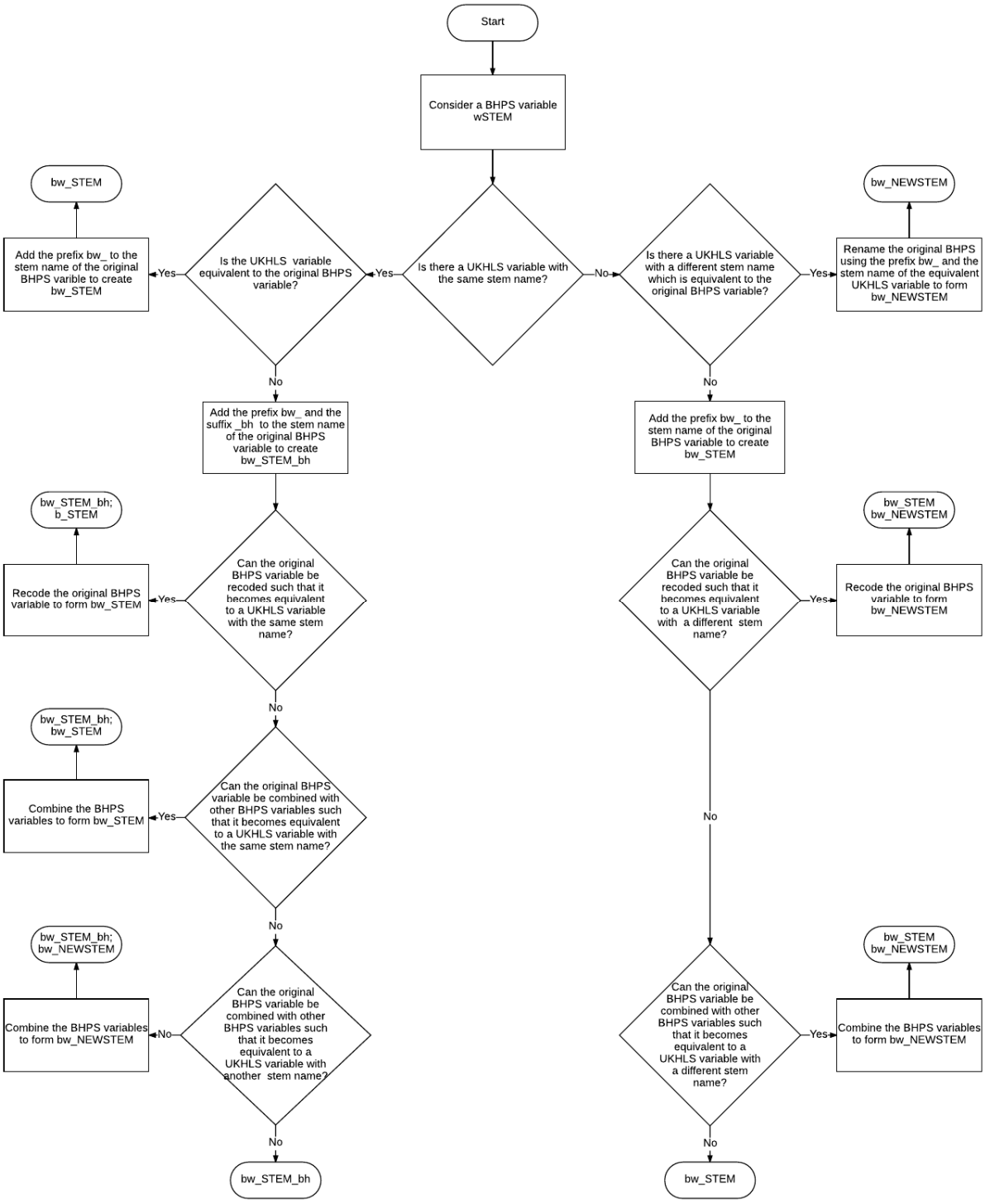
The aim of the project was to identify *Understanding Society* variables equivalent to one or more BHPS variables; this was done in the following steps:

1. We identified one (or more) *Understanding Society* candidate match variable(s) for a BHPS variable. This involved an automated search to match variables with the same or similar variable stem name or label in each study. Further matches were identified by experts (*Understanding Society* team members and beta users).
2. We determined whether the BHPS variable and its candidate *Understanding Society* match were sufficiently similar to be treated as equivalent. We checked for equivalence in question wording, question routing (i.e., the subset of respondents who are asked the question), response options and mode of administration.
3. We determined whether a modification could make the BHPS variable and its candidate *Understanding Society* match similar enough to treat as equivalent
4. We determined whether a set of BHPS variables could be combined to create a new variable equivalent to the candidate *Understanding Society* match.

Figure 2 illustrates the outcomes of the process above for the generic BHPS variable *w*stem.

- If stem name and content were the same in both studies or if no variable existed that conveyed the equivalent information in the UKHLS, *w*stem was renamed *bw*_stem.
- If equivalent information existed under a different variable stem name in the *Understanding Society*, *w*stem received the stem name of that variable and the wave prefix, i.e. *bw*_newstem.
- If the information conveyed in *w*stem and *w*_stem was different and no *Understanding Society* variable could be found or created to match the content of *w*stem, *w*stem was renamed to *bw*_stem_bh.
- If a BHPS variable could be modified to match the content of the *Understanding Society* variable *w*_stem, a clone of the BHPS variable was created and modified (*bw*_newstem), and the original renamed as *bw*_stem_bh.
- If a BHPS variable could be combined with other BHPS variables to match the content of the UKHLS variable *w*_newstem, a new variable *bw*_newstem was created by combining those variables.

Figure 2: Flow chart of harmonisation of variables



2.3.2. PROCEDURES FOR SPECIFIC TYPES OF VARIABLES

2.3.2.1. Identifiers and pointers to other household members

All sample members of the harmonised BHPS have a person and household identifier (**pidp**, ***bw*_hidp**). No changes were made to the BHPS within-wave person number ***w*pno**. The identifiers **pidp** and ***bw*_hidp** work across harmonised BHPS and *Understanding Society* files, and the ***bw*_hid *bw*_pno** can be used to link to other BHPS data currently not harmonised.

Both BHPS and *Understanding Society* provide a range of pointers to significant others, that is variables indicating the unique cross-wave person identifier or within-wave person numbers of significant others in the household. Pointers are based on information about within-household relationships. Although protocols for collecting this information varied across the two studies, there is a broad equivalence of constructs. Therefore, pointers have been considered in princple equivalent.

Table 5 shows identifiers and pointers in the stand-alone BHPS and how they have been named in the harmonised BHPS. Note that pointers containing the cross-wave person identifier have received the suffix _bh because pointers in BHPS report the **pid,** whilst pointers in *Understanding Society* report the **pidp**.

Table 5: Overview of identifiers and pointers in the harmonised BHPS

Stand-alone BHPS	Harmonised BHPS	Description
pid	pid	Unique person identifier (BHPS cohort)
	pidp	Unique person identifier (BHPS & UKHLS cohort)
*w*hid	*bw*_hid	Within-wave household identifier (BHPS cohort)
	*bw*_hidp	Within-wave household identifier (BHPS & UKHLS cohort)
*w*pno	*bw*_pno	Person number within current household
*w*hgspn	*bw*_ppno	PNO of spouse/partner (edited)
*w*hgra	*bw*_hgra	PNO of adult responsible for child under age 18 (HHGRID).
*w*rapid	*bw*_rapid_bh	ID of adult responsible for child under age 18 (HHGRID)
*w*hgmno	*bw*_hgbiom	PNO of natural mother (HHGRID)
*w*mpid	*bw*_mnpid_bh	ID of natural mother (edited)
*w*mnspno	bw_mnspno	PNO of nat./step/adopt. mother (edited)
*w*mnspid	*bw*_mnspid_bh	ID of nat./step/adopt. mother (edited)
*w*hgfno	*bw*_hgbiof	PNO of natural father (HHGRID)
*w*fpid	*bw*_fnpid_bh	ID of natural father (edited)
*w*fnspno	*bw*_fnspno	PNO of nat./step/adopt. father (edited)
*w*fnspid	*bw*_fnspid_bh	ID of natural/step/adoptive father (edited)
*w*cosapid	*bw*_cosapid_bh	ID of adult asked for consent to link child's administrative records (HHGRID)
*w*cosano	*bw*_cosano	PNO of adult asked for consent to link child's administrative records (HHGRID)

2.3.2.2. Batteries of questions

Batteries of questions are series of related questions with a common stem name. Some of these batteries of questions measure well-established constructs such as the General Health Questionnaire (GHQ).

Some batteries of questions used in the BHPS were entirely carried over into *Understanding Society*; others were carried over only partially. In cases where the batteries were entirely carried over into *Understanding Society* the variable-renaming followed the procedure explained in Figure 2. For example, the components of the battery measuring the GHQ were renamed from ***w*ghqa**-***w*ghql** to ***bw*_scghqa**-***bw*_scghql** in the harmonised BHPS, to match the *Understanding Society* equivalent variable names ***w*_scghqa**- ***w*_scghql**. In cases where the batteries of questions were only partially carried over into *Understanding Society*, the original battery of questions was kept with the original name, and the questions carried into *Understanding Society* were duplicated and given the *Understanding Society* stem name and the harmonised BHPS prefix.

An example for this case is shown in Table 6.

Table 6: Series wnvest and its treatment in the harmonised BHPS

stand-alone BHPS	harmonised BHPS	UKHLS
*w*nvesta	*bw*_nvesta	
*w*nvestb	*bw*_nvestb	*w*_svacts5
	*bw*_svacts5	
*w*nvestc	*bw*_nvestc	
*w*nvestd	*bw*_nvestd	
*w*nveste	*bw*_nveste	
*w*nvestf	*bw*_nvestf	
*w*nvestg	*bw*_nvestg	
*w*nvesth	*bw*_nvesth	*w*_svacts1
	*bw*_svacts1	
*w*nvesti	*bw*_nvesti	*w*_svacts2
	*bw*_svacts2	
*w*nvestj	*bw*_nvestj	
*w*nvestk	*bw*_nvestk	

2.3.2.3. Derived variables

Both the BHPS and *Understanding Society* include a range of derived variables. Derived variables are constructed using algorithms combining one or more pieces of information collected in the study (see Taylor, 2010, Appendix 2, and Knies 2017, Section 2.7). We harmonised derived variables already existing in the stand-alone BHPS and *Understanding Society* following the procedure described in Section 2.1 above. We did not harmonise the algorithms. The harmonised BHPS variables were given the *Understanding Society* stem names and prefixes.

2.3.2.4. Income variables

Both BHPS and *Understanding Society* collect information on earnings, investment income, and other sources of income such as benefits, pensions, and educational grants, alimonies, rent from boarders and lodgers, and payments from family members. Information on monthly earnings/income is comparable in the two studies. The top panel of Table 7 shows the harmonised derived income variables and associated imputation flags available in the harmonised BHPS. Again, harmonisation is at the variable- and construct- level. The algorithms used to construct these derived income variables and the imputation methods have not been harmonised. Full information on annual income is only available in BHPS. Therefore, information on annual income is largely not harmonised. A list of not harmonised income variables is provided at the bottom of Table 7.

Table 7: Income variables and their harmonisation

BHPS		**Description**
stand-alone	**harmonised**	
*w*payn	*bw*_payn_dv	Net pay per month in current job: last payment
*w*paynu	*bw*_paynu_dv	Usual net pay per month: current job
*w*paynui	*bw*_paynu_if	Usual net pay per month current job: Imputation flag
*w*payg	*bw*_payg_dv	Gross pay per month in current job: last payment
*w*paygu	*bw*_paygu_dv	Usual gross pay per month: current job
*w*paygui	*bw*_paygu_if	Usual gross pay per month current job: imputation flag
*w*fimnl	*bw*_fimnlabgrs_dv	Total monthly labour income
*w*fimn	*bw*_fimngrs_dv	Total personal income: gross
*w*j2pay	bw_j2pay_dv	Gross earnings from second jobs last month
*w*j2payi	*bw*_j2pay_if	Gross earnings from second jobs last month: imputation flag
wfrval	*bw_frval*	Benefits/Unearned Income: Last amount received
*w*j2payi	*bw*_j2pay_if	Gross earnings from second jobs last month: imputation flag
BHPS income variables not currently harmonised: fimnli frvali fimnnl fimnnli fimnsel fimnt fimnthi fimnti fiyeari fiyr fiyrl fiyrli fiyrnl fiyrnli jspayg jspaygi jsprof jsprofi paygli paygly paygti paygty paynli paynly paynti paynty sppayg tlfiyr tlfiyrl fihhmb fihhmbi fihhmp fihhmpi fihhmt fihhmti fihhyb fihhybi fihhyp fihhypi fihhyti fimnb fimnbi fimnp fimnpi fiyrb fiyrbi fiyrp fiyrpi fiyrt fiyrti. Note that the BHPS ***wj2pay*** has been also copied as ***bw_wj2pay_bh***.		

2.3.2.5. Sample design variables

The BHPS and *Understanding Society* are based on nine samples (see Figure 1). The description of these samples is provided in the stand-alone BHPS and *Understanding Society* user guides.

The variables describing the sampling unit and stratification are supplied in the **xwavedat** data file and have been harmonised. The primary sampling unit variable **psu** indicates the primary sampling unit (PSU) to which the sample member belongs. Its value does not change over time. The stratification variable **strata** indicates the sampling stratum from which the sample member was selected. Its value does not change over time. For new sample entrants, the psu and strata values are defined only from the wave at which they enter the sample. Note that some harmonised BHPS files include the variables **psu_bh** and **strata_bh**: these are the original primary sampling unit and stratification variables in the stand-alone BHPS.

The variable **hhorig** on data file **xwavedat** can be used to inspect the sample origin for individuals. The code frame for this variable has been adjusted in both studies to reflect samples across both studies (see Table 8). Note that some harmonised BHPS files contain the variable ***bw*_hhorig_bh**: this is the original sample origin variable in the stand-alone BHPS.

Please note that many of the samples of BHPS and *Understanding Society* are not designed as stand-alone samples and selecting them using variables such as **hhorig**, can lead to wrong results**.** We recommend that users choose the appropriate population weights provided in the studies: using weights automatically selects respondents from the samples of interest.

Table 8: Harmonised coding frame for the sample origin indicator (hhorig)

BHPS		**UKHLS**		**New harmonised coding frame**	
1	original sample	1	ukhls gb 2009-10	1	ukhls gb 2009-10
2	echp-scpr	2	ukhls ni 2009-10	2	ukhls ni 2009-10
3	echp-ons	3	bhps gb 1991	3	bhps gb 1991
4	echp-ni	4	bhps sco1999	4	bhps sco 1999
5	wales new sample	5	bhps wal1999	5	bhps wal 1999
6	scotland new sample	6	bhps ni 2001	6	bhps ni 2001
7	n.i. new sample	7	ukhls emboost 2009-10	7	ukhls emboost 2009-10
		8	ukhls iemb 2014-15	8	ukhls iemb 2015
		11	ip original sample	11	ip original sample
		12	ip4 refresher sample	12	ip4 refresher sample
		13	ip7 refresher sample 2014	13	ip7 refresher sample 2014
				14	echp-scpr
				15	echp-ons
				16	echp-ni

Notes: Greyed out samples do not have any cases in the harmonised BHPS and *Understanding Society* files.

2.3.2.6. Survey Weights

Both BHPS and *Understanding Society* provide survey weights and use specific naming conventions to help users identify the correct weights for their analysis. These study-specific naming conventions have been retained in the harmonised BHPS. In other words, weights in the harmonised BHPS have kept the same stem name, but have received the wave-prefix *bw_*.

2.3.3. VALUE AND VARIABLE LABELS

Value and variable labels have not been fully harmonised yet. While the content of the labels is the same, the exact wording may differ.

2.3.3.1. Missing values

The general missing value definition in the BHPS and *Understanding Society* is identical. In creating the harmonised BHPS files we retained the original BHPS values. Missing value definitions may vary across studies. Table 9 lists the missing value codes used in the harmonised BHPS.

Table 9: Missing value codes

Value	Description
-21	No data from the UKHLS
-20	No data from the BHPS Wave 1-18
-11	Only available for the IEMBS
-10	Not available for the IEMBS
-9	Missing by error or implausible
-8	Not applicable to the person or because of Routing
-7	Proxy respondent. The question was not asked of proxy respondents or derived variable cannot be computed for proxy respondents.
-2	Refused
-1	Don't know

3. DOCUMENTATION

The harmonised BHPS is documented as part of the main *Understanding Society* study website www.understandingsociety.ac.uk/documentation/. PDFs of the questionnaires used in the BHPS are provided in the questionnaire tab. Note that the question names in the questionnaires have not been updated to reflect the harmonised BHPS variable names. We provide a look-up file for BHPS variables that have been harmonised, allowing users to learn about the origin of harmonised variables, see https://www.understandingsociety.ac.uk/d/349/HarmoDocuAll.xls. The information is also provided in the variable level view of the online documentation.

Other fieldwork materials used in the BHPS are provided in the "Other fieldwork materials" tab.

4. DATA RELEASE

The data are released through the UK Data Service (UKDS) in SPSS, Stata and CSV formats. While documentation is released through the UKDS, we encourage users to consult the *Understanding Society* webpage. The online documentation will develop over time.

The *Understanding Society* harmonised BHPS is released as part of the main *Understanding Society*. The data are released according to the conditions of the regular UKDS End User Licence (EUL): http://ukdataservice.ac.uk/get-data/how-to-access/conditions.aspx#/tab-end-user-licence. These data are listed as SN 6614 - *Understanding Society*: Waves 1-7, 2009-2016 and harmonised British Household Panel Survey: Waves 1-18.

Some additional end-user licence data products available for the BHPS cases are listed in Table 10. These data can be linked using the BHPS unique identifiers (**pid** and ***w*hid**). For files at the household level, (e.g., the Derived Current and Annual Net Household Income Variables, SN 3909) the original format of the **hid** is used (e.g., **rhid** rather than **br**_**hid**) and these will need renaming in order to match the files.

All these data can be accessed directly by replacing ## by the Study number in the following URL: https://discover.*uk*dataservice.ac.*uk*/catalogue/?sn=##

Table 10: BHPS data products available through the UKDS

Study	Study no	Study Title
BHPS	5151	British Household Panel Survey: Waves 1-18, 1991-2009
BHPS	3909	British Household Panel Survey Derived Current and Annual Net Household Income Variables, Waves 1-18, 1991-2009
BHPS	5629	British Household Panel Survey Consolidated Marital, Cohabitation and Fertility Histories, 1991-2009
BHPS	7821	British Household Panel Survey: Programs for Generating Consistent Work-Life Histories: Waves 1-18, 1991-2009
BHPS	3954	British Household Panel Survey Combined Work-Life History Data, 1990-2005
BHPS	5354	Human Capital and Social Position in Britain: Creating a Measure of Wage-Earning Potential from BHPS Data, 1991-2004
BHPS	5356	British Household Panel Survey Calibrated Time Use Data, 1994-2004
UKHLS	6614	*Understanding Society*: Waves 1-7, 2009-2016. Includes harmonised BHPS Wave 1-18.
UKHLS	7251	*Understanding Society*: Waves 2-3 Nurse Health Assessment, 2010-2012. Includes cases from the BHPS.

A number of sensitive data files are released under Special Licence (SL). Currently, the only harmonised BHPS data available under SL are included in the core *Understanding Society* SL data. The data are listed as SN 6931 - *Understanding Society*: Waves 1-7, 2009-2016 and harmonised British Household Panel Survey: Waves 1-18: Special Licence Access. The data is a copy of the EUL data (SN 6614) that contains the month of birth, full occupational coding, rare country of

birth/nationality occurrences and uncapped income variables (for both harmonised BHPS and UKHLS files). Additional SL data products such as harmonised geographies will be added in due course.

Researchers can apply for access to SL data through a UKDS application procedure. Researchers will be required to justify their research objectives and explain why EUL data alone would be inadequate to reach those objectives. They will also be asked to report publications resulting from using the data. The conditions for using SL data are provided at http://ukdataservice.ac.uk/get-data/how-to-access/conditions/special-licence.aspx

5. EXAMPLE STATA CODE

In this section we provide examples for using the long run of panel data from the harmonised BHPS and UKHLS files. These examples illustrate how the data may be set up in principle. They do not present a perfect template for all types of analyses.

Figure 3 presents code to extract respondents' age, sex and highest qualification variables from the **indresp** files in all waves and both studies. Wave prefixes are removed, a new wave indicator that works across both sets of files is generated and the files are appended in long format. The resulting file includes information for everyone who has ever provided a full adult or proxy interview in the BHPS or in *Understanding Society*.

Figure 3: Example Stata Code: Merging individual files from harmonised BHPS and UKHLS in long format

```
// Load wave-specific info from BHPS files using loops. Generate wave identifier,
// remove wave prefixes, and save
foreach x in a b c d e f g h i j k l m n o p q r {
  use pidp b`x'_age_dv b`x'_hiqual_dv b`x'_sex using b`x'_indresp, clear
  gen wave=strpos("abcdefghijklmnopqr","`x'")
  save b`x'_indresp_junk, replace
  }
// Load wave-specific info from UKHLS files using loops. Generate wave identifier,
// remove wave prefixes, and save
foreach x in a b c d e f g  {
  use pidp `x'_age_dv `x'_hiqual_dv `x'_sex using `x'_indresp, clear
  rename `x'_* *
  gen wave=strpos("abcdefghijklmnopqr","`x'")+18
  save `x'_indresp_junk, replace
  }
// Append the data files in long format. Note that you can loop over all waves
use ba_indresp_junk, clear
foreach x in bb bc bd be bf bg bh bi bj bk bl bm bn bo bp bq br a b c d e f g {
  append using `x'_indresp_junk
  }
sort pidp wave
save all_indresp, replace
```

Figure 4 presents Stata code to undertake a longitudinal analysis of wellbeing in Britain for the continuing BHPS sample drawing on information from one BHPS and one UKHLS wave (data usage route 3, see Section 1.2.4 above). The appropriate longitudinal population weight is chosen, as well as the variables defining the complex survey design.

Figure 4: Example Stata code: Weighted longitudinal analysis of wellbeing in Britain (Usage 3)

```
use pidp bq_scghq1_dv bq_jbstat using bq_indresp, clear
rename bq_* *
// define a time variable
gen time=1
// Assure yourself that variables named the same are the same across all BHPS
// and UKHLS waves. The code frame for jbstat changed, as you can read in the
// online variable documentation and see by inspecting the variable in both
// studies. First see the BHPS JBSTAT code frame:
fre jbstat
save bq_indresp_junk, replace

use pidp c_scghq1_dv c_jbstat c_indin91_lw c_psu c_strata using c_indresp, clear
rename c_* *
// define a time variable
gen time=2
// Assure yourself that JBSTAT is the same across all BHPS and UKHLS waves.
// E.g. the code frame for jbstat changed in the UKHLS and needs to be recoded:
fre jbstat
recode jbstat (10 11=97)
save c_indresp_junk, replace

// Merge files from different waves. Note that labels in UKHLS are in better shape
// so merge harmonised BHPS files to UKHLS files (or make sure not to overwrite labels)
use c_indresp_junk, clear
append using bq_indresp_junk
// balanced panel: duplicates=1
duplicates tag pidp, gen(dupl)
keep if dupl==1
for var _all: replace X=. if X<0
// distribute unit's last wave longt'l weight, psu and strata to all waves
sort pidp wave
bys pidp: replace indin91_lw=indin91_lw[_N]
bys pidp: replace psu=psu[_N]
bys pidp: replace strata=strata[_N]
// tell Stata this is non-random sample
svyset psu, str(strata) singleunit(scaled)
// Run your stats taking account of complex survey design and attrition,
// e.g., mixed effects random effects regression of GHQ on employment status
svy: meglm scghq1_dv i.jbstat || pidp:, cov(un) pw(indin91_lw)
```

Figure 5 presents Stata code to merge information from the BHPS and UKHLS for a pooled cross-sectional analysis (data usage route 4, see Section 1.2.4 above). The key take-home point is that the variable names for the respective studies' cross-sectional population weights for the UK need to be aligned.

Figure 5: Example Stata code: Weighted cross-sectional time series analysis of wellbeing in the United Kingdom (Usage 4)

```
// Load employment status, GHQ and cross-sectional UK population weight from BHPS files
// Rename weight and compute wave identifier
foreach x in bq bp {
  use pidp `x'_scghq1_dv `x'_jbstat `x'_xrwtuk1 using `x'_indresp, clear
  rename `x'_* *
  gen wave=strpos("abcdefghijklmnopqr","`x'")
  clonevar ukpopw=xrwtuk1
  lab var ukpopw "UK population weight for current wave"
  save `x'_indresp_junk, replace
  }
// Load employment status, GHQ and cross-sectional UK population weight from UKHLS files
// Rename weight and compute wave identifier
foreach x in b c  {
  use pidp `x'_scghq1_dv `x'_jbstat `x'_indinub_xw using `x'_indresp, clear
  rename `x'_* *
  gen wave=strpos("abcdefghijklmnopqr","`x'")+18
  clonevar ukpopw=indinub_xw
  lab var ukpopw "UK population weight for current wave"
  // align code frames for jbstat
  recode jbstat (10 11=97)
  save `x'_indresp_junk, replace
  }
// Append BHPS and UKHLS files. Note that labels in UKHLS are in better shape so merge
// harmonised BHPS files to UKHLS
use c_indresp_junk, clear
foreach x in b bp qq {
  append using `x'_indresp_junk
  }
// sort by unit and time
sort pidp wave
for var _all: replace X=. if X<0
// tabulate population estimates of average GHQ by employment status for the two periods
table jbstat wave [pw= ukpopw], c(mean scghq1_dv)
```

6. THE *UNDERSTANDING SOCIETY* HARMONISED BHPS PROJECT: NEXT STEPS

The harmonised BHPS project is ongoing. We welcome user feedback on the data and look forward to your suggestions for improvement. Please email your feedback to: consult@understandingsociety.ac.uk

User feedback will be considered in the future development work. We plan to publish a first updated harmonised BHPS in Summer 2018 as part of a planned within wave release and a second updated version in November 2018 as part of the standard Wave 8 data release.

7. CITATIONS AND ACKNOWLEDGEMENTS

Any publication, *w*hether printed, electronic or broadcast, based *w*holly or in part on the *Understanding Society* data collection provided by the UK Data Service must be

accompanied by the correct citation and acknowledge the Institute for Social and Economic Research as the data provider and the UK Data Service as the data distributor. The acknowledgement, which gives credit to sponsors or distributors, is not a replacement for a proper citation. We recommend the following wording:

> "*Understanding Society* and BHPS are funded by the Economic and Social Research Council and various Government Departments, with scientific leadership by the Institute for Social and Economic Research, University of Essex, and survey delivery by NatCen Social Research and Kantar Public. The research data are distributed by the UK Data Service."

7.1. Citation of the Data

The format for bibliographic references is as follows:

> University of Essex. Institute for Social and Economic Research, NatCen Social Research and Kantar Public, [producers]: *Understanding Society*: Waves 1-7, 2009-2016 and harmonised British Household Panel Survey: Waves 1-18 [computer file]. 9th Edition. Colchester, Essex: UK Data Service [distributor], November 2017. SN: 6614, http://dx.doi.org/10.5255/UKDA-SN-6614-8.http://dx.doi.org/10.5255/UKDA-SN-6614-9.

7.2. Citation of the User Guide

This User Guide is to be cited as follows:

> Fumagalli, Laura, Knies, Gundi and Buck, Nick (2017): *Understanding Society*: The UK Household Longitudinal Study harmonised British Household Panel Survey (BHPS) User Guide. Colchester: University of Essex.

7.3. Acknowledgments

People who contributed to the writing of the *Understanding Society* harmonised BHPS User Guide for this release include, in alphabetical order: Stephanie Auty, Michaela Benzeval, Olena Kaminska and Peter Lynn. The *Understanding Society* harmonised BHPS project is led by Laura Fumagalli and Nick Buck with the support of the *Understanding Society* team. A very big "Thank you!" goes to the many people who have and continue to contribute to the unrelenting success and timely delivery of *Understanding Society* and particularly to the harmonised BHPS project They include, in addition to the aforementioned, (in alphabetical order): Geoff Angel, Gina Anghelescu, Glenn Dunn, Paul Fisher, Nicole James, Graham Jolliffe, Alita Nandi, Jonathan Nears, John Payne, Elaine Prentice-Lane, and Catherine Yuen. We would also like to thank our beta users for their valuable comments on test versions of the data. These beta users are, in alphabetical order: Giuseppe Forte, Vernon Gayle, Alessandra Guariglia, Andrew Henley, Gundi Knies, Kenisha Russell Johnson, Bertha Rohenkohl Cruz, Alireza Sepahsalari, Peter Spittal, Joanna Tyrowicz, George Ward, David Wright, Min Zheng, and Muzhi Zhou.

8. REFERENCES

Knies, G. (2017). *Understanding Society*: *Understanding Society*: Waves 1-7, 2009-2016 and harmonised British Household Panel Survey: Waves 1-18, 1991-2009, User Guide., November 2017. ISER, University of Essex.

Lynn, P., N. Buck, et al. (2006). Quality Profile: British Household Panel Survey. Waves 1 to 13: 1991-2003. P. Lynn. Colchester, University of Essex. Institute for Social and Economic Research.

Taylor, M. F. (2010). British Household Panel Survey User Manual Volume A: Introduction, technical report and appendices. Colchester, University of Essex.

9. APPENDIX

Appendix 1: Variables included in both the harmonised BHPS and *Understanding Society*: individual enumeration files (indall)

Topic	Variable stem name
Identifiers	hidp pid pidp pno ppno
Demographics	age_dv age_if
Household	birthm birthy depchl_dv hgbiof hgbiom hhmem jnmnth jnyear nchild_dv rach16_dv
Other variables	hhorig ivfho ivfio iviolw memorig sampst

Appendix 2: Variables included in both the harmonised BHPS and *Understanding Society*: household response files (hhresp)

Topic	Variable stem name
Identifiers	hidp ivh0-ivh16 rentp0-rentp16
Household	agechy_dv carown carval cduse1 cduse10 cduse11 cduse12 cduse13 cduse2 cduse3 cduse4 cduse5 cduse6 cduse7 cduse8 cduse9 cduse96 heatch hhsize hsbeds hscost hsctax hsjb hsowr0-hsowr116 hsrooms hsval hsyr04 hsyrbuy mgextra mglife mgnew mgold mgxty1 mgxty2 mgxty3 mgxty4 nch02_dv nch1215_dv nch34_dv nch511_dv ncouple_dv nemp_dv nkids_dv nonepar_dv npens_dv nwage_dv pcbroad pcnet rent rentf rentll solar1 solar2 solar3 tenure_dv xpgasy xphsdb xpmg xpoily xpsfly
Employment/ Income	fihhmngrs_dv fihhnegsei_if
Neighborhood	gor_dv
Other variables	hienddathh hienddatmm histrtdathh histrtdatmm intdated intdatem intdatey

Appendix 3: Variables included in both the harmonised BHPS and *Understanding Society*: adult response files (indresp)

Topic	Variable stem name
Identifiers	hidp pid pidp pno ppno pripn mpno hgbiof hgbiom debtpn0-debtpn16
Demographics	age_dv mstatsam sex plbornc yr2uk4 pgmrob pgprob
Parental background	macob maedqf maju masoc00 masoc00_cc masoc90 masoc90_cc mgmrob mgprob pacob paedqf paju pasoc00 pasoc00_cc pasoc90 pasoc90_cc
Household	agelh birthm birthy caidu1 caidu2 caidu3 caidu4 caidu5 caidu6 caidu7 caidu8 caidu96 caidu97 caruse ch1bm ch1by4 chaid1 chaid2 chaid3 chaid4 chaid5 chaid6 chaid7 chaid8 chaid96 chaid97 chfar chsee coh1bm coh1by coh1em coh1ey coh1mr drive farkid howlng huboss hubuys hufrys huiron humops husits ladopt lchmor lchmorn lcmarm lcmary4 lcmcbm lcmcby4 lcmcoh lcmspm lcmspy4 lcoh lmar1m lmar1y lnadopt lncoh lnprnt lprnt lvag16 mlstat nchild_dv paaid1 paaid2 paaid3 paaid4 paaid5 paaid6 paaid7 paaid8 paaid96 paaid97 paidu1 paidu2 paidu3 paidu4 paidu5 paidu6 paidu7 paidu8 paidu96 paidu97 parmar rach16_dv seekid wekid ynlp14
Employment/	ageret basnsa basrate basrest debtc1 debtc2 debtc3 debtc4

Income	debtc5 debtsj debtsk debty eprosh extnsa extrate extrest fimngrs_dv fimnlabgrs_dv finnow fiyrdb1 fiyrdb2 fiyrdb3 fiyrdb4 fiyrdb5 fiyrdb6 fiyrdia huruns j1boss j1mngr j1semp j1soc00 j1soc00_cc j1soc90 j1soc90_cc j2has j2hrs j2pay_dv j2pay_if j2semp j2soc00 j2soc00_cc j2soc90 j2soc90_cc jbbgd jbbgm jbbgy jbhad jbhas jbhrs jbisco88 jbisco88_cc jblkcha jblkchb jblkchc jblkchd jblkche jbmngr jboff jbonus jbot jbotpd jbpen jbpenm jbpeny4 jbperfp jbpl jbrise jbsect jbsemp jbsize jbsoc00 jbsoc00_cc jbstat jbterm1 jbterm2 jbttwt jbxpcha jbxpchb jbxpchc jbxpchd jbxpche jlboss jlendy jlisco88 jlisco88_cc jlmngr jlsemp jlsize jlsoc00 jlsoc00_cc jsaccs jsboss jshrs jspart jspayu jspayw jspl jsprbm jsprby4 jsprem jsprey4 jsprf jsprls jsprni jsprtx jspyni jspytx jssize jsttwt jsttwtb jstypeb jubgn julk4wk julk4x1 julk4x2 julk4x4 julk4x5 julk4x6 julkjb niserps ovtnsa ovtpay ovtrate ovtrest payg_dv paygl paygu_dv paygu_if payn_dv paynl paynu_dv paynu_if paytyp payu payug payusl penmcn penmex penmpy penmtp penspb pjulk4wk pppex pppexm retamt retsuf rtfnd1 rtfnd10 rtfnd2 rtfnd3 rtfnd4 rtfnd5 rtfnd6 rtfnd7 rtfnd8 rtfnd9 rtfnd96 rtpro5 rtpro6 save saved savlt savreg sppen svacts1 svacts2 svacts5 tuin1 tujbpl wktime
Education	edasp edtype feend hiqual_dv scend school
Health	adla adlad adlb adlbd adlc adlcd adld adldd adle adled adlf adlfd aglquit hl2gp hl2hop hlht hlhtc hlhtf hlhti hlwte hlwtk hlwtl hlwtp hlwts hosp hospch hospd ncigs scghqa scghqb scghqc scghqd scghqe scghqf scghqg scghqh scghqi scghqj scghqk scghql sf1 sf2a sf2b sf5 smcigs smever smncigs smnow smoker
Neighborhood	crdark crwora crworb gor_dv hood15 lkmove locserb locserc locserd locsere mvmnth mvyr plnowm plnowy4 xpmove
Caring	aidhh aidhrs aidhu1 aidhu2 aidxhh
Personality/ Attitudes	scopfama scopfamb scopfamd scopfamf scopfamh scptrt5a1 scptrt5a2 scptrt5a3 scptrt5c1 scptrt5c2 scptrt5c3 scptrt5e1 scptrt5e2 scptrt5e3 scptrt5n1 scptrt5n2 scptrt5n3 scptrt5o1 scptrt5o2 scptrt5o3
Politics	flyes1 flyes2 flyes3 flyes4 howlng impevent1 impevent1s impevent2 impevent2s impevent3 impevent3s impevent4 impevent4s nflyeu nflyin nflyos paperm2 vote1 vote2 vote3 vote4 vote5 vote6 vote7 vote8
Religion	oprlg oprlg1 oprlg2 oprlg3
Satisfaction	scghq1_dv scghq2_dv
Social Capital/ Leisure	netag_1 netag_2 netag_3 netkn_1 netkn_2 netkn_3 netph_1 netph_2 netph_3 netrl_1 netrl_2 netrl_3 netsx_1 netsx_2 netsx_3 netwr_1 netwr_2 netwr_3 org orga1 orga10 orga11 orga12 orga13 orga14 orga15 orga16 orga2 orga3 orga4 orga5 orga6 orga7 orga8 orga9 orgm1 orgm10 orgm11 orgm12 orgm13 orgm14 orgm15 orgm16 orgm2 orgm3 orgm4 orgm5 orgm6 orgm7 orgm8 orgm9 scssupr2r sctrust
Ethnic identity	wlsh1 wlsh2 wlsh3 wlsh4 wlsh96 wlshua wlshub wlshuc wlshud
Young adults	futra futrb futrc futrd futre futrf futrg futrh futri futrj futrk futrl ocimpa ocimpb ocimpe ocimpf ocimpi ocimpk ocimpl yasoc00 yasoc00_cc yasoc90 yasoc90_cc
Other topics	hhorig ienddathh istrtdatd istrtdathh istrtdatm istrtdaty ivfio ivprsnt memorig mobuse netuse remail_code sampst

Appendix 4: Variables included in both the harmonised BHPS and *Understanding Society*: individual youth response files (youth)

Topic	Variable stem name
Identifiers	hidp pid pidp pno fnspno mnspno
Demographics	ypdobm ypdoby ypsex
Family relationships	ypargf ypargm yptlkf yptlkm
Plans for the future	ypfuta ypfutb ypsoc00 ypsoc90 ypamar ypapar
Self-esteem	ypesta ypestb ypestc ypeste ypestf ypesti ypestj ypestk
Happiness	yphap yphfm yphfr yphlf yphsc yphsw
Other topics	ypcomp ypcrwra ypfght ypjfd yplate yplvhm ypmobu ypnbuks ypnpal yppchw ypsave yptrvl2sch yptrydiet ypvand ypvte6 ypwklw