



Understanding Society: Marital and Cohabitation Histories, 1991-2024

User Guide

Version 3, March 2026

Institute for Social and Economic Research
University of Essex



Contents

1. Introduction	1
2. Data and analysis guidance.....	3
2.1 File structure and description of variables	3
2.2 How the variables in this dataset were created	4
2.3 Imputing missing dates and spells.....	5
2.4 Missing values.....	6
3. Data file Description.....	7
4. Data Quality and reporting errors.....	13
5. Example of syntax to add characteristics for respondents and partners	15
6. Data Access	16
7. Citation.....	16
Citing this User Guide	16
Acknowledgements	16
Appendix A.....	17
Appendix B:.....	18

1. Introduction

[Understanding Society](#) is a longitudinal household panel survey that started in 2009 and attempts to interview the core sample members at approximately one-year intervals, as long as they are living in the UK. Each set of interviews are referred to as waves. To learn more about the survey please visit the [About the Study](#) web page and the [Main Survey User Guide](#)¹. The Main Survey data are available under [study number 6614](#) and can be linked to this dataset's data files **phistory_long** and **phistory_wide** using the variable **pidp**.

Understanding Society: Marital and Cohabitation Histories, 1991-2024 **SN 8473** (also known as the partnership history file) contains information about partnership spells reported by adult respondents in all Understanding Society (UKHLS) and British Household Panel Survey (BHPS) samples up to Wave 15. The [British Household Panel Survey](#) began in 1991 and is the base on which Understanding Society is built. Information about the different samples can be found in the [Main Survey User Guide](#).

The UKHLS sample initially collected partnership history and current partnership status information in Wave 1. In subsequent waves information is collected on changes since the last interview. For new entrants after Wave 1, only partial information about their past partnership history is collected (See section 2 for further information on how the variables were created).

For BHPS sample members, information about their partnership history until 2008 has been extracted from the single partnership history file created by Pronzato² (2009). For those BHPS sample members who were interviewed as part of Understanding Society (from 2010 onwards), their prior information (collected during the BHPS interviews) was combined with the information collected during these interviews.

The Understanding Society: Marital and Cohabitation Histories, 1991-2024 dataset is based on the following datasets³ which are available from the UK Data Service:

SN 6614: Understanding Society: Waves 1-14, 2009-2023 and Harmonised BHPS: Waves 1-18, 1991-2009

SN 5151: British Household Panel Survey: Waves 1-18, 1991-2009

SN 5629: British Household Panel Survey Consolidated Marital, Cohabitation and Fertility Histories, 1991-2009

¹ Institute for Social and Economic Research. (2025). Understanding Society: Waves 1-15, 2009-2024 and Harmonised BHPS: Waves 1-18, 1991-2009, User Guide, 23 February 2026, Colchester: University of Essex.

² [British Household Panel Survey Consolidated Marital, Cohabitation and Fertility Histories, 1991-2009](#).

³ See Appendix A for the source data files.

Data users are encouraged to read the user manuals for the datasets used to produce this data file to understand more about the underlying surveys and datasets (Taylor et al 2009, Pronzato 2009⁴). All individuals in this dataset can be linked to their Understanding Society data in the main study **SN 6614**, using the unique cross-wave identifier **pidp**. The unique cross-wave identifier **pp** in **SN 5629** is the same as **pid** (the unique cross-wave identifier for non-harmonised BHPS files) included in **SN 5151**⁵. The origin of different sample members can be identified in this dataset by the variable **hhorig**.

⁴ [Taylor, M., Brice, J., Buck, N., Prentice-Laine, E. \(ed.\) \(2009\). British Household Panel Survey User Manual Volume A: Introduction, Technical Report and appendices.](#)

⁵ Note: Those who joined the households of BHPS sample members after the BHPS ended, that is, as part of Understanding Society, do not have a valid existing PID. Therefore, PID equals -8.

2. Data and analysis guidance

2.1 File structure and description of variables

This dataset (**SN 8473**) includes two data files **phistory_long** and **phistory_wide**.

The first data file **phistory_long**, is in long format, i.e. each row represents one spell for each person in the dataset. A spell is the period of time a respondent spent in a partnership, including marriage, cohabitation, and civil partnership. Each reported change in partnership status is recorded as a new spell. In this section the long format file and its content is described. The second file, **phistory_wide** is in wide format, that is, each row represents one person. In this wide format file, variable names pertaining to spells have a number suffix which represents the spell number. Both files only include individuals who have provided information about at least one union or partnership, so all persons in the dataset have at least one partnership spell. The following variables are available in this dataset:

- **pidp**: This is the unique cross-wave person identifier available within the **SN 6614** data files. This identifier can be used to link information to the main survey **SN 6614** data files.
- **pid**: This is the unique cross-wave person identifier available within the **SN 5151** data files. This variable is also available in the **SN 6614** data files (except for any of the first wave files as these do not include BHPS sample members). This variable equals -8 for all UKHLS sample members and new BHPS sample members (that is, individuals who joined the BHPS sample members' households for the first time during the UKHLS survey period).
- **hhorig**: This variable identifies the sample origin for each respondent.
- **spellno**: This is the number of the reported spell of each individual, where the first spell is the earliest partnership.
- **spellnoR**: This is the number of the reported spell of each individual in reverse order, where the first spell is the most recent partnership.
- **status**: This variable identifies whether the partnership is a marriage, civil partnership or cohabitation (living together as a couple). It takes on the following values: 2 'Married', 3 'Civil Partner', and 10 'Living together as a couple'. Note the coding frame is the same as the marital status derived variable **w_mastst_dv** in the **w_indall** and **w_indresp** files.
- **partner**: This variable records the **pidp** of the partner reported for each spell, where this information was available.
- **starty & startm**: This is the year and month of the start date of the spell.
- **endy & endm**: This is the year and month of the end date of the spell.
- **divorcey & divorcem**: Where both marriage separation and divorce dates are available, the variable **enddate** represents the separation date, and the divorce date is recorded here. If only the separation date is available this is a valid skip (-8).
- **start_if**: This variable identifies whether the start date has been imputed. It takes on the following values: 0 is no imputation, 1 is month imputed, and 2 identifies that both month and year have been imputed.

- **end_if:** This variable identifies if the end date has been imputed. It takes on the following values: 0 is no imputation, 1 is month imputed, and 2 identifies that both month and year have been imputed.
- **divorce_if:** This variable identifies if the divorce date has been imputed. It takes on the following values: 0 is no imputation, 1 is month imputed, and 2 identifies that both month and year have been imputed.
- **spell_if:** This takes on the value 2 if the whole spell has been imputed. When the marital or partnership status changed between two consecutive wave interviews, but no spell was reported in the interview, the entire spell had to be imputed.
- **ongoing:** This binary variable identifies whether a spell is ongoing at the time of the data collection. It takes on the values of 0 and 1, with 1 indicating an ongoing spell and 0 a spell that has ended.
- **lastinty & lastintm:** This is the year and month of the last interview date.
- **startdate, enddate, divorcedate, lastintdate:** These are the start date, end date, divorce date, and the last interview date in Stata format, meaning that they are measured in months since January 1960. As Stata data formatting has been attached to the variables, these will appear as the correct date. For example, if the start date is January 1960 then **startdate** will have the value 2 and it will appear on screen as 1960m1.
- **ttd_married, ttd_civil_partnership, ttd_cohabit:** These variables measure the total number of marriage, civil partnership, and cohabitation spells for each person.
- **ever_married, ever_civil_partnership, ever_cohabit:** These variables identify if a person has ever been married, in a civil partnership, or in a non-marital cohabitation.
- **sampst:** This variable indicates sample status.
- **ever_proxy_ukhls:** This variable identifies whether the respondent has had completed at least one proxy interview during the UKHLS waves.
- **every_proxy_bhps:** This variable identifies whether the respondent has had completed at least one proxy interview during the BHPS waves 1-18.

2.2 How the variables in this dataset were created

In Wave 1 all adult respondents (aged 16+) were asked about their (i) current cohabitation, (ii) current marriage and the cohabitation spell preceding that with the same partner, (iii) past cohabitation and marriage history, if applicable. If the current marriage spell was preceded by a cohabitation spell with the same partner, then the cohabitation end date was not asked and was set to the month prior to the marriage start date. From the second wave onwards, they are asked about changes to their marital status and information about these changes. They are also asked about any new cohabitation spells since the last interview.

From the second wave onwards, new household members who join the study are asked about their first marriage and their total number of marriages, when they are interviewed for the first time, so that data on the start date of their first marriage is available. If the respondents are still married at the time of the interview, then their marriage end date has

been set to the interview date and listed as an ongoing spell. All other marriage dates are set to -9.

For ongoing spells, the end date has been set to the date of interview. The **ongoing** flag can be used to identify such cases. Please note that as respondents are not interviewed annually, the current interview date may vary across respondents. The variable **lastintdate** records the date that they were last interviewed. The month and year of this date are also recorded and are available as separate variables, **lastinty** (last interview year) and **lastintm** (last interview month).

2.3 Imputing missing dates and spells

In this section you can find information on how specific types of missing information were imputed. Please note that there were some missing dates that could not be imputed and have been left as is. Issues with the date data, such as overlapping spells or start dates being after the end dates of spells, have only been corrected where additional information was provided that led to a logical conclusion. In some cases, our imputation rules resulted in overlapping spells, some easy to fix, and others that required further investigation. For the latter, we looked at the entire partnership history and implemented corrections using information on the status, partner and logical timing. Where conflicting information has been reported between the annual events history (changes between interviews), and the partnership status identified at the time of the interview, priority was given to the information reported in the annual events history.

Information on past partnership history is incomplete or missing for new entrants and for those who have completed proxy interviews. Variables **ever_proxy_ukhls** and **ever_proxy_bhps** have been included to help identify cases where the participants have had a proxy interview at least once in the past, and their data might be incomplete. Variable **sampst** has also been included to help identify new entrants (**sampst** = 2 (PSM) & **sampst** = 3 (TSM)) for the same purposes.

The following points detail the imputations and corrections made in the case of missing or inconsistent data:

- If **startm** or **endm** of a spell was missing but a valid **starty** or **endy** was reported, the month was imputed by setting it to June. In a few cases this resulted in overlapping spells which required further correction. For example, if a respondent reported that a cohabitation spell ended in 2018 without specifying the month, and also reported another cohabitation spell that started in April 2018, our imputation rules would calculate the end date for the first cohabitation spell to be June 2018, resulting in overlapping spells. In cases like this we corrected the end date of the first cohabitation spell to be the month prior the start date of the second cohabitation spell. In the example this would be March 2018.
- If marriage spells followed cohabitation spells with the same partner, then the end date of the cohabitation spell was set to the month before the marriage start date, if that information was available. If the start date was missing, it was calculated to be

the mid-point between the cohabitation start date and the latest interview date, or the marriage end date.

- If a change in marital status was noted between waves, but the respondent did not report any changes in their latest interview (either because they were not asked the interview questions or because they simply did not report the changes), the dates were imputed. Namely, **startm** was set to June, and **starty** was set to the latest interview date, while **enddate** was set to the mid-point between the imputed **startdate** and **lastintdate**.
- Those who have never responded but their marital status was recorded initially as 'married' or 'cohabiting as a couple', could not be included in the partnership history file due to lack of data. Cases where adult respondents reported invalid dates for their current marriage and have no recorded past history, have also been excluded from the file. These cases can be identified through variables **evermar_dv** and **evercoh_dv** found in the **xwavedat** file, and variables **ever_married**, **ever_civil_partnership** and **ever_cohabit** in the partnership history files.

2.4 Missing values

The Marital and Cohabitation Histories dataset follows the same [missing values structure](#) as the BHPS and UKHLS study. That is:

- 1 Don't know – When the respondent does not know the answer
- 2 Refused – When the respondent refuses to answer a question
- 7 Proxy
- 8 Valid skip
- 9 Missing by error or implausible

However, as the date variables **startdate**, **enddate**, and **divorcedate** are in Stata date format, any -8 and -9 values would be interpreted by Stata as dates, 8 and 9 months before January 1960 respectively. To avoid any confusion, a value of -720 has been assigned if the date is missing (-9) and -721 if there is a valid skip (-8). This translates to January 1900 and December 1899 respectively, both impossible date values for the dataset.

3. Data file Description

A brief description of the **phistory_long** data file, including names, value labels, and variable labels, is shown in Figure 1. The data file includes 140,676 observations and 35 variables. This data file only includes adult respondents who have reported at least one partnership spell, so the file represents 85,871 adult respondents and 140,676 spells.

Figure 1: Description of data file “phistory_long”

variable name	storage type	display format	value label	variable label
pidp	long	%10.0f	pidp	cross-wave person identifier (public release)
pid	long	%12.0g	pid	person number identification
spellno	float	%9.0g		spell no, earliest first
spellnoR	float	%9.0g		spell no, recent first
status	float	%41.0g	status	partnership status
partner	long	%12.0g	partner	partner pidp
startdate	float	%tm	startdate	start date
starty	float	%9.0g	starty	start year
startm	float	%9.0g	startm	start month
start_if	float	%18.0g	start_if	start date imputation flag
enddate	double	%tm	enddate	end date
endy	float	%9.0g	endy	end year
endm	float	%9.0g	endm	end month
end_if	float	%18.0g	end_if	end date imputation flag
divorcedate	float	%tm	divorcedate	divorce date if available
divorcey	float	%16.0g	divorcey	divorce year
divorcem	float	%16.0g	divorcem	divorce month
divorce_if	float	%18.0g	divorce_if	divorce date imputation flag
mrgend	float	%10.0g	mrgend	how marriage/civil partnership ended
cohend	float	%9.0g	cohend	how cohabitation ended
ongoing	float	%13.0g	ongoing	ongoing spell indicator
t1l_spells	float	%9.0g		total number of any partnership spells
t1l_married	float	%9.0g		total number of marriage spells
t1l_civil_pa~p	float	%9.0g		total number of civil partnership spells
t1l_cohabit	float	%9.0g		total number of cohabitation spells
ever_married	float	%9.0g		ever married
ever_civil_pa~p	float	%9.0g		ever in civil partnership
ever_cohabit	float	%9.0g		ever cohabit
lastintdate	float	%tm		last interview date
lastinty	float	%9.0g		last interview year
lastintm	float	%9.0g		last interview month
hhorig	float	%27.0g	hhorig	sample origin, household
sampst	byte	%12.0g	sampst	final sample status
ever_proxy_uk~1	float	%28.0g	ever_proxy_ukhls	proxy interview during UKHLS waves?
ever_proxy_bh~1	float	%28.0g	ever_proxy_bhps	proxy interview during BHPS waves 1-18?

Summary statistics of all the variables included in the **phistory_long** data file can be found in Figure 2.

Figure 2: Summary statistics of data file “phistory_long”

Variable	Obs	Mean	Std. dev.	Min	Max
pidp	140,676	7.14e+08	4.98e+08	2727	1.67e+09
pid	140,676	1.49e+07	3.53e+07	-8	1.89e+08
spellno	140,676	1.637913	1.0246	1	14
spellnoR	140,676	1.637913	1.0246	1	14
status	140,676	5.337542	3.940637	2	10
partner	140,676	4.37e+08	5.16e+08	-9	1.67e+09
startdate	140,676	290.733	373.173	-720	779
start_if	140,676	.1080284	.3951842	0	2
enddate	140,676	488.2124	344.0839	-721	784
end_if	140,676	.1435995	.4592369	0	2
divorcedate	140,676	-633.8251	305.4265	-721	776
divorce_if	140,676	-7.368208	2.190156	-8	2
mrgend	140,676	-2.8284	4.472694	-8	4
cohend	140,676	-4.152265	4.578189	-8	2
ongoing	140,676	.4443046	.4968901	0	1
t1l_spells	140,676	2.275825	1.450711	1	14
t1l_married	140,676	1.121122	.7283792	0	7
t1l_civil~p	140,676	.0103145	.1051722	0	3
t1l_cohabit	140,676	1.144524	1.231696	0	12
ever_married	140,676	.8431644	.363647	0	1
ever_civil~p	140,676	.0099164	.0990865	0	1
ever_cohabit	140,676	.6615556	.4731822	0	1
lastintdate	140,676	680.2967	90.22351	377	784
lastinty	140,676	2016.233	7.571835	1991	2025
lastintm	140,676	6.497121	3.468945	1	12
hhorig	140,676	3.721864	5.021553	1	22
sampst	140,676	1.218452	.6006793	1	3
ever_prom~sl	140,676	.0805894	.2722045	0	1
ever_prom~psl	140,676	.0166837	.1280839	0	1
starty	140,676	1816.278	566.3738	-9	2024
startm	140,676	5.226748	5.323191	-9	12
endy	140,676	1888.221	473.9273	-9	2025
endm	140,676	5.543106	4.822652	-9	12
divorcey	140,676	145.8678	533.2757	-9	2024
divorcem	140,676	-6.897708	3.906088	-9	12

A brief description of the **phistory_wide** data file, including names, value labels, and variable labels, is shown in Figure 3. In the data file **phistory_wide**, there are 85,871 observations and 254 variables.

Figure 3: Description of the wide format of the data

variable name	storage type	display format	value label	variable label
pidp	long	%10.0f	pidp	cross-wave person identifier (public release)
status1	float	%41.0g	status	1 status
partner1	long	%12.0g	partner	1 partner
startdate1	float	%tm	startdate	1 startdate
starty1	float	%9.0g	starty	1 starty
startm1	float	%9.0g	startm	1 startm
start_if1	float	%18.0g	start_if	1 start_if
enddate1	double	%tm	enddate	1 enddate
endy1	float	%9.0g	endy	1 endy
endm1	float	%9.0g	endm	1 endm
end_if1	float	%18.0g	end_if	1 end_if
divorcedate1	float	%tm	divorcedate	1 divorcedate
divorcey1	float	%16.0g	divorcey	1 divorcey
divorcem1	float	%16.0g	divorcem	1 divorcem
divorce_if1	float	%18.0g	divorce_if	1 divorce_if
mrgend1	float	%10.0g	mrgend	1 mrgend
cohend1	float	%9.0g	cohend	1 cohend
ongoing1	float	%13.0g	ongoing	1 ongoing
status2	float	%41.0g	status	2 status
partner2	long	%12.0g	partner	2 partner
startdate2	float	%tm	startdate	2 startdate
starty2	float	%9.0g	starty	2 starty
startm2	float	%9.0g	startm	2 startm
start_if2	float	%18.0g	start_if	2 start_if
enddate2	double	%tm	enddate	2 enddate
endy2	float	%9.0g	endy	2 endy
endm2	float	%9.0g	endm	2 endm
end_if2	float	%18.0g	end_if	2 end_if
divorcedate2	float	%tm	divorcedate	2 divorcedate
divorcey2	float	%16.0g	divorcey	2 divorcey
divorcem2	float	%16.0g	divorcem	2 divorcem
divorce_if2	float	%18.0g	divorce_if	2 divorce_if
mrgend2	float	%10.0g	mrgend	2 mrgend
cohend2	float	%9.0g	cohend	2 cohend
ongoing2	float	%13.0g	ongoing	2 ongoing

Up to...

status14	float	%41.0g	status	14	status
partner14	long	%12.0g	partner	14	partner
startdate14	float	%tm	startdate	14	startdate
starty14	float	%9.0g	starty	14	starty
startm14	float	%9.0g	startm	14	startm
start_if14	float	%18.0g	start_if	14	start_if
enddate14	double	%tm	enddate	14	enddate
endy14	float	%9.0g	endy	14	endy
endm14	float	%9.0g	endm	14	endm
end_if14	float	%18.0g	end_if	14	end_if
divorcedate14	float	%tm	divorcedate	14	divorcedate
divorcey14	float	%16.0g	divorcey	14	divorcey
divorcem14	float	%16.0g	divorcem	14	divorcem
divorce_if14	float	%18.0g	divorce_if	14	divorce_if
mrgend14	float	%10.0g	mrgend	14	mrgend
cohend14	float	%9.0g	cohend	14	cohend
ongoing14	float	%13.0g	ongoing	14	ongoing
pid	long	%12.0g	pid		person number identification
t1l_spells	float	%9.0g			total number of any partnership spells
t1l_married	float	%9.0g			total number of marriage spells
t1l_civil_par~p	float	%9.0g			total number of civil partnership spells
t1l_cohabit	float	%9.0g			total number of cohabitation spells
ever_married	float	%9.0g			ever married
ever_civil_pa~p	float	%9.0g			ever in civil partnership
ever_cohabit	float	%9.0g			ever cohabit
lastintdate	float	%tm			last interview date
lastinty	float	%9.0g			last interview year
lastintm	float	%9.0g			last interview month
hhorig	float	%27.0g	hhorig		sample origin, household
sampst	byte	%12.0g	sampst		final sample status
ever_proxy_uk~1	float	%28.0g	ever_proxy_ukhls		proxy interview during UKHLS waves?
ever_proxy_bh~1	float	%28.0g	ever_proxy_bhps		proxy interview during BHPS waves 1-18?

Table 1 shows the distribution of spells across the respondents. There are 15,665 respondents who have not reported any marriage spell but have reported other types of partnership spells. About 82% of these 84,833 respondents have reported at least one marriage spell, 0.7% have reported at least one civil partnership spell, and 48.2% have reported at least one cohabitation spell.

Table 1: Distribution of partnership spells				
Number of spells	Number of adult respondents with these types of spells			
	Any partnership	Marriage	Civil partnership	Living as a couple
0		15,665	85,233	44,509
1	52,603	60,339	620	29,285
2	19,859	8,666	17	8,694
3	8,356	1,077	1	2,288
4	3,203	109		710
5	1,145	12		235
6	422	1		77
7	160	2		36
8	62			12
9	34			11
10	14			9
11	4			2
12	7			3
14	2			
Total	85,871	85,871	85,871	85,871

Table 2 shows the distribution of respondents with at least one marriage, civil partnership or cohabitation (living as a couple) spell across the different samples. Information about the different samples can be found in the [Main Survey User Guide](#) along with a [timeline](#) of when each sample was added. Note from 1997 to 2001 the [BHPS provided data for the ECHP](#).

Table 2: Distribution of respondents with at least one marriage, civil partnership or cohabitation (living as a couple) spell across the different samples	
Sample	Number of adult respondents
UKHLS General Population - GPS-GB sample	41,434
UKHLS General Population - GPS-NI sample	1,932
UKHLS Ethnic Minority Boost - EMB sample	6,500
UKHLS Immigrant and Ethnic Minority Boost IEMB sample	3,766
BHPS GB sample	14,498
BHPS Scottish Boost sample	2,940
BHPS Welsh Boost sample	3,162
BHPS Northern Irish Boost sample	3,266
European Community Household Panel– SCPR ⁶	550
European Community Household Panel– ONS⁷	508
European Community Household Panel- NI⁸	152
UKHLS General Population Boost - GPS2-GB sample	6,962
UKHLS General Population Boost - GPS2-NI sample	201
Total number of adult respondents	85,871

⁶ Low-income sub-samples of ECHP in Great Britain interviewed by SCPR (now NatCen Social Research). See [BHPS User Manual, Volume A](#), section A2-3.

⁷ Low-income sub-samples of ECHP in Great Britain interviewed by the Office for National Statistics (ONS). See [BHPS User Manual, Volume A](#), section A2-3.

⁸ All households in Northern Ireland in which all adult members responded at ECHP Wave 3 (1996). See [BHPS User Manual, Volume A](#), section A2-3.

4. Data Quality and reporting errors

There are 7 spells where the end date is later than the divorce date and 2,319 spells where the start date is earlier than the end date of the previous spell (overlapping spells). These are the cases that could not be corrected with the information available. Table 3 below shows the number of spells with missing start or/and end dates.

Table 3: Distribution of spells with missing start and/or end dates			
	Spell end data missing		
Spell start date missing	No	Yes	Total
No	122,655	5,676	128,331
Yes	9,765	2,580	12,345
Total	132,420	8,256	140,676

Table 4 shows the number of spells where the start and/or end dates were imputed.

Table 4: Distribution of spells imputed start and/or end dates				
	Spell end date imputed			
Spell start date imputed	No	Yes, month imputed	Yes, whole date imputed	Total
No	120,247	4,924	4,515	129,686
Yes, month imputed	3,899	2,777	107	6,783
Yes, whole date imputed	2,513	132	1,562	4,207
Total	126,659	7,833	6,184	140,676

Comparing the partnership status and partnership **pidp** in this dataset with that reported in the individual adult respondent files for cases where the partnership spell was ongoing at the time of the interview, **w_indresp**, some mismatches were found in partner **pidp** (**partner vs w_ppid**) and current marital status (**status vs w_mastat_dv**), as seen in Table 5.

Table 5: Mismatches in information at the time of interview between this partnership history file and the information collected at the interview			
	Respondent is living with a partner at the time of this interview based on phistory_long but not according to w_indresp	Respondent is living with a partner at the time of this interview based on w_indresp but not according to phistory_long	Mismatch in partnership status at the time of this interview
Wave 1	0	0	76
Wave 2	0	23	79
Wave 3	0	21	72
Wave 4	0	20	65
Wave 5	0	22	61
Wave 6	0	16	78
Wave 7	0	17	136
Wave 8	0	21	88
Wave 9	0	19	73
Wave 10	0	24	73
Wave 11	0	26	75
Wave 12	0	28	59
Wave 13	31	21	71
Wave 14	52	41	94
Wave 15	463	234	554

5. Example of syntax to add characteristics for respondents and partners

The variables **pidp** and **partner** are the id of the respondent and their reported partner and can be used to add individual level characteristics from the main survey (**SN 6614**). Below is an example of Stata syntax to add sex, year of birth, and ethnic group from the **xwavedat** file in the main survey. This file contains stable characteristics of individuals such as sex, year of birth, and ethnic group, which are typically collected only once when someone is interviewed or enumerated⁹ for the first time. As **xwavedat** includes all sample members ever enumerated in either Understanding Society or the BHPS, it includes more individuals than in the partnership history datafiles. To know more about this data file or other datafiles in the main survey see this [Main Survey User Guide section](#).

```
use phistory_long, clear
merge m:1 pidp using xwavedat, nogen keep(1 3) ///
keepus(racel_dv sex birthy)
save temp, replace
```

```
use phistory_long, clear
rename pidp kpidp
rename partner pidp
drop if pidp<0
merge m:1 pidp using xwavedat, nogen keep(1 3) ///
keepus(racel_dv sex birthy)
rename racel_dv p_racel_dv
rename sex p_sex
rename birthy p_birthy
rename pidp partner
rename kpidp pidp
keep pidp partner p_*
save partner, replace
```

```
use temp, clear
drop if partner<0
merge m:m pidp partner using partner
assert _m==3
drop _merge
save phistory_long_partnerinfo, replace
```

```
erase temp.dta
erase partner.dta
```

⁹ Being enumerated refers to an individual being recorded as part of a responding household. Not everyone who is enumerated is interviewed – that may be because they are not eligible (10 to 15-year-olds are eligible for youth interviews, and 16+ year-olds are eligible for adult interviews, but 0 to 9-year-olds are not interviewed) or they are eligible but did not participate in the interview for various reasons.

6. Data Access

These data are released through the UK Data Service. The End User Licence (EUL) version, SN 8473, can be found on the UK Data Service website:

<https://beta.ukdataservice.ac.uk/datacatalogue/studies/study?id=8473>.

7. Citation

The bibliographic reference for this study is as follows:

University of Essex, Institute for Social and Economic Research. (2026). Understanding Society: Marital and Cohabitation Histories, 1991-2024. [data collection]. 7th Edition. UK Data Service. SN: 8473, <http://doi.org/10.5255/UKDA-SN-8473-7>.

Citing this User Guide

Institute for Social and Economic Research. (2026). Understanding Society: Marital and Cohabitation Histories, 1991-2024 User Guide, Version 3, March 2026, Colchester: University of Essex.

All works which use or refer to these materials should acknowledge these sources by means of bibliographic citation. To ensure that such source attributions are captured for bibliographic indexes, citations must appear in footnotes or in the reference section of publications.

Acknowledgements

Understanding Society is an initiative 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.

If you have any questions or would like to provide feedback please email
usersupport@understandingsociety.ac.uk

Appendix A

Table A1: Description of the variables from the source data files used to create the variables in this datafile			
UKHLS Waves 1 – current wave			
Variable name	Waves	Data File	Survey
pidp, hhorig, pid	-	xwavedat	
pidp, w_ppid, w_marstat, w_mastat_dv, w_cohab_dv, w_istrtdaty, w_istrtdatm	All Waves	w_indresp, w = a, ..., 0	SN 6614
pidp, w_nmar, w_cohab_dv, w_lcmcb4, w_lcmcbm, w_lcmcoh, w_lcmspy4, w_lcmspm, w_lcsby4, w_lcsbm	Initial Wave (Wave 1 for GPS & EMBS, Wave 6 for IEMBS)	w_indresp, w = a, f	
pidp, w_marno, w_lmarm, w_lmarmy4, w_lmcoh, w_lmcbm, w_lmcb4, w_lmend, w_lmwwm, w_lmwwy4, w_lmdvm, w_lmdvy4, w_lmspm, w_lmspy4		w_marriage, w = a, f	
pidp, w_cohabno, w_lcsbm, w_lcsby4, w_lcsem, w_lcsey4		w_cohab, w = a, f	
pidp, w_statcy4*, w_statcm*, w_lmcb4, w_lmcbm, w_coh1by, w_coh1bm, w_lmar1y, w_lmar1m	Annual event history, Waves 2-15	w_indresp, w = b, ..., 0	
BHPS Waves 1-18			
Variable name	Waves	Data File	Survey
pidp, bw_marstat	1-18	bw_indresp (w = a, ..., r)	SN 6614
pp→pid, partner, separation, start_date, stop_date, marital, cohabitation		family	SN 5629

Appendix B:

Description of the questions including routing about marriage, civil partnership and cohabitation in UKHLS Wave 1 onwards.

Wave 1 for GPS & EMBS & Wave 6 for IEMB sample

W_INDRESP file:

LCMARM, LCMARY4: Start month and year of current marriage

LMCOH: Did they cohabit with their spouse before marriage?

LMCBM, LMCBY4: Start month and year of cohabitation before marriage [If LCMCOH=1]

NMAR: “Altogether, how many times have you been married or in a civil partnership?”

If NMAR>1, then ask about past marriages [W_MARRIAGE file]:								
MARNO	LMARM, LMARY4	LMCOH	LMCBM, LMCBY4	LMEND	LMWWM, LMWWY4	LSPWWD	LMDVM, LMDVY4	LMSPM, LMSPY4
Marriage no	Start month and year of current marriage	Did they cohabit with their spouse before marriage?	Start month and year of cohabitation before marriage [If LMCOH=1]	Did that marriage/CP end in divorce/dissolve, widowhood or something else?	Month and year partner died [if LMEND=2]	Were you and your partner living together at that time? [if LMEND=2]	Month and year marriage ended in divorce / CP dissolved [If LMEND=1]	In what month and year did you stop living together? [If LMEnd = 1 97] LSpWwd =2]

LCOH: “As you know some couples live together without actually getting married. Aside from any marriages or civil partnerships you have just told me about, have you ever lived with someone as a couple for three months or more?”

LNCOH: “In all, how many partners have you ever lived with for three months or more outside of marriage?” [If LCOH=1]

If LCOH=1, then ask about all cohabitations [W_COHAB file]:		
COHABNO	LCSBM, LCSBY4	LCSEM, LCSEY4
Cohabitation number	Start month and year of cohabitation	End month and year of cohabitation

Wave 2 onwards, for continuing sample members, any sample

W_INDRESP file:

Changes in marital status since last interview

MSTATSAM: When we last interviewed you, your legal marital status was [ff_marstat] . Has your legal marital status changed at all since [ff_intdate]

- 1: Yes
- 2: No
- 3: Last wave marital status incorrect

LWWRONG: Can you tell me what your legal marital status was last time we interviewed you on [ff_IntDate] [**IF MSTATSAM=3**]

MSTATSAMN: Has your legal marital status changed at all since [ff_IntDate] ? [**IF MSTATSAM=3**]

- 1: Yes
- 2: No

MSTATCH1: What did your legal marital status become with this change? Were you... [**MSTATSAM=1 OR MSTATSAMN=1**]

STATCM1, STATCY41: Month and year of marital status change

[If **MSTATCH1** = 5 or 8]:

DIVCHK: You say you became divorced {if MStatCh = 5} / your civil partnership was dissolved {if MStatCh = 8} . Is the date you have given me the

date your divorce {if MStatCh = 5} / dissolved partnership {if MStatCh = 8} became final or the date you separated and stopped living with your husband/wife/partner?

- 1: Date became final
- 2: Date stopped living together

DIVFIN: Is your divorce {if MStatCh = 5} / dissolved civil partnership {if MStatCh = 8} legally final?

1: Yes

2: No

DVM, DVY4: In what month and year did your divorce or annulment {if MStatCh = 5} / dissolved civil partnership {if MStatCh = 8} become final?

CMLSTAT1: And has your legal marital status changed again since then?

1: Yes

2: No

LOOP UNTIL CMLSTATn=2

Cohabitations since last interview

COHAB: If any periods of cohabitation lasting three months or more since last interview [including cohabitation with current husband/wife/partner]

COHABN: Number of cohabitation lasting three months or more since last interview, including periods of cohabitation with current husband/wife/partner

LMCBMn, LMCBY4n: Start month and year of cohabitation no. n

CURRPARTn: Check if this is the current husband/wife /partner.

LMSPM, LMSPY4n: End month and year of cohabitation no. n [If **CURRPARTn** =2]

Wave 2 onwards, new entrants, any sample

LCOHNPI: ever cohabited without marrying

COH1BM, COH1BY4: first cohabitation start month and year

COH1MR: did marry the person cohabiting with?

COH1EM, COH1EY4: first cohabitation end month and year

NMAR: number of times married

LMAR1M, LMAR1Y4: Start month and year of first marriage

