Understanding Society: Transition to Adulthood Living Arrangements Event History Datafile, User Guide

Eleni Karagiannaki

Centre for Analysis of Social Exclusion, LSE

March 2024

Abstract: This user guide describes the content and the methodology underlying the construction of the *Transition to Adulthood Living Arrangements Event History Datafile (TTA-LA)* – a dataset that was constructed based on the British Household Panel Survey (BHPS) and the Understanding Society the UK Household Longitudinal Study (UKHLS) and contains detailed data that capture the living arrangement histories of different cohorts of 16-year-olds as they transition from childhood to adulthood for different cohorts of children in the UK.

1. Introduction

This user guide provides information of a dataset which is constructed based on the British Household Panel Survey (BHPS) and the Understanding Society the UK Household Longitudinal Study (UKHLS) and contains detailed data that capture the living arrangement histories of different cohorts of 16-year-olds as they transition from childhood to adulthood for different cohorts of children in the UK. The creation of this dataset supplements "The Transition To Adulthood (TTA)" dataset which was constructed as part of an Understanding Society Fellowship project "*Transition to adulthood in the UK in an intergenerational context*". This project had the overarching aim to provide information to enable a better understanding of how the transition from adolescence into adulthood in the UK has changed over time and whether and to what extent intergenerational family

Corresponding author: Eleni Karagiannaki, Centre for Analysis of Social Exclusion, London School of Economics, Houghton Street, London, WC2A 2AE, e.karagiannaki@lse.ac.uk

context plays an increasingly more important role in determining the types of pathways followed in the transition to adulthood.

As in the original *TTA* dataset the sample included in the *Transition To Adulthood – Living Arrangements event histories dataset (TTA-LA)* includes 10,778 children of BHPS and UKHLS sample members who had turned 16 years of age over subsequent BHPS and UKHLS waves. Following this sample longitudinally throughout the annual BHPS and UKHLS panel data and using supplementary data from the Understanding Society Marital and Cohabitation history file, I compile information about the marital and partnership histories, the living with parents status histories, as well as the fertility and living with children histories of each sample member from 16 years of age up to the maximum age at which each sample member is observed in the surveys. In addition, I constructed a set of variables to capture episodes in which sample members lived with others individuals beyond their parents and partners/spouses as well as some key variables which capture children's socio-economic background characteristics.

The dataset is in discrete-time format (long person-time), with one record per one-month time interval. The sample contributes 722,774 monthly observations. The key variable in the dataset are a series of variables indicating the state of each sample member for each domain of interest, a spell indicator for the relevant state identified by each status variable, a time of episode indicator (t), the duration of each episode along with the corresponding censoring indicators (in each monthly interval identification). This user guide and the associated STATA do-files can be used to replicate the dataset. The following documentation gives a description of the input variables and the consequent preparation of the output variables included in the dataset.

2. UKHLS and BHPS

The data underlying the construction of the dataset includes the 18 waves of the BHPS and 11 waves of the UKHLS that were available at the time this study was conducted (University of Essex, 2021) [1]. The BHPS was an annual survey which run continuously from 1991 to 2008 (with a total

[1] Understanding Society: Waves 1-11, 2009-2020 and Harmonised BHPS: Waves 1-18, 1991-2009 (SN 6614)

of 18 waves). The initial 1991 BHPS sample included a nationally representative sample of about 5,500 households (around 10,000 individuals). Several extension samples were added over the years to the initial BHPS sample.[2] The total sample size for the BHPS including the extension samples was around 10,000 households across the UK.[3] The individuals in the sampled households (including new members joining sampled households) were re-interviewed in successive waves even if they split-off from their original households to form new households (in which case all adult members of these new households were also interviewed). Children of BHPS sample members were interviewed once they reached the age of 16 and they were also followed over subsequent survey sweeps using the same rules as the original sample members.[4]

BHPS came to an end in 2008 and was replaced by *Understanding Society* in 2009, with former BHPS sample members were invited to participate in the UKHLS from Wave 2. Similarly, to the BHPS, the *Understanding Society* is a nationally representative panel survey collecting annual information about individuals and households in the UK. Compared to BHPS, the UKLHS has a much larger sample size (for example in wave 1, over 50,000 adults in 30,000 households completed face-to-face interviews). The Understanding Society main survey sample consists of three sub-samples: a new large General Population Sample (GPS) (around 26,000 households at wave 1); the Ethnic Minority Boost Sample (EMBS) (approximately 4,000 households in wave 1); the former BHPS sample (approximately 8,000 households); and the Immigrant and Ethnic Minority Boost Sample (IEMBS) added in wave 6 (around 2,400 households).[5]

---

[2] These samples included 1,500 households in each of Scotland and Wales which were added in 1999 and a sample of 2,000 households was added in Northern Ireland in 2001. In addition, from Wave 7 to Wave 11 the BHPS began providing data for the United Kingdom European Community Household Panel (ECHP) incorporating a sub-sample of the original UKECHP (for further information on BHPS sub-samples see Taylor et al, 2018).

[3] For further information for the various BHPS sub-samples Taylor, Marcia Freed (ed). with John Brice, Nick Buck and Elaine Prentice-Lane (2018) British Household Panel Survey User Manual Volume A: Introduction, Technical Report and Appendices. Colchester: University of Essex. https://www.iser.essex.ac.uk/wp-content/uploads/bhps/documentation/volumes/5151userguide_vola.pdf

[4] The sample - Institute for Social and Economic Research (ISER) (essex.ac.uk)

[5] For further details about the UKHLS sample see Lynn and Knies, (2016).

3. Information on event history construction

The event histories included in the "Transition to adulthood: Living arrangement Histories" dataset fall into three main domains: 1) living with parents histories; 2) marital and cohabitation history; 3) the fertility and living with children domain; and the 4) living with others (beyond parents and partners/spouses) domain.

Marital and Cohabitation Histories, 1991-2019

The marital and cohabitation event history data were extracted from the Understanding Society marital and cohabitation history file.[6] This 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 10 (Nandi et al, 2021)[7]. As described in detail in Nandi et al (2021) for UKHLS samples, their marital and cohabitation histories were compiled from information about their partnership history and their current partnership status that collected in their initial (wave 1) interview along with information on changes since last interview that was collected in subsequent interviews. For respondents joining Understanding Society in subsequent waves (after wave 1), only partial information about their past partnership history was collected (Nandi et al, 2021). In BHPS, the marital and cohabitation histories were collected in waves 2, 11 and 12, respectively. As in Understanding Society these data are updated with information collected in the the panel annually. For BHPS sample members, information about their partnership history until 2008 (i.e. when they were interviewed as part of the BHPS) has been extracted from the single partnership history file created by Pronzato (2009). For those BHPS sample members who continued into Understanding Society (from onwards 2010), their histories up 2008 from this file, were combined with that collected in Understanding Society. The marital and cohabitation history dataset is supplied in both long and wide formats. For constructing the transition to adulthood

[6] University of Essex, Institute for Social and Economic Research. (2021). Understanding Society: Marital and Cohabitation Histories, 1991-2019. [data collection]. 3rd Edition. UK Data Service. SN: 8473, http://doi.org/10.5255/UKDA-SN-8473-3

[7] Nandi, Alita; Bellamio, Sara; Smith, Robert and Menon, Seetha. (2021). Understanding Society: Marital and Cohabitation Histories, 1991-2019 User Guide. Colchester: University of Essex

datafile I use the wide format dataset. In this file each marital or cohabitation spell along with its start and end dates is captured by different variables. The variable names for each of the spells have a number suffix that represents the spell number.

Living with parents histories

Histories of movements in and out of parental home were derived based on information collected in the panel mainly as part of the household grid. Based on this information I determine for each sample member when they moved in and out of their parental home during their life up to their last interview in the panel. In particular, we can observe when they first moved out of the parental home, as well as subsequent moves back into and out of the parental home. From these we can also determine how many moves each respondent have had during the panel. The maximum number of movements for our sample member are 8. The variables used in the derivation of living with parents histories for each of the BHPS and UKHLS sample members are listed in Table 1.

The process of deriving the full history of movements in and out of parental home for each respondent proceeds in several steps. First I derive the variable indicating for each respondent whether he or she lived with at least one of his parents (natural, step or adoptive) in each wave. The derivation of this variable uses information about the father or mother's personal identifier or the personal number variable whichever is available for each respondent (fnpid_bh/mnpid_bh, fsnpid/msnpid, fnspno/mnspno). Following each individual through the panel, I can then determine when/if he/she moved out the parental home as well as all subsequent movements back into and out of the parental home. Note that for respondents whose parents did not live together moves between parents are not captured as movement out or back into the parental home.

Then and in order to identify each move and the date that this has occurred I rely on information about the date at which each individual has been interviewed in each wave, the individual interview outcome in each wave (which codes household members who moved out of the household with ivfio=12) and the year and the month that the individual moved out or joined the household. The latter information is available in all BHPS waves as well as the first seven waves of the UKHLS. The variables that contain the information about the year and the month when an individual left the household are respectively the variables moveyr and movemnth while the variable capturing the

month and year at which an individual joined a household are the jnyear and the jnmonth respectiely.

For cases where information about the year and the month that an individual left or joined a household are not missing (i.e. the variable moveyr, movemnth jnyear, jnmonth) the date of the move (identified by end of the spell variable) is set to the date identified by these variables. For moves out of parental home the relevant variables are the moveyr and movemth while for moves back into the parental home the jnyear and jnmonth variables. For individuals with missing data on either the year and/or the month of the move I set the spell end date at the midpoint between the date of the current interview and the previous interview. When no individual interview is provided the household interview date (of the original household) is used.

Using these variables, I identify the number of spells that the individual lived in and out of the parental home as well the start and end date of each spell (with the start date of each spell set as the end of the previous spell). For all individuals the start date of the first spell (which involves the living with parent status) is set at the birth date. Note that this variable does not capture either moves between parents not living together and or moves that involve moves between households that do include parents. The former moves are captured by the detailed living with parents status variable which disaggregates whether the respondent lives with both parents, father only or the mother only.

Table 1. Variables used in the living with parents histories

Variable name/description	Survey/wave availability
hgbiof/hgbiom: pno of biological father/mother	BHPS waves 1-18 and UKHLS waves 1-11
fnpid_bh/mnpid_bh: pid of biological father/mother	BHPS waves 1-18
fsnpid/msnpid: Nat/step/adopt father/mother: Cross-wave person identifier (PIDP) of natural/adoptive/step father. PIDP	UKHLS waves 1-11
fnspid_bh	BHPS x-wave data (only youth)
fnspno/mnspno: person number of father/mother (nat/step)	BHPS 4-18, UKHLS 1-11
npns_dv/npns_dv: Number of respondent's biological/step/adoptive parents in hh	UKHLS waves1-11
livpar: respondent is living at home with parent/s	UKHLS waves 3-11
ivfio: individual interview outcome	BHPS wave 1-18 and UKHLS waves 1-11
moveyr: year left household	BHPS waves 1-18, UKHLS waves 1-11 but available 1-7
movemnth: month left household	BHPS waves 1-18, UKHLS waves 1-11 but available 1-7
jnyear: year joined the household	BHPS waves 1-18, UKHLS waves 1-11
jnyear: year joined the household – new entrants	BHPS waves 1-18, UKHLS waves 1-11
jnmnth: month joined household – new entrants	BHPS waves 1-18, UKHLS waves 1-11
intdatm_dv/ intdaty_dv: interview date: month/year derived	UKHLS waves 1-11,
Intdatey: household interview year	BHPS waves 2-18 and UKHLS waves 1-11
Intdatem: household interview month	BHPS waves 1-18 and UKHLS waves 1-11

Note: For details about these variables see https://www.understandingsociety.ac.uk/documentation/mainstage/dataset-documentation

Fertility and living with children

Fertility and living with children history variables were derived based on information collected in the panel as part of the household grid and main interview questionnaire. The key variables included in the dataset to capture respondents' fertility and living with children history are a variable indicating whether respondents live with children (either natural or stepchildren), a new birth indicators variable (i.e. a variable indicating whether the respondents had a new baby in each time period), and indicators of the number of respondents' children (overall and of different ages). Histories are constructed for each sample member from age 16 up to their last interview in the panel.

The variables used in the derivation of the living with children histories are listed in Table 2. To derive the fertility and living with children history for each respondent I first derive the variable indicating the number of children who live in the same household as the sample member in each wave. The derivation of this variable is based on linking each respondent to their children who live in the household based on the variables indicating children's father or mother's personal identifier (fnpid_bh/mnpid_bh, fsnpid/msnpid). Following each respondent through the panel, I can then determine whether there has been a change between waves in the number of children who live in the same household as the respondent and to identify the reason behind the change i.e. whether the change was due to new birth, a move with a partner/spouse with children, or a move out of partner/spouse with children. Based on this information I then construct a variable to indicate the date that each change has occurred. For new births (which are captured based on the variable indicating the reason why a new household member joined the household i.e. ynew=1 in UKHLS and newhy=1 in BHPS) the date is based on information about the date on which the new joiner joined the household (i.e. jnmnth and jnyearr) or in case that this is missing on the date of birth of each new child. If the change in the number of children was associated with partnership formation or dissolution the date is set at the date that the partnership spell has started or ended (based on information from the marital and partnership history file). When either the dates are missing the date is set at the date that the marital/partnership spell ended based on information from the marital and partnership history file. For all sample members the start date of the first spell is set at the birth date.

Table 2. Variables used in the living with children history variables

Variable name/description	Survey/wave availability
fnpid_bh/mnpid_bh: pid of biological father/mother	BHPS waves 1-18
fsnpid/msnpid: Nat/step/adopt father/mother: Cross-wave person identifier (PIDP) of natural/adoptive/step father PIDP	UKHLS waves 1-11
nchild_dv: Number of own children in household	BHPS waves 1-18; UKHLS waves1-11
ivfio: individual interview outcome	BHPS wave 1-18 and UKHLS waves 1-11
moveyr: year left household	BHPS waves 1-18, UKHLS waves 1-11 but available 1-7
movemnth: month left household	BHPS waves 1-18, UKHLS waves 1-11 but available 1-7
newhy: unlisted joiners	BHPS waves 1-18
ynew: New joiner reason	UKHLS: waves 1-11
jnyear: year joined the household	BHPS waves 1-18, UKHLS waves 1-11
jnyear: year joined the household – new entrants	BHPS waves 1-18, UKHLS waves 1-11
jnmnth: month joined household – new entrants	BHPS waves 1-18, UKHLS waves 1-11
intdatm_dv/ intdaty_dv: interview date: month/year derived	UKHLS waves 1-11,
intdatey: household interview year	BHPS waves 2-18 and UKHLS waves 1-11
intdatem: household interview month	BHPS waves 1-18 and UKHLS waves 1-11

Note: For details about these variables see https://www.understandingsociety.ac.uk/documentation/mainstage/dataset-documentation

Living with others

In addition to the above, the dataset also includes a set of variables that capture episodes that the respondents lived with people beyond their spouses and parents. To derive these variables I first derived a variable which indicated whether the sample member lived with other people beyond their parents or spouse. Two variants of this variable has been constructed. The first takes the value of 1 if the respondent lived with other people beyond their parents, and partner or spouse and 0 otherwise (i.e. if they lived alone or with either parents or spouse/partner). The second variants takes the value of 1 if the respondent lived with other people beyond their parents and their partner or spouse (i.e. under this definition respondents may live with their parents or spouse and still be assigned as living with others) and 0 if they did not live with other people (or with either parents or spouse/partner). The variable included in the dataset is based on the first definition. After defining the variable which indicates the living others status of each sample member, a variable is constructed to indicate whether there has been a change in status between waves. Based on this a spell indicator is derived to indicate the spell number of each living with other spell. Then for each living with other spell, a variable is constructed to indicate its start and the end date (based on interview date in which the change occurred).

4. Sample

The sample included in the dataset consists of all individuals who turn 16 years of age over the course of the BHPS and UKHLS panels. The sample includes overall 10,777 individuals (after excluding children who did not live with their parents at age 16, the ECHP samples as well temporary sample members who did not provide interviews). We use the yearly panel data from all available BHPS and Understanding Society waves to follow these samples of 16 years olds and construct socioeconomic sequences of life events that this sample has experienced from age 16 to the maximum age at which each cohort can be observed based on methodology described in the previous section.

Overall, 18 cohorts can be constructed based on BHPS and 11 based on Understanding Society. Table 1 shows the sample size of each cohort and the number of cohort sample members who give

full or proxy interviews at each wave. In total among the 28 cohorts, 20 can be observed up to age 25, 17 up to 27 and 15 up 30.

As is common in longitudinal surveys there is a high degree of attrition in the samples. The magnitude of attrition can be seen in Table 2 which gives for each cohort the proportion of sample members who give full or proxy interviews as a proportion of all who gave interviews at age 16. Among the 5,444 individuals who belong to cohorts that can be observed at age 25 (which includes all cohorts who turned 16 years of age in waves 1-18 of BHPS and in waves 1 and 2 of Understanding Society) 3,131 are lost at some point before their age 25 interview (or 57.5% of 5,444). In general attrition is higher in Understanding Society than in the BHPS. Moreover, in BHPS attrition is concentrated in later waves while in Understanding Society the problem is more acute in the earlier sweeps and especially in waves 1 and 2. So, for example while 10% of the BHPS cohort 1 sample is lost between age 16 and 17 the respective proportion among the cohort the cohort of 16 years old identified by Understanding Society wave 1 who are lost between their age 16 and age 17 interviews is 24%.

Attrition often correlates with the occurrence of major life course events, such as leaving the parental home, partnership formation or dissolution (Pelikh, 2019; Bayrakdar et al, 2018). As these types of life-course events are more likely to occur in early adulthood there is an increased risk of attrition for this age group. Indeed, as shown by Lynn and Borkowska (2018) attrition in *Understanding Society* is slightly higher among people under the age of 30 years old. As mentioned above among the 5,444 individuals belonging to cohorts that can be observed up to 25 years of age, 3,131 cases (57.7% of the total sample) are lost at some point before their age 25 interview. The majority of the individuals who left the panel were living with their parents in the year prior leaving the panel (2,656 cases or 85% of all those left). For 749 of cases the exit from the panel was associated with a move out of the parental home (24% of all lost). For another 702 cases this was because of non-interviews while for 1,214 (51% of all those left the panel) due to the entire origin household (i.e. the household in which they belonged to in previous waves) was lost. For individuals who left the panel because the entire origin household was lost we cannot determine whether they have left the parental home while for individuals who were living with their parents in the wave prior the attrition occurred we can safely assume that this involved a move out of parental home.

Table 3. Number of each cohort giving full or proxy interviews at each age

Wave	BHPS 1 (1991)	BHPS2 (1992)	BHPS3 (1993)	BHPS4 (1994)	BHPS5 (1995)	BHPS6 (1996)	BHPS7 (1997)	BHPS8 (1998)	BHPS9 (1999)	BHPS1 0 (2000)	BHPS1 1 (2001)	BHPS1 2 (2002)	BHPS1 3 (2003)	BHPS1 4 (2004)	BHPS1 5 (2005)	BHPS1 6 (2006)	BHPS1 7 (2007)	BHPS1 8 (2008)	USOC W1 (2009/1 0)	USOC W2 (2010/ 11)	USOC W3 (2011/ 12)	USOC W4 (2012/ 13)	USOC W5 (2013/ 14	USOC W6 (2014/ 15)	USOC W7 (2015/ 16)	USOC W8 (2016 /17)	USOC W9 (2017 /18)	USOC W10 (2018 /19)	USOC W11 (2019 /20)
Age/Year of birth	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993/ 94	1994/ 95	1995 / 96	1996/ 9 7	1997/ 98	1998/ 99	1999/ 00	2000/ 01	2001/ 02	2002/ 03	2003/ 04
16	155	148	142	127	135	163	135	122	222	180	283	254	251	243	252	253	274	213	895	967	782	744	729	651	636	537	513	384	368
17	143	132	134	120	130	155	126	113	204	162	245	214	223	220	237	229	249		677	747	664	641	564	513	501	413	411	298	368
18	128	128	122	118	118	144	115	106	182	145	219	194	208	199	211	210		162	561	682	596	509	510	455	421	378	374	682	
19	122	119	125	113	116	137	109	101	170	134	193	179	187	184	191		168	152	492	608	483	454	411	386	385	323	1298		
20	119	118	122	108	114	127	108	94	166	130	187	172	168	171		152	155	129	442	492	410	378	363	352	330	1651			
21	115	116	117	102	109	118	90	85	153	127	184	155	158		139	141	136	116	356	430	369	328	345	322	2273				
22	115	110	114	98	95	114	92	83	139	123	164	139		126	109	128	122	98	356	382	319	305	312	2679					
23	112	104	108	96	94	102	83	76	134	116	153		100	112	94	112	107	82	314	341	288	280	3234						
24	104	94	98	90	90	100	78	75	122	108		100	92	100	85	95	93	77	262	313	287	3639							
25	101	84	97	81	90	95	71	69	108		110	84	78	95	72	78	83	73	248	273	4198								
26	100	79	89	79	90	92	72	64		76	91	78	72	77	68	77	72	68	229	5235									
27	94	75	89	75	88	93	67		77	65	96	74	62	81	63	72	67	59	4832										
28	94	74	85	76	79	89		49	74	55	81	68	61	71	51	66	63	1229											
29	90	71	82	68	79		50	48	59	51	61	61	53	66	50	64	1589												
30	88	65	80	68		73	44	42	59	51	59	55	51	68	45	1677													
31	86	65	73		58	64	39	42	56	46	55	53	47	61	1667														
32	82	61		53	53	63	36	40	52	42	56	47	44	1874															
33	81		61	49	50	62	34	37	50	40	55	51	1855																
34		50	55	48	48	52	34	35	48	38	49	1978																	
35	60	48	52	48	48	43	30	33	41	36	2341																		
36	57	42	44	45	47	42	31	35	39	1725																			
37	47	43	39	44	42	43	23	32	2155																				
38	48	41	37	42	40	44	25	1381																					
39	44	38	36	40	39	34	1492																						
40	43	39	36	36	36	2049																							
41	43	37	33	39	1888																								
42	42	33	37	1863																									
43	36	33	2107																										
44	39	2047																											
	2388																												

Note: Each cell includes sample members who did not leave the panel at t-1 (may include individuals who did not provide interviews at t-1).

Table 4. Proportion of each cohort providing full or proxy interviews at each age

Wave	BHPS 1 (1991)	BHPS 2 (1992)	BHPS 3 (1993)	BHPS 4 (1994)	BHPS 5 (1995)	BHPS 6 (1996)	BHPS 7 (1997)	BHPS 8 (1998)	BHPS 9 (1999)	BHPS 10 (2000)	BHPS 11 (2001)	BHPS 12 (2002)	BHPS 13 (2003)	BHPS 14 (2004)	BHPS 15 (2005)	BHPS 16 (2006)	BHPS 17 (2007)	BHPS 18 (2008)	USOC W1 (2009/10)	USOC W2 (2010/11)	USOC W3 (2011/12)	USOC W4 (2012/13)	USOC W5 (2013/14	USOC W6 (2014/15)	USOC W7 (2015/16)	USOC W8 (2016/17)	USOC W9 (2017/18)	USOC W10 (2018/19)	USOC W11 (2019/20)
Age/Year of birth	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993/94	1994/95	1995/96	1996/97	1997/98	1998/99	1999/00	2000/01	2001/02	2002/03	2003/04
16	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
17	0.92	0.89	0.94	0.94	0.96	0.95	0.93	0.93	0.92	0.90	0.87	0.84	0.89	0.91	0.94	0.91	0.91		0.76	0.77	0.85	0.86	0.77	0.79	0.79	0.77	0.80	0.78	
18	0.83	0.86	0.86	0.93	0.87	0.88	0.85	0.87	0.82	0.81	0.77	0.76	0.83	0.82	0.84	0.83		0.76	0.63	0.71	0.76	0.68	0.70	0.70	0.66	0.70	0.73		
19	0.79	0.80	0.88	0.89	0.86	0.84	0.81	0.83	0.77	0.74	0.68	0.70	0.75	0.76	0.76		0.61	0.71	0.55	0.63	0.62	0.61	0.56	0.59	0.61	0.60			
20	0.77	0.80	0.86	0.85	0.84	0.78	0.80	0.77	0.75	0.72	0.66	0.68	0.67	0.70		0.60	0.57	0.61	0.49	0.51	0.52	0.51	0.50	0.54	0.52				
21	0.74	0.78	0.82	0.80	0.81	0.72	0.67	0.70	0.69	0.71	0.65	0.61	0.63		0.55	0.56	0.50	0.54	0.40	0.44	0.47	0.44	0.47	0.49					
22	0.74	0.74	0.80	0.77	0.70	0.70	0.68	0.68	0.63	0.68	0.58	0.55		0.52	0.43	0.51	0.45	0.46	0.40	0.40	0.41	0.41	0.43						
23	0.72	0.70	0.76	0.76	0.70	0.63	0.61	0.62	0.60	0.64	0.54		0.40	0.46	0.37	0.44	0.39	0.38	0.35	0.35	0.37	0.38							
24	0.67	0.64	0.69	0.71	0.67	0.61	0.58	0.61	0.55	0.60		0.39	0.37	0.41	0.34	0.38	0.34	0.36	0.29	0.32	0.37								
25	0.65	0.57	0.68	0.64	0.67	0.58	0.53	0.57	0.49		0.39	0.33	0.31	0.39	0.29	0.31	0.30	0.34	0.28	0.28									
26	0.65	0.53	0.63	0.62	0.67	0.56	0.53	0.52		0.42	0.32	0.31	0.29	0.32	0.27	0.30	0.26	0.32	0.26										
27	0.61	0.51	0.63	0.59	0.65	0.57	0.50		0.35	0.36	0.34	0.29	0.25	0.33	0.25	0.28	0.24	0.28											
28	0.61	0.50	0.60	0.60	0.59	0.55		0.40	0.33	0.31	0.29	0.27	0.24	0.29	0.20	0.26	0.23												
29	0.58	0.48	0.58	0.54	0.59		0.37	0.39	0.27	0.28	0.22	0.24	0.21	0.27	0.20	0.25													
30	0.57	0.44	0.56	0.54		0.45	0.33	0.34	0.27	0.28	0.21	0.22	0.20	0.28	0.18														
31	0.55	0.44	0.51		0.43	0.39	0.29	0.34	0.25	0.26	0.19	0.21	0.19	0.25															
32	0.53	0.41		0.42	0.39	0.39	0.27	0.33	0.23	0.23	0.20	0.21	0.18																
33	0.52		0.43	0.39	0.37	0.38	0.25	0.30	0.23	0.22	0.19	0.21																	
34		0.34	0.39	0.38	0.36	0.32	0.25	0.29	0.22	0.21	0.17																		
35	0.39	0.32	0.37	0.38	0.36	0.26	0.22	0.27	0.18	0.20																			
36	0.37	0.28	0.31	0.35	0.35	0.26	0.23	0.29	0.18																				
37	0.30	0.29	0.27	0.35	0.31	0.26	0.17	0.26																					
38	0.31	0.28	0.26	0.33	0.30	0.27	0.19																						
39	0.28	0.26	0.25	0.31	0.29	0.21																							
40	0.28	0.26	0.25	0.28	0.27																								
41	0.28	0.25	0.23	0.31																									
42	0.27	0.22	0.26																										
43	0.23	0.22																											
44	0.25																												

Note: Each cell includes sample members who did not leave the panel at t-1 (may include individuals who did not provide interviews at t-1).

5. Combining the three event history files and restructuring the dataset into a discrete-time monthly format

The steps to construct the file:

- Merge and append the various files from all available waves BHPS and Understanding Society waves to generate a single file with all available data from all BHPS and Understanding Society waves
- Construct the relevant samples and deriving the living with parents event history variables
- Merge the relevant variables from the marital and cohabitation history files to create a dataset with all the relevant data in long format with one record per individual for each wave observed in the surveys.
- Restructure (expand) the data in order to organise the datafile in discrete-time monthly format i.e., with one record per one-month time interval from the date when each sample member became eligible for their first interview as adult respondents till their last interview.
- Use the relevant spell variables for each of the domains (living with parents, marital and cohabitation status, living with children, and living with other) along with the start and end date of each spell to construct the relevant status variables to indicate the various living arrangements statuses (i.e. the marital/cohabitation status, the living with parents status, the living with children status, and the living with other status) of each individual for each month.
- In addition to the event history data, the dataset includes a set of baseline characteristics of the respondents. These include:
    1. the equivalised household disposable income when the respondent was 16 years of age;
    2. a set of derived categorical variables indicating the social class (NS-SEC) of respondents' parents (for mother and father separately as well as a combined parental NSSEC variable indicating father's or mother's social class whichever was higher);
    3. a set of derived categorical variables indicating the education of respondents' parents when the respondent was 16 years of age (for mother and father separately as well as combined parents' variable indicating father's or mother's education whichever was higher).

In total, the sample of 10,778 respondents contribute 722,774 monthly observations.

6. Data-file structure

As mentioned above, the dataset is organised in discrete-time format (long person-time), with one record per one-month time interval spent in each state. For each domain of interest, the dataset includes a variable indicating the status of the individual in the relevant domain, a spell indicator (i.e. indicating the spell number of each episode), the time of the episode indicator (indicating the time elapse since the beginning of each episode), the duration of each episode (indicating the total duration of each episode) along with the corresponding censoring indicators (in each monthly interval identification). The variables included to capture the living with parents, the marital and cohabitation, the living with children and the living with other histories, along with their coding and explanatory notes are shown in Tables 5, 6, 7 and 8 respectively.

Living with parents history: The 10,778 sample members contribute 11,698 movements out of the parental home and 4,430 moves back into the parental home (16,128 moves in total). 6,968 individuals had only one spell (i.e., they lived with their parents throughout the period observed in the panel), 2,989 had two spells (they left the parental home once and they did not move back over the period observed in the panel), 278 had three spells (left parental home and moved back once over the observation period), 452 had four spells (left parental home, moved back and left again) and 91 had five or more living/not living with their parents spells. For the "detailed living with parents" status indicator which disaggregates the living with parents into whether the respondents "live with both parents" "live with the mother only" and "with the father only", there are 16,963 spells. 6,574 individuals had only one spell (i.e., they lived with the same parents or with both parents as in their age 16 interview throughout the panel), 3,077 had two spells, 517 had three spells, 469 had four spells and 141 had five or more spells.

Partnership history: The sample members have 11,425 spells not living with partner or spouse and 2,409 spells living with a partner (13,834 moves in total). 9,094 individuals had only one spell (i.e. they were never married or in cohabitation over the period observed in the panel), 832 had two spells (i.e. one spell being single and one spell with a partner), 560 had three spells, 145 had four spells and 147 had five or more.

Fertility history: The sample members also had 1,092 spells of living with children (this include spells in which they start living with children because of new births among respondents who did not live with children before as well as spells associated with moving in with a partner who had children of their own), 10,818 not living with children spells (11,925 spells in total). 9,736 individuals had only one spell (i.e.

they never lived with children during the observation window), 958 had two spells (i.e. one spell without and one spell with children) and 84 had three or more spells. Moreover, there were 1,797 new births in the panel among 1,067 respondents. 531 respondents had one new baby, 391 had two new babies, 105 three new babies, and 40 four or more.

Living with others: The 10,778 sample members, had 937 spells of living with other people beyond their parents and/or partners. Among the 803 who had at least one spell of living with other people 227 had two spells (one living with others and one that they did not), 470 had three spells and 112 more than four.

Table 5. Living with parents event history variables

Variable name	Description	Coding	
lwpstatus	Living with parents status	0=not living with parents 1=living with parents	
lwpstatusd	Living with parents status - detailed	0=not living with parents 1=living with both parents 2=living with father 3=living with mother only	
lwpstatus_flag	Flag of living with parent status variable	1. Ongoing spell 2. Censored 3. Spell ended 4. Spell ended but date unknown	
lwpspell	Living with parent spell identifier	Reset to 1 when pidp changes	
lwpevent	Living with parents transition indicator	0= no change in status 1= change in status (includes both movements in and out of parental home but not move between parents)	
lwpt	Month for each episode	Reset to 1 for each episode	
lwpd	Total duration of each episode (*months*)	N of months	
lwp_spellsum	Total number of spells for each respondent		
leftpar1m	Indicator of first transition out of parental home		
ageleftpar1m	Age of first transition out of parental home		

Table 6. Marital and cohabitation history variables

Variable name	Description	Coding	
marital_statusm	Marital of partnership status	1. Not married/in civil partnership or cohabitation 2. Married 3. In civil partnership 4. Cohabiting	
marital_spellm	Partnership status spell identifier	Reset to 1 when pidp changes	
marital_eventm	Marital/partnsership status transition indicator	0= no change in marital/partnership status 1= change in partnership status	
marital_tm	Month for each episode	Reset to 1 for each marital/partnership episode	
mariatl_dm	Total duration of each marital/partnership episode	N of months	
Marital_spellsum	Total number of partnership spells	N	
age_partn1m	Age first partnership/marriage or cohabitation		
partn1m	Indicator of first partnership event		
partn1d	Indicator of being in partnership		

Table 7. Fertility and living with children history variables

Variable name	Description	Coding	
lwcstatusm	Indicator of living with children	0. Not living with children 1. Living with children (own or step children)	
lwcspell	Living with children spell identifier	Reset to 1 when pidp changes	
lwcevent	Living with children status transition indicator	0= no change in living with children status 1= change in living with children status	
lwct	Month for each episode	Reset to 1 for each live with children episode	
lwcd	Total duration of each living with children episode	N of months	
lwc_spellsum	Total number of living with children spells for each respondent		
age_firstchild1m	Age of first living with children transition		
firstchild1m	Indicator of first living with children transition	Takes the value of 0 before the transition 1 at the month of the living with children transition and setting to missing after that	
newbirth	Indicator of new birth	0=no new birth 1=new birth (respondent had a new born)	
sumofnewbirth	Sum of new births within the panel		

Table 8. Living with others variables

Variable name	Description	Coding	
lwostatusm	Indicator of living with others individuals (beyond partner and parents)	0. Not living with others 1. Living with others (beyond parents and/or partner) * does not include respondents who live with others but also with parents or partners	
lwospellm	Living with others spell identifier	Reset to 1 when pidp changes	
lwoeventn	Living with others status transition indicator	0= no change in living with others status 1= change in living with others status	
lwot	Month for each episode	Reset to 1 for each live with others episode	
lwod	Total duration of each living with others episode	N of months	
lwo_spellsum	Total number of living with children spells for each respondent		

References

Bayrakdar, S, Coulter, R. Parents, local house prices, and leaving home in Britain. Popul Space Place. 2018; 24:e2087. https://doi.org/10.1002/psp.2087

Lynn, P., and Borkowska, M. (2018). "Some Indicators of Sample Representativeness and Attrition Bias for BHPS and Understanding Society" Understanding Society at the Institute for Social and Economic Research.

Lynn, P., Knies, G. (ed.) (2016) Understanding Society: the UK Household Longitudinal Study: Wave 1-5 Quality Profile. Colchester: University of Essex.

Nandi, A., Bellamio, S., Smith, R., and Menon, S. (2021). Understanding Society: Marital and Cohabitation Histories, 1991-2019 User Guide. Colchester: University of Essex

Pelikh, A.(2019) Birth cohorts in Understanding Society: a description, Understanding Society Working Paper 2019-13, Colchester: University of Essex

Pronzato, C. (2011). British Household Panel Survey Consolidated Marital, Cohabitation and Fertility Histories, 1991-2009. [data collection]. 3rd Edition. University of Essex. Institute for Social and Economic Research, [original data producer(s)]. UK Data Service. SN: 5629, http://dx.doi.org/10.5255/UKDA-SN-5629- Essex

Taylor, Marcia Freed (ed). with John Brice, Nick Buck and Elaine Prentice-Lane (2018) British Household Panel Survey User Manual Volume A: Introduction, Technical Report and Appendices. Colchester: University of Essex. https://www.iser.essex.ac.uk/wp-content/uploads/bhps/documentation/volumes/5151userguide_vola.pdf

Understanding Society: Waves 1-11, 2009-2020 and Harmonised BHPS: Waves 1-18, 1991-2009 (SN 6614)

University of Essex, Institute for Social and Economic Research. (2021). Understanding Society: Marital and Cohabitation Histories, 1991-2019. [data collection]. 3rd Edition. UK Data Service. SN: 8473, http://doi.org/10.5255/UKDA-SN-8473-3