

Understanding Society
The UK Household Longitudinal Study
Innovation Panel, Waves 1-17,

User Guide

Version 1, August 2025



**Economic
and Social
Research Council**

Table of Contents

1. Data access.....	1
1.1 Innovation Panel datasets accessible through the UK Data Service	1
1.2 Additional data collections	2
Life events	2
Household finances.....	2
Linked credit reference agency data	2
Time-use diaries.....	3
Social media data	3
Biomeasures, cognition, and wellbeing	3
1.3 Citation.....	5
Citing this User Guide	5
Why cite data?	5
1.4 Ethics	6
2. Study overview.....	7
2.1 Innovation Panel competition	7
2.2 Sample design	7
Design of Wave 1 Innovation Panel sample	7
Refreshment sample in Wave 4	8
Refreshment sample in Wave 7	8
Refreshment sample in Wave 10.....	8
Refreshment sample in Wave 11.....	8
Refreshment Sample in Wave 14.....	8
Sample for Wave 14 LAT survey	9
Sample design variables.....	9
2.3 Annual interview process.....	9
Interview modes: web, CAPI and telephone	9
2.4 Fieldwork procedures	10
2.5 Response outcomes	10
Using Innovation Panel response outcome data	10
Stata code	11
3. Data structure	12
3.1 Files released every wave	13
3.2 Variable naming conventions	15
3.3 Additional data files in some waves	16
3.4 Derived variables	17

Variable search.....	18
Identifiers and other useful variables	18
Occupation codes.....	19
3.5 Derived income variables.....	19
Top-coding of income variables	19
3.6 Fieldwork paradata	20
Call records	20
Address response form	20
Timings data files	21
Interviewer characteristics	21
Keystroke paradata	21
3.7 Known data issues.....	21
Cross-wave issues	21
Wave 1	21
Wave 2	22
Wave 3	22
Wave 5	23
Wave 6	26
Wave 7	26
Wave 11	26
Wave 13	26
Wave 14	27
Wave 15	27
Wave 16	27
Wave 17	27
3.8 Stata code for matching files	29
Example 1: Distributing household level information to individual level	29
Example 2: Summarising individual level information at the household level	29
Example 3: Matching individuals within a household	30
Example 4: Using the EGOALT file to create household composition variables	31
Example 5: Merging individual files across waves into long format	31
Example 6: Merging individual files across waves into wide format.....	32
3.9 Changes to employment variables from Wave 14 and Wave 17	33
4. Questionnaire	34
4.1 How to read the questionnaire.....	34
4.2 New studies and questionnaire content carried this wave	36
5. Weighting.....	37

5.1 Weights for analysis of Wave 14 and later	37
Creating your own longitudinal weight	38
5.2 Weights for analysis of waves before Wave 14	39
Selecting the correct weight for your analysis	39
5.3 Naming conventions for weighting variables	41
5.4 Technical details.....	43
Weights for Wave 14 onwards	43
Weights before Wave 14	44
6. Experiments, methodological studies and non-experimental new content carried in the Innovation Panel	57
Submit a proposal for an experiment or other new survey content.....	57
6.1. Overview of Innovation Panel experiments and studies.....	58
Procedural experiments.....	58
Questionnaire design experiments: General issues	63
Questionnaire design experiments: Specific topics.....	66
Non-experimental studies	75
7. Procedural experiments.....	77
7.1 Respondent incentives to encourage participation and web response	77
7.2 Conditional and unconditional incentives	95
7.3 Mixed mode experiments: telephone and face-to-face.....	95
7.4 Mixed mode experiments: web and face-to-face	98
7.5 Paper vs CASI self-completion	104
7.6 Advance materials: letters vs. cards	106
7.7 Advance materials: content of advance letters	106
7.8 Targeted advance letters	107
7.9 Number of mailings between interviews.....	108
7.10 Different ways of asking respondents to register on participant website	108
7.11 Improving interviewer observations about characteristics of address	109
7.12 Effect of content of re-issue letter on refusal conversion.....	109
7.13 Early bird scheduling: encouraging respondents to call interviewer to schedule appointment	110
7.14 Targeted weekday invitation emails.....	110
7.15 Using prospect theory in wording of advance letters	112
7.16 Invitation letters for mixed mode survey	113
7.17 Spending Study 2	114
7.18 Invitation to complete pre-interview blood pressure measure: information treatment vs pro-social appeal	115
7.19 Fieldwork compression experiment	116

7.20 Event triggered data collection	117
7.21 Consent to send survey questions by SMS	118
7.22 Wellbeing app study	119
7.23 Asking for Living Apart Together partner details.....	120
7.24 Contact protocols for IP14 refreshment sample	121
Prenotification letters and reminders	121
Logo/branding on outside of envelope	122
Include a gift incentive.....	122
Explain the longitudinal nature of the survey	122
‘Early-bird’ incentive and deadline	122
Explain household survey	123
Collecting email addresses of household members	123
Include encouraging messages during the web survey to signal progress	123
7.25 PERKs for unconditional incentives	123
7.26 Youth online survey	124
7.27 Test of Government logo on envelopes for continuing sample	125
7.28 Youth survey invite mailing.....	126
Allocations.....	126
8. Questionnaire design experiments: general issues	127
8.1 Subsetting the questionnaire content	127
8.2 Showcards vs. no showcards	130
8.3 Impact of question wording and context on measuring change	134
8.4 Dependent interviewing wording	137
8.5 Branched vs. unbranched rating scales for measuring attitudes	138
8.6 Smiley faces vs. text-based scales in child self-completion.....	139
8.7 Quality of recall data with web vs. face-to-face	139
8.8 Methods of reducing item non-response in web surveys	139
Error in f_fiyrdia and f_fiyrdia_fu	142
8.9 Separating systematic measurement error components using MTMM in longitudinal studies	142
Template for the questionnaire.....	143
8.10 Replicating classic response order experiments across countries	146
8.11 Impact of response scale direction on responses	147
8.12 Enhancing respondent engagement with the survey through tailored interesting questions.....	148
8.13 Grid design in mobile surveys	150
8.14 Don’t know/prefer not to answer response presentation	151

8.15 Collecting mobile phone numbers.....	152
9. Questionnaire design experiments: specific topics	154
9.1 Measures of consumption and expenditure	154
9.2 Measuring satisfaction.....	155
9.3 Reference groups in measuring satisfaction.....	159
9.4 Measuring identity.....	159
9.5 Measures of wealth	160
9.6 Context of questions about consent to data linkage with administrative records	161
9.7 Respondent preferences about mode of data collection.....	162
9.8 Feasibility of directly measuring household energy use	163
9.9 Context effects in fertility decisions	164
9.10 Vignettes: measuring partner satisfaction with division of household labour	164
9.11 Subjective expectations about the returns to higher education and decisions to attend university.....	166
9.12 Measuring change in self-assessed disability	168
9.13 Associated Study: measuring time and risk preferences.....	169
9.14 Assessing how people think about environmental taxes	171
9.15 Validity of interviewer ratings of respondent health	172
9.16 Social desirability bias in attitudes towards immigration.....	172
9.17 Measuring sexual identity using direct and indirect questioning	173
9.18 What do the general population regard as “successful ageing”?	174
Allocations.....	175
Modules affected	178
9.19 Benefits unit finances module	178
9.20 Presentation of response options in satisfaction questions	179
9.21 Improving consent to link to the electoral register	179
9.22 Financial management within couples	180
9.23 Non-resident parents and reasons for separation	181
9.24 Variations of the EQ-5D questions.....	183
9.25 HMRC data linkage consent experiment	183
9.26 Does competition over public services decrease support for residency rights of immigrants?	186
9.27 Reporting of height and weight	186
9.28 Biomarker and sample collection	187
9.29 LinkedIn consent	193
9.30 Vignettes: intentions to prepare for automation	194
9.31 Comparison of ReQoL-10 to other scales	195

9.32 Proxy nomination.....	196
9.33 Twitter consent	197
9.34 Marginal propensity to consume	197
9.35 Informal care.....	198
9.36 Alcohol consumption	200
9.37 Body Volume Index app and body measurements.....	201
9.38 National identity	204
9.39 Vignettes: measuring flexibility stigma - double whammy or femininity stigma.....	205
9.40 Cognitive reflection and politically motivated reasoning.....	206
9.41 Robustness of climate change worries measurement	207
9.42 Understanding of the long-term future	207
9.43 Mental health questions comparisons	208
9.44 Asking for child red book pictures	209
9.45 Spatial cognition mobile app game	210
9.46 Consent decision process.....	211
Allocations.....	211
Follow-up questions.....	212
Modules affected	212
9.47 Identification of informal caregiving.....	212
Allocations.....	213
Modules affected	213
9.48 Labour market expectations	213
Inclusion criteria.....	214
Allocations.....	214
Modules affected	214
9.49 Indoor residential environment: consent for in-home sensor	215
Allocations.....	215
Modules affected	216
9.50 Domestic energy use: consent for smart meter data linkage	216
Allocations.....	216
Modules affected	217
10. Non-experimental studies	218
10.1 Questions about twins	218
10.2 Measuring finger length ratios as indicator of prenatal testosterone exposure	218
10.3 Associated Study: time-use diary.....	218
Format of the time diary.....	219
Fieldwork procedures	220

Response to time diaries.....	220
10.4 Spending Study 1	221
10.5 Consent to link Twitter data	221
10.6 An investigation of children’s consistency in reporting their parents’ occupations	221
10.7 Panel conditioning	222
10.8 Living Apart Together (LATs) partner and survey	222
10.9 Asking for Parents Living Apart other parent details.....	224
10.10 What constitutes an interesting size effect when measuring people’s psychological state?.....	224
10.11 People’s expectations of gender discrimination related to work	225
10.12 The extent and nature of the use of domestic workers	225
10.13 Judging the passage of time	225
10.14 Adaptation to eco-climate emergency	225
Modules affected.....	225
10.15 Youth online survey	226
10.16 Indoor residential environment and energy use: background information.....	226
New questions	226
Modified or adapted questions	227
11. Support and resources.....	228
Useful documentation links	228
Variable search.....	228
Creating syntax with the Code creator	228
Topic page	228
11.1 User Forum	228
Email.....	229
Online helpdesk	229
11.2 Training	229
References – general	230
References – summary results from IP experiments and new data.....	234
Appendix A: Controlling variables for conjoint preference experiment (9.26)	237
Appendix B: Specification for the Body Volume Index app	244

1. Data access

1.1 Innovation Panel datasets accessible through the UK Data Service

The Innovation Panel data are released through the [UK Data Service](#) and are available in SPSS and Stata formats.

There are two primary versions of the Innovation Panel data, both available from the UK Data Service. One is available under the standard End User Licence (EUL) agreement (SN 6849), and the other is a special licence version (available under SN 7083). The special licence version contains month and year of birth variables in addition to age, county variables, more detailed country and occupation coding for a number of variables, and various income variables have not been top-coded (see the documentation available with the special licence version for more detail on the differences). There is also a secure access version of the dataset (SN 7332), which includes British National Grid postcode references at 1m resolution, with assigned grid references mostly relating to the building of the matched address closest to the postcode mean.

We also release special licence geographical datasets for use in conjunction with the primary datasets. Low- and medium-level geographical identifiers are also available subject to special licence access conditions and fine detail geographic data are available under more restrictive Secure Access conditions that contains British National Grid postcode grid references (at 1m resolution) for the unit postcode of each household surveyed.

The access principles and guidance outlined in this section are derived from the [Understanding Society Data Access Strategy](#).

Along with the data files, the zipped folder will contain this Innovation Panel User Guide at the time of release and the questionnaires. However we recommend using the [web-based version of the Innovation Panel User Guide](#) as the documentation will develop over time, with the latest versions updated to these pages.

Researchers who would like to use the Innovation Panel need to [register with the UK Data Service](#) to download the dataset. Researchers should also read the UK Data Service [Research Data Handling and Security: Guide for Users](#) before downloading the data. Under the terms and conditions of access, this is required reading for users who obtain Special Licence data or Secure Access data, where disclosure risk is increased.

In addition to the main Understanding Society data, we have produced data files with information to link the Innovation Panel survey data with geographic units including Local Authority Districts, Area Classification for Output Areas, Travel to Work Areas, Westminster Parliamentary Constituencies, Local Education Authorities, and Primary Care Trusts. See the list of the available datasets on the [Access the Innovation Panel data page](#).

1.2 Additional data collections

We have conducted additional information from the Innovation Panel sample and released the data as separate studies by the [UK Data Service](#). These datasets contain the individual identifier **pidp**, which can be used to link the additional datasets listed below to the data from the annual IP interviews.

Life events

Between Waves 11 and 13 of the Innovation Panel, respondents were asked to answer a short monthly survey for one year. The monthly survey asked about the occurrence of certain life events in the previous calendar month. If respondents had not experienced any of the list of events, the survey ended after the first question. If they had experienced one or more of the events, follow-up questions were asked about the event.

Two experiments were embedded in this data collection: one on incentives and a second on the timing of the reminders. Data from the life events study is available from the UK Data Service:

- SN 8990 [Understanding Society: Innovation Panel Life Events Study, 2020](#)

Read the research: [Trialling event triggered data collection in the Understanding Society Innovation Panel: response and measurement quality](#)

Household finances

Two studies have been run in the Innovation Panel exploring household spending. In the first, respondents were invited to download an app to report daily expenditure for a month. App users were asked to upload shopping receipts or report spending directly into the app. This study included two experiments on incentives. Spending Study 1 data are available from the UK Data Service:

- SN 8749 [Understanding Society: Spending Study 1, 2016-2017](#)

Read the research: [Participation in a mobile app survey to collect expenditure data as part of a large-scale probability household panel](#)

The second spending study invited respondents to keep a diary on an app for one month, recording their expenditure including daily spending, direct debits and standing orders. This study included an experiment on how respondents were invited to take part. Spending Study 2 data are available from the UK Data Service:

- SN 8909 [Understanding Society: Spending Study 2, 2018-2019](#)

Read the research: [Increasing participation in an mobile app study: the effects of a sequential mixed-mode design and in-interview invitation](#)

Linked credit reference agency data

Data are available from the UK Data Service:

- SN 9259 [Understanding Society: Innovation Panel, Linked Credit Reference Agency \(CRA\) Dataset, 2012-2018: Secure Access](#)

Time-use diaries

In Wave 7 of the Innovation Panel, respondents were asked to keep a time diary on two separate days. The diary asked them to record how they spent their 24-hour day in 10-minute periods. Respondents were randomly allocated one weekday and one weekend day for their household, although different household members could complete the diary in different weeks. The diary was designed to be 'light touch', with respondents indicating their activities against a pre-coded list. These data are available as part of the IP7 dataset and documented in this User Guide. (See 10.3 Associated Study: time-use diary.)

Download the data:

- SN 6849 [Understanding Society: Innovation Panel, Waves 1-17, 2008-2024](#)

Social media data

At Wave 10 of the Innovation Panel, respondents were asked for permission to link their Twitter data to their survey responses. A dataset is available containing data collected from consenting respondents retrieved through the Twitter API between June 2007 and February 2023. Data in this dataset can be linked to data on the same individuals from previous and future waves of the Innovation Panel.

Download the data:

- SN 9208 [Understanding Society: Innovation Panel Twitter Study, 2007-2023](#)

Biomeasures, cognition, and wellbeing

Several studies have been carried out collecting new types of health and wellbeing data.

Finger measurement

In IP16, all respondents were asked to provide measurements of their second and fourth finger to investigate prenatal testosterone exposure effects on human development. These data are available in the IP16 data release and documented in this User Guide.

Hair sample

Respondents were asked to provide a small sample of their hair, either taken by themselves or by a visit from a professional nurse. The data coded from these samples includes hair cortisol, hair cortisone, hair progesterone, and hair testosterone. These data were released as part of the IP12 data release and documented in this User Guide.

Read the research: [Collection of biomarkers using nurses, interviewers, and participants: the design of IP12](#)

Dried blood sample

Respondents were asked to provide a dried blood sample, which they either collected themselves or was collected by a professional nurse. The data coded from these samples include dried blood triglycerides, dried blood cholesterol, dried blood high-density lipoprotein cholesterol, and dried blood glycated haemoglobin. These data were released as part of the IP12 data release and documented in this User Guide.

Read the research: [Understanding Society: health, biomarker and genetic data](#)

Pre-interview blood pressure measure

This was an advance letter experiment, where all eligible respondents were requested to take their blood pressure prior to their annual interview and then to provide the measurement in the interview. One third of the sample was provided with information on their nearest pharmacy, one third were provided with an altruistic appeal text in the invitation letter, and the last third acted as a control group. The data from these measurements were released as part of the IP12 data release and documented in this User Guide.

Child ‘red book’ data

Respondents were asked to send photographs of pages of their child’s personal child health record (‘red book’) showing information on height and weight. These data were released as part of the IP16 data release and are documented in the [IP16 Working Paper](#).

Read the research: [Understanding Society Innovation Panel wave 16: results from methodological experiments and new data](#)

Body volume index

Respondents in IP15 were asked to download an app which asked profile questions and invited them to take two pictures of their body using their device camera. The app calculated body measures, such as body fat and waist/hip ratio. Experiments in this study included unconditional versus conditional incentives, and whether offering feedback in the app affected response. These data were released as part of the IP15 data release and are documented in this User Guide.

Read the research: [Understanding Society Innovation Panel wave 15: results from methodological experiments](#)

Spatial navigation

Respondents in IP16 were invited to use a cognition app – Sea Hero Quest. This navigation game measured spatial cognition over a series of increasingly difficult levels. An experiment on incentives was included in this study. These data were released as part of the IP16 data release and documented in the [IP16 Working Paper](#).

Read the research: [Understanding Society Innovation Panel Wave 16: results from methodological experiments and new data](#)

Wellbeing app

Respondents in IP13 were invited to download an app and every evening for days respond to a set of questions on their interactions with loved ones, stressors they had experienced that day, and their mood. This app study included experiments on bonus incentives, the length of the daily questionnaire, and the position of the invitation in the IP interview. Data from the wellbeing app study is available from the UK Data Service:

- SN 9065 [Understanding Society: Innovation Panel Wellbeing App Study, 2020](#)

Read the research: [Participation of household panel members in daily burst measurement using a mobile app: effects of position of the invitation, bonus incentives, and number of daily questions](#).

1.3 Citation

The citation changes at each release to reflect the addition of the data from the new wave:

Please visit <https://www.understandingsociety.ac.uk/documentation/citation> for the citation for the latest version of the data. Please cite each dataset that you use.

If you use Understanding Society data, you must acknowledge this. 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.

Citing this User Guide

When citing this User Guide you can use the citation of this particular version quoted below. Alternatively, you can cite a previous version if required by replacing the date with that on the front of the User Guide you wish to cite. Note that the version available on the Understanding Society website is always the most up to date.

Institute for Social and Economic Research (2025). *Understanding Society – The UK Household Longitudinal Study, Innovation Panel, Waves 1-17, User Guide August 2025*. Colchester: University of Essex.

Why cite data?

Experts from across the University of Essex, the Institute for Social and Economic Research, Understanding Society and the UK Data Archive come together [in this video](#) to explain the importance of data citation. **#CiteTheData**

1.4 Ethics

The Understanding Society study protocols and research programme — including for the Innovation Panel — are scrutinised by appropriate research ethics committees to provide assurance that ethical obligations are respected at all times.

Ethical approval for the Innovation Panel has been gained for all waves. For waves prior to Wave 12, ethical approval for the Innovation Panel was received as part of the approvals for the Main survey data collection:

- By letter dated 6 July 2007 for Waves 1 and 2
- By letter dated 17 December 2010 for Waves 3 to 5
- By letter dated 20 August 2013 for Waves 6 to 8
- By letter dated 4 October 2016 for Waves 9-11

Since Wave 12, specific ethical approvals have been gained separately for each wave of the Innovation Panel:

- IP12: approval from the NHS Ethics Committee: East of England – Essex Research Ethics Committee reference: 19/EE/0146
- IP13: approval from the University of Essex Ethics Committee reference ETH1920-0863
- IP14: approval from the University of Essex Ethics Committee reference ETH2021-1115 (9 April 2021)
- IP15: approval from the University of Essex Ethics Committee reference ETH2122-0939 (6 May 2022)
- IP16: approval from the University of Essex Ethics Committee reference ETH2223-1128 (4 May 2023)
- IP17: approval from the University of Essex Ethics Committee reference ETH2324-1225 (7 June 2024)
- IP18: approval for the refreshment sample from the University of Essex Ethics Committee reference ETH2425-0780 (6 March 2025); approval for the main data collection from the NHS Ethics Committee: South Central - Hampshire A Research Ethics Committee reference 24/SC/0316.

2. Study overview

The Innovation Panel (IP) is a separate survey, conducted as part of the UK Household Longitudinal Study, [Understanding Society](#). It is designed for experimental and methodological research relevant to longitudinal surveys. The IP contains a sample of around 1,500 households used by researchers as a test-bed for innovative ways of collecting data and developing new areas of research. Its design, content, and data collection procedures are similar to the main Understanding Society survey.

The IP has multiple experimental studies in which individuals, households, interviewers or areas are randomly assigned to a particular instrument or survey procedure. Care is taken to ensure that the experiments do not affect each other, either by making treatments for different experiments explicitly orthogonal, or by using independent randomisation.

Researchers might also be interested in using the IP for non-experimental research. As an example, researchers might use measures incorporated in the Innovation Panel but not in the main Understanding Society.

2.1 Innovation Panel competition

Each year, Understanding Society runs a competition, giving researchers the opportunity to [submit ideas for a study](#) to be carried in the subsequent wave of the Innovation Panel.

Since the outset, the Innovation Panel has accepted proposals for experimental studies and other studies aiming to advance survey methods. Since IP16, the competition has explicitly also accepted submissions for non-experimental studies proposing new content to be carried in the Innovation Panel for the relevant wave.

Questionnaire space to accept studies through the competition is generally made available by not carrying Understanding Society's rotating content — those modules that are carried in the Main Understanding Society survey every few years.

2.2 Sample design

Design of Wave 1 Innovation Panel sample

The Innovation Panel (IP) is similar to the Understanding Society survey in having a stratified and geographically clustered [sample design](#). The sample covers England, Scotland and Wales but unlike the Understanding Society survey areas north of the Caledonian Canal and Northern Ireland are excluded.

In Wave 1 post code sectors from the Postcode Address File (PAF) were ordered by Government Office Region, the percentage of household heads classified as National Statistics Socio-Economic Classification (NS-SEC) categories 1 and 2 (non-manual) and population density. A systematic random sample of 120 sectors was drawn with probability proportional to population size. Within each selected sector, 23 addresses were selected by systematic random sampling, making a total sample of 2,760 addresses from 120 sectors. The final stage of sampling was carried out in the field by interviewers. Additional detail can be found in the working papers on the [sample design](#) and [weighting strategy](#) for Understanding Society.

All household members aged 16 years or older were eligible for the adult interviews. As with the main survey fieldwork design, all persons resident at the address, including children, were defined as original sample members (OSM) to be followed throughout the life of the study. In contrast to the main general population sample, the IP does not attach absent household members living in institutional accommodation to the IP sampled households. This introduces some degree of coverage error since these persons do not otherwise have an independent chance of selection through the Postcode Address File. Despite these differences, the IP has many similarities to the overall sample design of Understanding Society.

Refreshment sample in Wave 4

An additional 960 addresses, eight in each of the original 120 PSUs, were added as the refreshment sample. Addresses were selected using systematic random sampling from amongst addresses not already selected for the IP. If a household at one of the added addresses responded, members were defined as Original Sample members, from Wave 4 onwards.

Refreshment sample in Wave 7

An additional 1,560 new addresses were added as a further refreshment sample. This includes an initial refreshment sample of 1,080 households (9 in each of the 120 original PSUs) and a boost refreshment sample of 480 households (4 in each PSU). The boost sample was added midway through fieldwork due to the low response rate achieved on the initial refreshment sample. If a household at one of the added addresses responded, members were defined as Original Sample members, from Wave 7 onwards.

Refreshment sample in Wave 10

An additional 960 new addresses, 8 in each of the original 120 PSUs, were added as a further refreshment sample. In addition, a further 455 reserve refreshment sample households were issued in July 2017. Addresses were selected in the same way as for the Wave 4 refreshment sample, and members of respondent households defined as Original Sample members, from Wave 10 onwards.

Refreshment sample in Wave 11

An additional 1,680 new addresses, 14 in each of the original 120 PSUs, were added as a further refreshment sample. In addition, a further 832 reserve refreshment sample households were issued in September 2018. Addresses were selected in the same way as for the Wave 4 refreshment sample, and members of respondent households defined as Original Sample members, from Wave 11 onwards.

Refreshment Sample in Wave 14

An additional 6,047 addresses were added as a further refreshment sample. Unlike previous refreshment samples, this was recruited by web only. Addresses were selected from the Postcode Address File (PAF), clustered in 32 areas. These addresses were then contacted by post and invited to complete the survey online.

Sample for Wave 14 LAT survey

All IP13 and IP14 participants who were in a living apart together (LAT) relationship, and those partners for whom they had provided contact details, were invited to take part in a web survey between March and May 2022. Information on the sample, contact attempts and outcome is in **n_latsamp_ip**. For partners who took part, a **pidp_sosm** was created for them in the data file of their web survey **n_latsurvey_ip**. Note this file is organised by the sample members' **pidp**.

Sample design variables

The sample design is described by three variables, indicating sampling stratum, primary sampling unit and design weight.

Sampling stratum and primary sampling unit (**psu**) are identified by **w_strata** and **w_psu** respectively. They are on individual level enumeration and response files, as well as **w_hhsamp_ip**. The individual level enumeration files are **w_indall_ip**. The individual level response files are **w_indresp_ip** and **w_youth_ip**.

The person-level design weight is **a_psnenip_xd** and it is on **a_indall_ip**, **a_indresp_ip** and **a_youth_ip**. The household level design weight is **a_hhdenip_xd**. It is on **a_hhresp_ip** and **a_hhsamp_ip**.

2.3 Annual interview process

Interviewing for the IP commences in the spring of the year before the corresponding wave of the main [Understanding Society survey](#). The IP shares the same basic interview structure as the main survey which includes:

- Household roster and household questionnaire
- Individual questionnaire, with a brief proxy interview for any respondents who cannot be interviewed in person
- Adult self-completion instrument
- Youth self-completion instrument

Interview modes: web, CAPI and telephone

In the earliest waves of the IP, the primary mode of data collection was face-to-face interviews using computer-assisted personal interviewing (CAPI). However, Wave 2 of the IP experimented with a mixed-mode design in which a portion of the sample was interviewed via telephone using computer-assisted telephone interviewing (CATI). In IP Wave 5 of the IP, another mixed-mode experiment began: using an online web survey along with the CAPI-only design. This mixed-mode design has been used since, with some variations described in the [fieldwork procedures](#).

In the CAPI instrument, the questionnaire is a computer program in which the computer shows the questions on the screen and the interviewer reads them to the respondent and records the respondent's answers. CATI is the same but with the interview taking place over the telephone. For web instruments, the questionnaires are similar to the CAPI version,

though designed to be completed online by the sample member themselves without an interviewer present.

The self-completion instruments for adults and youth in IP1 and IP3 were paper-based. There is no adult self-completion interview in Wave 2 of the IP. In IP2, the youth self-completion was available in two formats: a web questionnaire and a paper self-completion. The web address and individual log-on passwords were printed on the advance letter/card of the responsible adult. Interviewers also had paper copies of the self-completion for those that had not done it online. This was not experimental. In IP4 and from IP6 onwards, the youth self-completion instrument was paper based. At Wave 5, the youth self-completion was again available online as well as paper copies. Alongside paper self-completion youth surveys, web youth surveys were again available at Waves 16 and 17 of the IP (see 7.26 Youth online survey and 10.15 Youth online survey).

From Waves 4 to 6 adults in a random half of households completed a paper self-completion, while adults in the other half of the sample were asked to complete their questionnaire using the interviewer's laptop (CASI: computer-assisted self-interviewing). Beginning at IP7, the self-completion part of the adult interview was moved entirely to computer, either as CASI as part of the face-to-face interview or as part of the web survey.

Other [fieldwork documents](#) include advance letters for communication with participants, other correspondence with participants, information leaflets, interview instructions, consent forms, and showcards. The fieldwork documents are organised by wave.

2.4 Fieldwork procedures

The [fieldwork procedures](#) document provides information on how the fieldwork was conducted for each wave of the Innovation Panel.

2.5 Response outcomes

The [fieldwork response tables](#) document shows the household and individual response outcomes (response rates) for each wave of the Innovation Panel.

Using Innovation Panel response outcome data

The above analyses rely on sample data from the IP. There are a number of response outcome codes that were combined to identify respondents, non-respondents, and ineligible units. There are separate files for the household sample, individual sample, and individual response outcomes. The allocation of units to conditions, such as the mode experiments, are all found on the household sample file (allocation is done at the household level). The files will need to be linked to examine individual differences by allocations.

The table below presents the naming convention for the data files and variables used. In this table **w_** denotes the wave of interest, and will need to be replaced with a letter to represent the wave, for example **c_** for Wave 3. Note that at Wave 1, the file **a_indall** is used for all individuals (the 'indall' file is generated from the household enumeration grid and includes all individuals enumerated in respondent households); for later waves the file **w_indsamp** is used, as these are the individuals actually issued to field, including individuals in respondent and non-respondent households. The mixed-mode allocation **b_ff_modew2** is

specific to the Wave 2 experiment. The variable **w_hidp** is used to link household data to individual data, while **pidp** is used to link between individual respondent datasets. Sample identifier variables appear only in Waves 4 forward.

Variables used for Innovation Panel response outcomes

	Data set	Response variable	Sample variable	Mixed-mode variable	Linking variable
<i>Household sample</i>	w_hhsamp_ip	w_ivfho	w_hhorig	b_ff_modew2, w_ff_gridmodew5 w_ff_gridmodew8 w_ff_gridmodew9 w_ff_lowwebw8	w_hidp
<i>Individual sample</i>	a_indall_ip w_indsamp_ip	w_ivfio	w_memorig		w_hidp, pidp
<i>Individual response</i>	w_indresp_ip			w_indmode	w_hidp, pidp

Stata code

The [annotated Stata code](#) used to calculate all of the IP7 response rates presented shows the outcome codes used for responding, non-responding, and ineligible units. This logic can also be expanded to define more categorisations (e.g. separating refusals and non-contacts). The code also shows a worked example of how IP data sets are used in combination generally.

3. Data structure

Multiple files are released for each wave of the Innovation Panel. The table below summarises the general content and level of data files. The questionnaire data, sample information, and paradata files are released in each wave. The Life History files exist only for some waves. The table summarises which units are contained in the file, which instrument the data are from, and points to key variables useful in linking data files for analysis. Variants in the files released are summarised on the [Additional data files](#) page.

3.1 Files released every wave

List of data files released in most waves of the Innovation Panel

Data file	Description of units	Source of data	Unique identifiers*
Questionnaire data			
INDALL	all enumerated persons at respondent households	household grid module	w_hidp, w_pno, pidp
EGOALT	derived file containing one record for each pair of enumerated individuals within each household	derived from household grid module	w_hidp, w_epno, w_apno, pidp, w_apidp
HHRESP	respondent households	household questionnaire	w_hidp
INDRESP	Respondent individuals, including proxy interviews (w_ivfio=2)	individual questionnaire including self-completion	w_hidp, w_pno, pidp
CHILDCARE	a record for each child receiving childcare	childcare module in individual questionnaire	w_childpno, w_hidp, w_pno, pidp
INCOME	a record for each income source reported by an individual	household finance module in individual questionnaire	w_hidp, w_pno, w_ficode
YOUTH	youth respondents (10-15 years old)	youth self-completion questionnaire	w_hidp, w_pno, pidp
Retrospective data from Life History modules (RELEASED ONLY IN SOME WAVES)			
ADOPT	a record for each adopted or step-child	fertility history module in individual questionnaire	w_adoptno, w_hidp, w_pno, pidp
COHAB	a record for each cohabitation spell outside of legal marriage except those that preceded marriage	partnership history module in individual questionnaire	w_cohabno, w_hidp, w_pno, pidp
MARRIAGE	a record for each reported legal marriage – from PARTNERSHIPHISTORY module	partnership history module in individual questionnaire	w_marno, w_hidp, w_pno, pidp
NATCHILD	a record for each biological child	fertility history module in individual questionnaire	w_childno, w_hidp, w_pno, pidp

NEWBORN	a record for each child born since previous interview	newborn module in individual questionnaire	w_childno, w_hidp, w_pno, pidp
Sample information			
INDSAMP	a file with every person enumerated at the current or previous wave	if a household has split then there are two observations for each person: one with the identification number (w_hidp) for the original household and the individual interview outcome (w_ivfio) indicating that the individual has moved, and one with the individual interview outcome and identification number for the new household	w_hidp, w_pno, pidp
ISSUE	households issued to the interviewers	fieldwork information	w_hidp, w_issue
HHSAMP	all sampled addresses, includes experimental allocations (variables with prefix "w_ff_")	sampling information, response status, non-respondent household characteristics	w_hidp
Paradata, Information about the collection of data			
CALLREC	a record for each visit made by interviewers to attempt to interview the household	information about date, time and status of each visit	w_hidp, w_issueno, w_int_num, w_callno
PTIMINGS^	respondents to individual adult questionnaire	amount of time for individual interview modules; for web interviews (Waves 5, 6) the timings are per screen	pidp
HTIMINGS^	households who responded to household questionnaire	amount of time for household interviews	w_hidp

* "w_" is a placeholder for the wave prefix ("a_" for wave 1, "b_" for wave 2, etc.)

^ From wave 7 onwards timing data are released in CSV files due to the variable naming structure. See below.

3.2 Variable naming conventions

Understanding Society has a distinct naming convention for its datafiles to identify which wave the data is from and the source of the data.

The naming convention for variables follows the same rules as file names. Variable names have the same root name which is fixed over time, and begin with a prefix to reflect the wave the data are collected “a_” for Wave 1, and “b_” for Wave 2, etc. In this user guide we have used the wave prefix as “w_”. The names of records in the Innovation Panel have the “_ip” suffix. We have attempted to keep the names of variables that came from the British Household Panel Survey (BHPS) the same.

Variables ending in “_dv” are derived variables.

Variables ending in “_code” are text variables that have been coded, such as indicators for whether the respondent gave address information, such as an email address.

The variable labels for some variables begin with “mc”. This indicates that the variable is related to a multi-code question, where respondents can select all response options that apply.

Additional codes denote different reasons for the lack of a valid response. These values have not been specified as missing in Stata or SPSS. However, these statistical packages have commands to assign values to missing for many variables simultaneously. The codes are:

- 1 Don't know
- 2 Refused
- 7 Proxy respondent
- 8 Not applicable to the person or because of routing
- 9 Missing by error.

The meaning of substantive codes is explained with each variable's value labels.

Note that in Waves 5, 7 and 8 there are some variables in the **w_youth_ip** file with a value of 9 which is labelled as "Not Applicable". The variables are:

Datafile	Variable
e_youth_ip	e_ycawiout, e_ycapiout
g_youth_ip	g_ypestc, g_ypestk, g_ypfythclub, g_ypdistv, g_ypfadmus, g_ypfhmwrk, g_yphmwkhlp, g_yplvsc2do, g_ypsrhlth, g_yphlwtr, g_yptrvl2sch, g_ypsmrsk2, g_ypalcrsk2, g_ypmjrsk1, g_ypmjrsk3, g_ypersk1, g_ypamrsk1
h_youth_ip	h_ypcomp, h_yppchw, h_ypconstm, h_yupset, h_ypsibverab, h_ypstealsib, h_ypverabsib, h_yphlf, h_ypllknbrd, h_ypacvwell, h_yplvsc2do, h_yp2uni, h_ypfrobulli, h_ypsave, h_ypdklm, h_ypfght

3.3 Additional data files in some waves

There is substantial variation in the data files produced in Wave 3 from the **fertility** and **partnership history** modules, which produce **marriage**, **natchild**, **adopt** and **cohab**. This is related to experiments where random subsets of the questionnaire content were presented to respondents (see 8.1 Subsetting the questionnaire content).

Generally, we obtain a complete fertility history in the first wave and then ask only about updates. New entrants to the sample are asked a somewhat shorter version of the module about biological (natural) and adopted children. At IP1, we gave the full fertility history to half of the sample, with the other half receiving the full history at IP2. At IP3, we used the “new entrant/never interviewed” short version. At IP4, the refreshment sample members were all OSMs and so got the full history.

The same logic follows for issues related to Marriage and Cohabitation from the **partnership history** module.

Some files have only been produced for a few waves:

- In Wave 4, there is a file related to the Early Bird appointment experiment (See the Experiments section) **d_admineb_ip**.
- Files related to the Measures of Wealth experiment (see the Experiments section) in Wave 3 – **c_hholdinvest_ip** and **c_wealth_ip**.
- Files **d_parstyle_ip** and **d_chdev_ip** and **d_breastfed_ip** are related to questions about parenting style, child development, and breastfeeding behavior. They appear in Wave 4. File **w_chdev_ip** also appears in waves 10 and 11. File **e_parstyle_ip** is also released in Wave 5. **f_breastfed_ip** is also released in Wave 6 and **j_breastfed_ip** in Wave 10.
- File **c_wealth_ip** is related to types of assets and is released for Wave 3.
- File **f_meterreading_ip** contains data from the postal/telephone follow-up collection of electricity, gas and odometer readings in Wave 6. See the Experiments section for the description of this experiment.
- File **g_timediary_ip** contains data from paper time diaries collected in Wave 7 (see the Experiments section).
- File **i_bufind_ip** is an individual level file. It contains the computed monthly income from benefits, pensions, and other unearned income sources which feeds into the Benefit Unit Finances module in Wave 9.
- File **i_benefitsum_ip** is a benefit unit level file. It includes the computed monthly income of the benefit unit (individuals or couples and their dependent children) from benefits, pensions and other unearned income sources in Wave 9.
- File **i_bufinance_ip** is a benefit unit level file. It contains data from the Benefit Unit Finances module in Wave 9.
- Files **j_nonresch2_ip** and **j_nonressum_ip** and **j_nonresid_ip** contain data from the ‘non-resident parents and reasons for separation’ experiment in Wave 10. For the description of this experiment, see 9.23 Non-resident parents and reasons for separation.

- File **k_keystroke_paradata** contains data captured automatically by the questionnaires script, while respondents completed the HMRC consent module and the HMRC consent follow-up module. See description of paradata files below.
- In Wave 12, there are additional files containing the biomeasures collected.
- File **l_experience_ip** contains data from one off questions about the respondents' experience of the interview, such as questions about the kit supplied, whether the respondent would provide blood samples, and the experience of being in IP12.
- File **l_hair_ip** contains data on the collection of hair during the Nurse interview visit.
- File **n_latsurvey_ip** contains data collected with a web survey from participants in a Living Apart Together (LAT) relationship and LAT partners for whom we have contact data.
- File **n_latsamp_ip** contains the sample of participants in a Living Apart Together (LAT) relationship and LAT partners for whom we have contact data invited to the web survey, information about the contact attempts and survey outcome.
- File **n_latsurvey_ip** contains data from participants in a Living Apart Together (LAT) relationship and LAT partners for whom we have contact data. The data were collected with a web survey of sample members who are in LAT relationships and their partners. The partner's data is listed under the sample member's pidp. If they completed the survey a pidp has also been created for them (**pidp_sosm**).
- File **o_bviapp_ip** contains additional data on body measurements. IP15 respondents were asked to install the BodyVolume app on their smartphone or tablet (iOS or Android) and use it after the interview to take two photos of themselves. The app used the body outlines along with profile information that the respondent entered in the app (age, sex, height, weight, level of activity) to calculate measures including waist and hip circumference, total body fat, visceral body fat, and lengths of different body parts. See experiment 9.37 Body Volume Index app and body measurements.
- File **p_redbook_ip** contains health related data on children under 16 collected from the children's red books. For details of the experiment associated with this data collection, see 9.44 Asking for child red book pictures. See the Child Development Measures from the 'red book' for details on the content of this data file.
- File **p_shq_results_ip** contains data from a mobile app game that tested spatial cognition in IP16. The file contains one observation for every respondent who downloaded the game on their mobile device. For details on an experiment varying incentives for participating in this game, see 9.45 Spatial cognition mobile app game. See Coutrot et al (2024) for details on the content of this data file.

3.4 Derived variables

Derived variables are variables that are computed from one or more variables. Some are computed during the interview to control the routing within the questionnaire and can be identified in the questionnaires by searching for "Compute". Others are computed post-field for the purpose of analysis and are positioned last in the data files with the suffix **_dv** in the

variable name. There are exceptions to this rule. Pointers to significant others in the household (such as the natural parents), based on edited information in the household grid, end on the familiar **_pidp** and **_pno**.

Variable search

The derived variables are documented as part of the online [variable search](#) and contain notes giving information on how they were derived, for example variable [agegr5 dv](#).

Information about how a derived variable is produced is shown in the Derived Variable Note field of the variable. The [variable search](#) provides descriptive statistics for each variable and, in the Origin field, lists the variables used in the computation of the derived variable. For variables that were computed during the interview, additional information is available in the questionnaires.

Identifiers and other useful variables

Households are identified by “**w_hidp**”, a wave-specific variable with a different wave-specific prefix for each wave. As shown in the table below, “**w_hidp**” can be used to link information about a household from different records within a wave. “**w_hidp**” cannot be used to link information across waves. Since the composition of households changes between waves, the data do not include a longitudinal household identifier.

Individuals are identified by the personal identifier (“**pidp**”), which is constant in all waves, and can be used to link information about a person from different records belonging to one wave, or to link information from different waves. Individuals are also identified by “**w_pno**” – the person number within the household. The combination of “**w_hidp**” and “**w_pno**” is unique for each individual and can also be used to link information about individuals within a wave.

Useful variables

Variable	Description	Available in File
w_hidp	Household identifier	All files
pidp	Cross wave person identifier	All EXCEPT w_hhsamp_ip, w_hhresp_ip
w_gor_dv	Government office region (Wave 1)	w_hhsamp_ip, w_hhresp_ip, w_indall_ip, w_indresp_ip
w_pno	Person number within the household	All EXCEPT w_hhsamp_ip, w_hhresp_ip
w_sex	Sex	w_indall_ip, w_indresp_ip
w_dvage	Age	w_indall_ip, w_indresp_ip
w_hgpart	PNO of spouse/civil partner	w_indall_ip, w_indresp_ip
a_psnenip_xd	cross-sectional person design weight	a_indresp_ip, a_indall_ip, a_youth_ip

a_hhdenip_xd	cross-sectional household design weight	a_hhsamp_ip, a_hhresp_ip
w_indinip_lw	longitudinal adult main interview weight	w_indresp_ip
w_psu	primary sampling unit	w_hhsamp_ip, w_hhresp_ip, w_indall_ip, w_indsamp_ip, w_indresp_ip, w_youth_ip
w_strata	sampling strata	w_hhsamp_ip, w_hhresp_ip, w_indall_ip, w_indsamp_ip, w_indresp_ip, w_youth_ip

Occupation codes

Understanding Society collects free text information on respondents' job titles and industry. Industry descriptions are coded to ONS Standard Industry Code 2007, or SIC 2007. Job titles are coded to the ONS Standard Occupational Classification 2000, or SOC 2000. Coding is undertaken using the Computer Assisted Structured Coding Tool (CASCOT) system. We use look-up files between SOC 2000 and other classifications provided on the [CAMSIS website](#) to derive further occupational classifications.

We provide the following classifications: International Standard Classification of Occupations (ISCO88), Registrar General Social Class (RGSC), National Statistics Socio-economic Classification (NS-SEC), Employment Status (ES), and Socio-economic Group (SEG).

3.5 Derived income variables

Income from a variety of sources is measured through a number of questions, some of which are used to also create derived variables (these are discussed on the [Derived variables page](#)).

Top-coding of income variables

Because extremely high incomes are rare, releasing this information can increase the risk of disclosure. Therefore, to ensure people with very high incomes are not identifiable in any way, reported income beyond a certain amount is top-coded to a maximum annual value of £180,000. Some income variables can be reported on rates other than an annual basis. In these cases, values are top coded to an equivalent amount for the relevant period (for example, £15,000 for a monthly reported value being equivalent to £180,000 per year). The variables affected by this top coding and the maximum annual values are reported in the table below.

Top-coded variables and maximum values

Variable	Description	Top-coded value
w_fiyrdia	Amount received in interest/dividends	180,000
w_fiyrdic_dv	Income from savings and investments (annual)	180,000
w_j2pay	Gross earnings from seconds jobs last month	15,000
w_paygu_dv	Usual gross pay per month: current job	15,000
w_paynu_dv	Usual net pay per month: current job	15,000
w_payu_dv	Usual pay per month	15,000
w_payu	Usual pay	180,000
w_payg_dv	Gross pay per month in current job: last payment	15,000
w_payn_dv	Net pay per month in current job: last payment	15,000
w_jsprf	Self-employed: net profit in last yearly account	180,000
w_jspayu	Average income from job/business	15,000
w_seearngrs_dv	Self-employment earnings – gross	15,000
w_fimnlabgrs_dv	Total monthly labour income – gross	15,000
w_fimngrs_dv	Total personal income – gross	15,000
w_paygl	Gross pay at last payment	180,000
w_paynl	Net pay at last payment	180,000

3.6 Fieldwork paradata

Additional data collected during the interview process (paradata) are available. These consist of call records, timings data and other information collected by the interviewers during the interview.

Call records

Call record files have information on the number of calls made as well as the issue number, interviewer identifier (scrambled), time and date and the outcome of each call. This is available in the dataset **w_callrec_ip**.

Address response form

Information collected in the address response form (ARF) by interviewers while contacting each household and requesting household members to participate in the survey is available in **w_hhsamp_ip**. This includes data on the area surrounding the address, the type of accommodation and other information that the interviewer can observe for both responding and non-responding households. Reasons for refusal are also available. Interviewers also record some information about the quality of the interview and persons present during the interview process. This is available along with substantive data collected during adult individual interviews (including proxy interviews) in **w_indresp_ip**. From Wave 7 onwards the ARF was no longer used.

Timings data files

Timings data files (**w_ptimings** and **w_h timings**) include data on the time taken to complete each question and module in the individual and household questionnaires. In IP1, the start and end times are given for blocks of questions, where blocks are one or more question modules. The times are given in seconds. From IP2 onwards the times are given in seconds for individual questions. If the variables are asked in a loop or multi-choice format, the variable name is suffixed with the multi-choice item number or loop iteration count. In Waves 5 to 9 the timings data for interviews completed by web are per screen rather than per question, although most screens contain only a single question. Where there are multiple questions per screen this is documented in the pdf questionnaire. Waves 7 onwards are released in CSV format because the variable names are long strings that are truncated when imported into Stata. From Waves 7 to 16 the timings files are **w_hhgrid_timings**, **w_hhint_timings**, and **w_indint_timings**. From Waves 17 onwards the timing files are **w_hh_module_timings**, **w_hh_question_timings**, **w_iv_moudle_timings** and **w_iv_question_timings**.

The IP11 timings data included an error which has been corrected (see the example 3.8 Stata code for matching files3.8 Stata code for matching files).

Interviewer characteristics

The interviewer **id w_intnum** can be linked to the mainstage cross-wave file **xivdata** which contains interviewer characteristics. This file is available from the UK Data Service as a [separate dataset \(SN 8579\)](#), under Special Licence agreement.

Keystroke paradata

For IP11 there is an additional paradata file (**k_keystroke_paradata**), which contains information automatically recorded from CAPI and web respondents, while they answered the questions in the modules “HMRC consent” and “HMRC consent follow-up” (early and late versions). For each question in these modules the strings in the variables **k_keystrokes1** and **k_keystrokes2** record the question name, the response category selected, and the timestamp when the interviewer or respondent clicked ‘next’. The variable **k_keystrokes1** is truncated for some cases and the remainder of the string can be found in **k_keystrokes2**.

3.7 Known data issues

Some of the known problems relate to problems in implementing the experiments.

Cross-wave issues

In all waves, the benefit income data has not been edited for outliers.

Wave 1

In Waves 1 and 2, we asked participants for consent to link administrative records to survey data. We will not be linking the administrative records because some of the consent forms have been lost. In this data release we are including Wave 1 consent variables. Wave 2 consent variables have been restructured to improve clarity.

Wave 2

In Wave 2, a variable for **w_ivtrans** (translator used) was not collected. However, there is a related variable available in waves 1-4, **w_ivaffct22** (in what way was the respondent influenced: Other helped in translation, reading showcards, and other survey tasks).

Wave 3

In Wave 3, the Showcard experiment required at least some interviewers to use showcards for some participants and not for others. There are doubts about whether interviewers correctly followed the instruction about which sample members should have showcards. This situation could create errors and there is no check that would tell us whether or not the respondent saw the showcards.

In Wave 3, some respondents were incorrectly asked the experimental IP2 satisfaction questions in addition to the IP3 questions in relation to the satisfaction experiment. This happened with respondents with values 7, 8, 9, or 10 on the IP2 treatment indicator **b_ff_lifesatw2**. The responses to the IP3 questions for the respondents are potentially affected by having answered similar questions earlier in the interview. The questions that should not have been asked are **c_lfsat** variables ending in **_g** to **_j**. The **c_lfsat** variables ending in **a_** to **_f** are correct.

Variable **c_conddateh**, which is about strategies used to recall dates for a health condition beginning says it is a “check all that apply” variable. However, it was implemented as “select one”. Similarly, the variable **c_pldateh**, which is about strategies used to recall dates for a move is documented as a “check all that apply” variable, but was implemented as “select one”.

Variables related to **nssec** in Wave 4 for current and last job (previously not included) are in the current release. These include the 3, 5 and 8 category classifications.

A variable for highest qualification is not released because there has been a change in the response categories for educational and vocational qualifications.

In the Wave 3 Annual Events questions about employment, there are inconsistencies. For the first job, the question on the type of employment (**nxtjbes**) is less detailed than the one in the loop if they have additional jobs after this (**nextjob**). **nxtjbes** only asks if they were employed or self-employed, whereas **nextjob** asks if they were doing a different job for the same employer, working for a different employer or working as self-employed. This is only a problem for the first job reported in the annual events.

There are inconsistencies in variable names and variable labels in employment histories between IP2 and IP3/IP4 (because of changes in the way the histories are collected). From wave 3 the loop through jobs starts at the second employment spell, whereas in IP2 the loop begins at the first spell. As a result the variable names are slightly inconsistent between IP2 and IP3/IP4. At IP3, the variable **nxtst** is supposedly equivalent to the variable **nextstat1** at IP2 – i.e. it's the first employment spell. However, it seems the variable **nextstat1** has been incorrectly labelled as this first spell (it is in fact the 2nd employment spell after **nxtst**).

Wave 5

Errors in the Wave 5 questionnaire

The grid, household questionnaire and individual questionnaires were all programmed as separate web instruments, whereas the CAPI questionnaire was programmed as one combined instrument. In previous waves, the feed-forward data sat within the household grid, and any text fills or routing in the household or individual questionnaires were programmed via a reference to the household grid data. In IP5, because the web instruments were programmed separately, the feed-forward data needed to be copied into these instruments, so that it could be referenced within the household or individual instrument. Each feed forward variable was copied individually (using code). There were mistakes in the code copying feed-forward data into the household and individual questionnaires. For subsequent waves, the whole feedforward is copied as a block, to ensure that all feedforward variables are copied correctly.

Feed-forward variables determine which experimental questions are asked in an interview, so the copying errors corrupted some of the experiments. This section describes their effects.

Household questionnaire. At the household level, three feed-forward variables: **e_ff_rentwc**, **e_ff_metersw5** and **e_ff_diw5** were improperly copied. The related variables about gas or electric meter reading were not asked and were not released in the data.

Additionally, the **e_ff_diw5** variable did not have the correctly assigned experimental values. This meant that the dependent interviewing (DI) experimental variables in the household questionnaire were confounded, in that some DI questions were asked, but not the ones that should have been according to the experimental design. There were four sets of questions affected by this confounding: **hsrooms/hsbeds** (number of bedrooms and other rooms at the address); **hsownd** (tenure); **xpmg** (monthly mortgage payments) and **rent/rentwc** (amount and frequency of rent). Some variables were combined to facilitate analysis; others were not released (see summary below).

The affected variables in the household questionnaire were:

Summary of household variables affected by errors	
Variable	Impact
e_ff_metersw5	Blank due to programming error
e_ff_diw5	Incorrect values due to programming error
e_hsroomchk	Combined version released
e_hshowndchk	Combined version released
xpmg_a	Asked, but wrong experimental version, not released
xpmg_b	Asked, but wrong experimental version, not released
xpmg_c	Asked, but wrong experimental version, not released

xpmg_d	Asked, but wrong experimental version, not released
ff_rentwc	Blank due to programming error
rentchk_a	Asked, but wrong experimental version, not released
rentchk_b	Asked, but wrong experimental version, not released
rentchk_c	Asked, but wrong experimental version, not released
rentchk_d	Asked, but wrong experimental version, not released
gasuse	Not asked due to programming error in FF_MetersW5
gasuse_cawi	Not released
gasmeter	Asked, but wrong experimental version, not released
gasest	Asked, but wrong experimental version, not released
elecuse	Asked, but wrong experimental version, not released
elecmeter	Asked, but wrong experimental version, not released
elecest	Asked, but wrong experimental version, not released

Errors in the Wave 5 individual questionnaire

There was an error in the code copying three feed-forward variables in the employment modules of the individual questionnaire which meant that they were blank, namely: **ff_jbmng**, **ff_jbsize** and **ff_jbterm1**. This affected multiple variables which were not released. See the summary below.

Due to an error in the code, none of the **e_ff_bentype01** to **e_ff_bentype37** variables were copied into the individual questionnaire. This affected the **nfh01** to **nfh37** variables about benefit income. It only affected those people who did not mention a benefit that they said they were receiving the previous year. Such people will not have received the additional prompt question reminding them of last year's answer. Our estimate is that around three-quarters of respondents were not eligible to be asked any additional prompt questions in the first place; of those who were eligible to be asked any, a large majority (around 70 per cent) only missed out on one such question, 20 per cent missed out on two, and ten per cent missed out on three or more.

The **e_ff_casiw5** variable was not copied into the individual questionnaire at the start of fieldwork. The variable controls the mode of the self-completion questionnaire. The problem was resolved part way through the fieldwork period (after June 11). We created a variable **e_scflagip5** (on **e_indresp_ip**) to show the status of mode of completion for the self-completion questionnaire in Wave 5. The effect of the error is that around 50 per cent of those eligible to receive the questions in face-to-face CASI mode did not get asked the experimental questions (313 people, based on unedited data). It should be noted that this does not confound the experiment (i.e. no respondents were asked questions in the wrong

mode), but the reduced numbers mean that it does reduce its power to detect mode differences.

The affected variables in the Wave 5 individual questionnaire were:

Summary of individual level variables affected by errors in feed forward variables	
Variable	Impact
ff_jbmngr	Blank due to programming error
jbmngrchk	Not asked because ff_jbmngr was blank
ff_jbsize	Blank due to programming error
jbsizechk_a through to jbsizechk_d	Not asked, not released
ff_jbterm1	Blank due to programming error
jbterm1_a through to jbterm1_d	Not asked, not released
ff_bentype01 to ff_bentype37	Blank due to programming error
nfh01 to nfh37	Not asked because FF_BENTYPE01 – FF_BENTYPE37 were blank
ff_cawiw5	not released
SF12 Module	Not asked of some respondents (identified by variable e_casiflagger) due to programming error that meant that some respondents were not asked part of the self-completion questions
GHQ Module	Not asked of some respondents (identified by variable e_casiflagger) due to programming error that meant that some respondents were not asked part of the self-completion questions
Parental Relationships Module	Not asked of some respondents (identified by variable e_casiflagger) due to programming error that meant that some respondents were not asked part of the self-completion questions
Alcohol Module	Not asked of some respondents (identified by variable e_casiflagger) due to programming error that meant that some respondents were not asked part of the self-completion questions
Personality Module	Not asked of some respondents (identified by variable e_casiflagger) due to programming error that meant that some respondents were not asked part of the self-completion questions

Wave 6

In Wave 6, four households in the £10 incentive treatment group became aware of the £30 treatment. To compensate they were offered an extra £20. The households are identified by the variable **f_incentcomp** on the record **f_hhsamp**.

Wave 7

In Wave 7, there are a few households with missing values for the experimental treatment allocations in the **hhsamp** file. The initial IP7 sample used to generate the experimental allocation variables was based on what was the latest IP6 data delivery at that time. Later data deliveries included some additional households. Most of these extra households were untraced and we had no addresses for them to go out into the field. For the few extras that did have an address we generated randomisations for the experimental variables separately. For households with missing address information the experimental variables remained missing.

Wave 11

In Wave 11, the variables related to height and weight have been removed from all waves in the **w_youth_ip** record, due to measurement problems with these variables.

The Wave 11 individual interview question timings file (**k_indint_timings.csv**) contained two errors which have been corrected with the version released with IP14. The first error was the following. When the timings file was created, respondents with the same Serial ID (a household-level fieldwork identifier) were updated with values from another interview that had the same Serial ID, overwriting values (e.g. if they both answered ConsentQ3, both interviews would have the same value for ConsentQ3 and should have been different). If they were routed to different versions of questions – e.g. ConsentQ3 and ConsentQ4 – one respondent would have timings for both questions. This has been corrected by ensuring that the serial ID is used with other identifiers to uniquely identify individual cases. This means that the updated timings data has changed across all the timings variables and the derived summary variables. The second error was that some observations for questions in modules that the respondent was not routed into contained the value “12/30/1899 0:00:00” instead of being blank. This has also been corrected.

Wave 13

In Wave 13, there was an error in the sample file. The variable **ff_eventtrigw12** was erroneously set to missing for all sample members. As a result, the question “**eventdebrief**” that should have been asked for all sample members invited to the event-triggered data collection during 2020/2021 (**ff_eventtrigw12=1**) was not asked of anyone. In addition, in the introduction to the Annual Events History (“**calintro**”), the text fill “Please tell us about all changes, even if you have already reported them in the monthly questions about life events that we have been trialling. The reason for asking you again is that in this interview we are interested in different aspects of any changes you have experienced.”, which should have been shown to all respondents invited to the event-triggered data collection, was not displayed to anyone.

Wave 14

In **Wave 14**, the variable **father** ("Fathered children since last interview") is not populated for about 300 cases where it should be. This error occurred because an age filter was left active from a prior question, so respondents aged > 64 years were not asked.

Wave 15

In **Wave 15**, there is a household that participated in the wave but had no allocations for the experimental conditions (**ff_** variables in the **hhsamp** file). This was a late re-joiner household, that was lost at Wave 13 but re-joined the panel for wave 15 but after the allocations had been made.

Wave 16

In **Wave 16**, there are three households that participated in the wave but had no allocations for the experimental conditions (**ff_** variables in the **hhsamp** file). These were late re-joiner households, that were lost previously but re-joined the panel for Wave 16 but after the allocations had been made.

In **Wave 16**, no respondents were routed into the proxy questionnaire module that asked about respondents who had moved into a care home (module "**carehomeproxy**" in the IP16 questionnaire). The corresponding variables were therefore dropped from the file **p_indresp_ip**. Similarly, in the household grid, no respondents were routed into the questions about household members who were reported as having moved into a care home at the previous wave. The variables **chomestill** – **chmrespdp** were therefore dropped from the file **p_indall_ip**.

Wave 17

Lost household responses

At **Wave 17**, an error in the data collection process resulted in lost household responses for some households.

A scripting error related to face-to-face fieldwork led to some web data being unintentionally overwritten. This occurred where the household grid had been completed online, but individual interviews were incomplete or missing. If such cases were later accessed by a face-to-face interviewer, the data already provided via the web could be lost, as it was overwritten by (blank) face-to-face data.

The issue was not apparent during fieldwork and only became clear during data processing, when inconsistencies were found in cases thought to have complete household data. This problem was particularly relevant for households that were not fully completed online (for example, where individual interviews remained outstanding). Fully complete households were automatically 'locked' by the script, preventing overwriting, but this was not done for partially completed households as access to the household record was necessary for interviewers to complete follow-up work face-to-face.

This issue affected 156 households (5.3% of the issued sample). All were marked in the data returned by the fieldwork agency with a 'data lost' outcome.

These cases are identified in the file **q_hhsamp_ip** via their value of the household level outcome variable, **q_ivfho**. Most (147 cases) have **q_ivfho == 17** 'individual interviews only (no grid)'; the remaining 9 cases have been allocated **q_ivfho == 39** 'lost on laptop' as they have neither a household grid nor an adult interview.

We have been able to partially recreate data for some of these cases using some weekly data flows received during fieldwork. To facilitate prompt delivery of sensors to households consenting to have in-home sensors (see 9.49 Indoor residential environment: consent for in-home sensor) and prompt data collection from smart meters (see 9.50 Domestic energy use: consent for smart meter data linkage), the fieldwork agency sent us weekly data on the outcomes of the relevant consent questions. As we retained these weekly data files, we have been able to reassemble sparse records providing the consent responses in relation to the households where the main household data was lost.

Consent question envsenscons2 asked in CATI mode

At Wave 17, a consent question was inadvertently asked of some telephone respondents. We generally do not ask questions seeking consent for data linkage or to complete additional tasks in the CATI (telephone) mode.

The IP17 household questionnaire included two variants of a consent question to place environmental sensors in people's homes. Households were pre-allocated via variable **ff_esensinfow17** to receive either version 1 (**envsenscons1**) or version 2 (**envsenscons2**) of the question.

Households responding in the CATI mode and allocated **ff_esensinfow17=1** were (correctly) **not** asked **envsenscons1**.

However, households responding in the CATI mode and allocated **ff_esensinfow17=2** were asked **envsenscons2** when the question should have been skipped based on the mode of response.

Inconsistency between mode of allocation variable and mode of completion

At Wave 17, different allocation variables were computed within the questionnaire script (i.e., at the point the respondent was completing the questionnaire) depending on whether the respondent was completing online or face-to-face. Face-to-face respondents were allocated a value of **congrpftf** while web respondents were allocated a value of **congrpweb**. (For further details, see 9.46 Consent decision process.)

There are 12 respondents who have values for **congrpweb** — indicating they completed online — but also have a mode of completion (**indmode**) value describing them as having completed face-to-face. These cases are believed to all be respondents who switched mode at some point during their completion of the survey. They would have completed part of the survey online, receiving an allocation of **congrpweb**, then having not completed the full survey an interviewer would have contacted them and they would have finished the survey face-to-face.

The apparent data inconsistency, therefore, arises because the two variables are indicators of different things. The presence of a value for variable **congrpweb** indicates they completed the survey online up to at least the point where that variable was allocated, while the face-to-face value in **indmode** indicates they finished the survey in the face-to-face mode.

3.8 Stata code for matching files

The syntax examples below show how to perform some common data management tasks useful in analysing the Innovation Panel data files. Each task is illustrated with code for Stata. Statements beginning with `//` are comments. The six tasks are:

- Distributing household level information to individual level
- Summarising individual level information at the household level
- Matching individuals within a household
- Using the egoalt file to create household composition variables
- Merging individual files across waves into long format
- Merging individual files across waves into wide format

Example 1: Distributing household level information to individual level

In this example we will distribute household level information to individuals in those households. We can do this by merging household level file (such as **w_hhresp_ip**) with an individual level file (such as **w_indresp_ip**) within the same wave.

```
// open the household level file
use a_hidp a_hhsize using a_hhresp_ip, clear
// sort it on the household identifier, w_hidp
sort a_hidp
// save this temporary file
save hhinfo, replace
// open the individual level file
use pidp a_hidp a_marstat using a_indresp_ip, clear
// sort it on the household identifier, w_hidp
sort a_hidp
// merge it with the earlier saved file on w_hidp. The output
shows how many cases matched
merge m:1 a_hidp using hhinfo
// drop this variable - essential step
drop _merge
save finall, replace
// clean up unwanted files
erase hhinfo.dta
```

Example 2: Summarising individual level information at the household level

In this example we will summarise individual level information within a household (number of 18-24 year olds in the household) and then match that onto the household level file.

```
use a_hidp a_hhsize using a_hhresp_ip, clear
sort a_hidp
```

```

save hhinfo, replace
use pidp a_hidp a_dvage using a_indall_ip, clear
// create a variable that counts the number of 18-24year olds in
each household
bysort a_hidp: egen n1824= sum(a_dvage>=18 & a_dvage<=24)
// keep only first observation for every household
bysort a_hidp: keep if _n==1
// keep only household level information
keep a_hidp n1824
// now merging this household information with the household level
file
sort a_hidp
merge 1:1 a_hidp using hhinfo
drop _merge
save final2, replace
erase hhinfo.dta

```

Example 3: Matching individuals within a household

In this example we will match the information of wives onto that of their partners/spouses.

```

/* Open the dataset with information on all persons in responding
households and keep only those persons who have a spouse/partner
in the household*/
use a_hidp a_pno a_hgpart a_sex a_dvage using a_indall_ip if
a_hgpart>0, clear
// rename the prefix a_ to something that would indicate that this
information relates to the spouse or partner
renprefix a_ sp_
/* rename the spouse/partner pno variable to the respondent pno
variable as this will be used to match on to the respondent
information. Then sort and save the data*/
rename sp_hgpart a_pno
rename sp_hidp a_hidp
drop sp_pno
sort a_hidp a_pno
save spousepartner, replace
/* Again open the data with information on all persons in
responding households*/
use a_hidp a_pno a_hgpart a_sex a_dvage using a_indall_ip if
a_hgpart>0, clear
/* rename the prefix a_ to something that would indicate that this
information relates to the respondent */

```

```
renpfix a_ r_
/* as we want to match on a_hidp and a_pno rename r_hidp and r_pno
back to these */
rename r_hidp a_hidp
rename r_pno a_pno
// Now sort and merge with the spouse partner file
sort a_hidp a_pno
merge 1:1 a_hidp a_pno using spousepartner
drop _merge
save final3, replace
erase spousepartner.dta
```

Example 4: Using the EGOALT file to create household composition variables

In this example we will create a variable that measures the number of siblings in the household using the **w_egoalt_ip** file.

```
use b_hidp b_epno b_relationship using b_egoalt_ip, clear
// create a variable that counts the number of siblings in the
household
bysort b_hidp b_epno: egen nsiblings = sum(b_relationship>=14 &
b_relationship<=17)
lab var nsiblings "number of siblings in household"
// keep one observation per person
bysort b_hidp b_epno: keep if _n==1
sort b_hidp b_epno
save final4, replace
```

Now this information can be merged with any individual level file.

Example 5: Merging individual files across waves into long format

To match individual level files across *two* waves into a long format do the following (for more waves add wave specific prefix in the foreach statement):

```
foreach w in a b {
    // open the individual level file
    use pidp `w'_jbhas using `w'_indresp_ip, clear
    // drop the wave prefix from all variables
    renpfix `w'_
    // create a wave variable
    gen wave=strpos("`ab", "`w'")
    // save one file for each wave
    save temp`w', replace
}
```

```

}

// open the file for the first wave (wave a_)
use tempa, clear
foreach w in b {
    // append the files for second wave onwards
    append using temp`w'
}
// save the long file
save final5, replace
// erase temporary files
foreach w in a b {
    erase temp`w'.dta
}

```

Example 6: Merging individual files across waves into wide format

The following code shows how to match individual level files across *two* waves into a wide format. The code can be adapted to handle more waves by adding wave specific prefixes in the foreach statement:

```

use pidp a_jbhas using a_indresp_ip, clear
sort pidp
save temp, replace
foreach w in b {
    use pidp `w'_jbhas using `w'_indresp_ip, clear
    sort pidp
    merge 1:1 pidp using temp
    drop _merge
    sort pidp
    save temp, replace
}
save final6, replace
erase temp.dta

```

3.9 Changes to employment variables from Wave 14 and Wave 17

Questions about respondents' employment underwent substantial changes from **Wave 14** of the Innovation Panel and Wave 13 of the Main survey, to better reflect how people now work in the UK. For Waves 1 to 13, questions focused on one main job and one (if any) secondary job. Since Wave 14 of the Innovation Panel (and Wave 13 of the Main survey), respondents have been able to report details about multiple jobs (initially up to 16 jobs, subsequently reduced to up to 10 jobs). These changes and the resultant variables are described in full in the [Main Survey User Guide](#).

From **Wave 17** of the Innovation Panel, there has been a further change to the data format for releasing these variables, although the underlying questions and resultant data have not changed. When the questions were amended at Wave 14 of the Innovation Panel to allow reporting of multiple additional jobs, the resulting variables were initially released as extra columns within the **indresp** file. For most respondents, these columns were empty for all jobs except the first job. To reduce the number of nearly empty columns in the main file, while still making the same data available to data users, from Wave 17 the variables related to second and subsequent jobs have been moved to a separate **employment** file.

Moving the variables to a separate file has also enabled them to be presented in 'long' rather than 'wide' format. Users of the **employment** file will find each row represents a certain job for a certain person. For example, a person reporting 4 jobs will have 4 rows in the **employment** file, with each of those rows containing variables about one of their jobs numbered 1 to 4; the variables related to job 1 will still be reported in the **indresp** file too.

4. Questionnaire

The [Innovation Panel questionnaires](#) are based on the questionnaires for the main Understanding Society survey, with some variations. For each IP wave, the equivalent [main survey wave questionnaires](#) are used as a starting point.

Questionnaires include the household enumeration grid and the household questionnaire (completed by one household member), the individual adult questionnaires (completed by all household members aged 16+) including questions that are administered as a self-completion module for those interviewed face-to-face, and the youth self-completion questionnaire (completed by children aged 10-15).

Questionnaires are divided into topical modules, with approximately half appearing yearly, and the remainder being devoted to topical modules that appear on a rotating basis every few years. In recent waves, many of the rotating modules have been dropped from the IP, to generate space for studies proposed through the annual Innovation Panel competition.

There is minor variation in the names of modules used in the different waves. Also, the content of modules may vary despite having the same name. Researchers should examine the specific questions to ensure comparability across waves. The Questionnaire modules tab within the [Variable search](#) can help with this.

Researchers should also be aware that there are modules that are asked only of new entrants, e.g., the **Initial Conditions** module after Wave 1, or of newly eligible respondents, e.g. respondents who have turned age 16. The universe or characteristics of respondents eligible for a question is indicated in the questionnaire.

The content of the main survey questionnaires is summarised in the [long-term content plan](#).

The additional content of the IP is summarised in the experiments section (see 6. Experiments, methodological studies and non-experimental new content carried in the Innovation Panel).

4.1 How to read the questionnaire

There are multiple resources for learning about the study variables to plan analyses. These include the questionnaires, variable summaries prepared for each data file released, the code books for each data file, and the summary of derived variables.

You can also find the variables you need for your research by using the [variable search](#). Here you can search by variable name, by data file or questionnaire module. This search provides links between questions, variables and data files.

Many of the basic (non-derived) variables can be learned about directly from the questionnaires. The example below shows a marked-up excerpt of the individual questionnaire from IP Wave 3. You can see that although the variable name consists of the listed variable name and the wave prefix, the wave prefix does not appear in the questionnaire. The document also shows the brief variable label, text of the question, source of the question and value labels for the response options. Showcards to help the respondent in answering are also marked as part of the questionnaire. For the Innovation

Panel, it is particularly important to note that experimental groups may receive different versions of questions and that the responses may be recorded in different variables. For details on each experiment, see the relevant experiment sections (8. Questionnaire design experiments: general issues; 9. Questionnaire design experiments: specific topics).

Example from Innovation Panel questionnaire

Jbsat_a. Job satisfaction, FL, CASI

Variable Name and Variable label
Note that wave prefix does not appear

Source
BHPS adapted

Interviewer Instruction
TURN SCREEN AROUND TO RESPONDENT, ASK THEM TO READ THE QUESTION AND ENTER THE NUMBER WHICH REPRESENTS THEIR ANSWER

Text
All things considered, which number best describes how **d**issatisfied or satisfied you are with your present job overall?

Options

1	Completely Dissatisfied
2	Mostly Dissatisfied
3	Somewhat Dissatisfied
4	Neither Satisfied nor Dissatisfied
5	Somewhat Satisfied
6	Mostly Satisfied
7	Completely Satisfied

Response options, value labels

Use
Ask JbSat_A

Modules
Module Jobsatisfaction_ip3. Job satisfaction module

Universe
If (CURRENTEMPLOYMENT.JBHAS = 1 | CURRENTEMPLOYMENT.JBOFF = 1) // Worked in the last week or did not work last week but has a job
And If (ff_jobsatv3 = 1) // Fully labelled scale CASI treatment

Question is from Wave 3 Job Satisfaction module

Question asked of these respondents. Note that it is part of the CASI experiment

The [online data documentation](#) contains sections on:

- Citation – information on how to cite Understanding Society datasets
- Variable search – information about data files and all variables, as well as a search tool to find questions/variables
- Questionnaires – all the questionnaires used for each wave, in PDF format
- Technical reports – PDF files containing an overview of methodology used at each wave
- Fieldwork documents – all materials used in fielding the survey, such as advance letters, consent forms, and showcards

Help and Support for using the survey can be found in the [User support forum](#). After a short registration data users can read past issues, FAQ's and report any issues or queries of their own.

4.2 New studies and questionnaire content carried this wave

Various studies, including experiments and proposals for new survey questions, are submitted via the [Innovation Panel Competition](#) at each wave. Most of these result in new content being carried in the Innovation Panel questionnaire for a single wave.

At Wave 17, the studies carried were:

- 7.28 Youth survey invite mailing
- 9.18 What do the general population regard as “successful ageing”?
- 9.46 Consent decision process
- 9.47 Identification of informal caregiving
- 9.48 Labour market expectations
- An overarching study looking at indoor environments and energy use, comprising two experimental studies and some additional background questions:
 - 9.49 Indoor residential environment: consent for in-home sensor
 - 9.50 Domestic energy use: consent for smart meter data linkage
 - 10.16 Indoor residential environment and energy use: background information
- 10.14 Adaptation to eco-climate emergency
- 10.15 Youth online survey

A full list of studies carried over the waves is included in the tables in 6.1. Overview of Innovation Panel experiments. Each is described in full in sections 7 to 10 (7. Procedural experiments, 8. Questionnaire design experiments – general issues, 9. Questionnaire design experiments – specific topics, 10. Non-experimental studies).

5. Weighting

Weights are provided in order to adjust for differential nonresponse, and for unequal selection probabilities and potential sampling error. Weighted analysis will adjust for response rate differences between subgroups of the sample. The appropriate weight to use will depend on the nature of the analysis being undertaken. Weights should be selected carefully following advice provided below.

If you aim to generalise results to the population of Great Britain, our advice is to always use weights.

5.1 Weights for analysis of Wave 14 and later

Innovation Panel data should be analysed taking account of its sample design. This is achieved by specifying clustering (W_psu) and stratification (W_strata) and a weight. Such specification ensures representation of the population and/or inference to other similar surveys.

A number of weights are provided for users of IP data corresponding to different instruments and waves of use. From Wave 14 onwards, two main types of weights are provided: cross-sectional weights related to the current wave, and issue weights which can be used to create your own longitudinal weight tailored to your analysis.

To study the effect of an experiment or calculate an estimate in one wave, a cross-sectional weight should be used. The specific cross-sectional weight would depend on the origin of the variables used in the analysis:

Origin of variables used in analysis	Weight for cross-sectional analysis
All variables used are from the household enumeration grid (INDALL file)	w_psnenip_xw
All variables used are from the household enumeration grid (INDALL file) and/or the household questionnaire (HHRESP file) and your analysis is at an individual (not household) level	w_psnenip_xw
All variables used are from the household questionnaire (HHRESP file) and your analysis is at a household level	w_hhdenip_xw
All variables used are from the proxy questionnaire (INDRESP file) and possibly from the household enumeration grid (INDALL file) and household questionnaire (HHRESP file)	w_indpxip_xw
At least one of your variables is from an individual adult questionnaire (INDRESP and associated files) and possibly also from the household enumeration grid (INDALL file) and household questionnaire (HHRESP file)	w_indinip_xw
At least one variables is from an individual self-completion questionnaire (INDRESP file) and possibly also from the	w_indscip_xw

household enumeration grid (INDALL file) and the household questionnaire (HHRESP file)	
Variables from the youth self-completion questionnaire (YOUTH file)	w_ythscip_xw

If your analysis includes multiple time points or waves of Innovation Panel interviews you should use a longitudinal weight. From Wave 14 onwards we do not provide specific longitudinal weights. Instead, we provide starting points for longitudinal weights, called issue weights (**w_psneni#.li**). The issue weights are provided at an enumeration level and are created for each refreshment (including those before wave 14). It takes into account selection probabilities and corrects for nonresponse up until and including refreshment year, and joins all the samples at the time of refreshment. It is expected that a user creates their own longitudinal weight starting with our issue weight and corrects for attrition themselves tailoring it to their own analysis model. See below for details on how to create your own longitudinal weight.

Creating your own longitudinal weight

Longitudinal analysis assumes using multiple waves of Innovation Panel data. Follow these steps to create your own longitudinal weight. We provide an example analysing data from the self-completion questionnaire in waves 5 and 14.

1. Choose a base weight.

For this think of the earliest wave that your analysis includes. Use the inclusion weight from this wave (if it is present) or closest earlier wave.

Example: for an analysis of waves 5 and 14, the earliest wave is 5. Wave 5 does not have an inclusion weight, but the earlier closest wave that has it is wave 4. Use the inclusion weight from wave 4 (**d_psnenip_li**) as your starting point.

2. Create a response indicator.

The response indicator is a 0/1 variable with 0 indicating lack of response and 1 response. Set up your model – those sample members who have valid information in your model (are included in your model) are respondents (value 1). All other sample members should have a value of 0. Assign missing values to those who have become ineligible for the survey (died or moved abroad) since the wave of the issue weight you are using.

Example: for our example you would use the self-completion questionnaire from wave 5 and 14. Thus those who responded to the self-completion at both wave 5 and 14 will have a value of 1 for the response indicator (let's call it "resp").

3. Create predictors for your weighting model.

Use the INDALL file from the same wave as your inclusion weight to create your predictors that will be used to correct for nonresponse. Choose variables that are related to the subject that you study and potentially related to nonresponse. Use as many variables as you can but avoid multicollinearity. Keep only significant predictors in the model. Make sure there is no missingness in predictors.

Example: predictors would come from the wave 4 INDALL file (most recent refreshment before wave 5).

4. Predict response.

In order to correct for remaining attrition you need to predict probability to respond between the issue weight at the enumeration stage and the last wave in your analysis. The easiest way to do this is through a logistic regression where the response variable created earlier is dependent variable in the model. Use the predictors created earlier. Make sure your model is weighted by the inclusion weight.

Example: In Stata run the following model:

```
logistic resp predictor1 predictor2 ... [pw=d_psnenip_li]
predict prob1 // this part saves predicted probabilities to respond (have valid
information in your model) conditional on enumeration at wave 4
replace prob1=0 if resp==0
```

5. Create your weight.

In order to create your weight multiply the predicted probability by your inclusion weight.

Example: in Stata:

```
gen myweight=d_psnenip_li*prob1
```

For more information take a look at our online training material [Creating tailored weights for UKHLS](#). This page provides details on [how to enrol](#).

5.2 Weights for analysis of waves before Wave 14

Selecting the correct weight for your analysis

A number of different weights are provided to meet different needs of users. The weight for your analysis reflects the survey instrument(s) which is/are the source of the data, the analysis level (household or individual) and the wave(s).

All weights follow a naming convention, designed to help users to pick the correct weight. The name of each weight reflects the wave for which the weight is calculated, level of analysis, data source and its nature (design weight, cross-sectional analysis weight or longitudinal analysis weight). The rules are described in the “Naming Conventions for Weighting Variables” section below.

We have presented variable names and segments of variable names in bold so they stand out from the text.

If your analysis uses only data from one wave, select the “**xw**” (cross-sectional) version of the weight. This weight is defined for all sample members¹ who responded to the relevant survey instrument at that wave. If your analysis uses data from two or more waves, select the “**lw**” (longitudinal) version of the weight for the most recent wave included in your analysis. This weight is defined for sample members who responded to the relevant survey instrument at each wave.

For individual level analysis you may want to combine information from different questionnaire sources. In this situation, please select the weight suitable for the lowest level according to the hierarchy below:

Level of analysis	Questions available for
4	Household level (all enumerated individuals)
3	Proxy and main adult interview
2	Main adult interview only (no proxy)
1	Self-completion interview, adult or youth

For example, if in one model you use wave 1 data from the adult main and proxy interview as well as from the self-completion, then the correct weight will be **a_indscip_xw** – the weight for the self-completion questionnaire as its level (1) is lower than the level for main and proxy interview (3).

Table: List of available weights for the Innovation Panel

Analysis level	Wave(s)	Data source	Analysis weight
HH	<i>n</i>	HH grid and/or HH interview	<i>n_hhdenip_xw</i>
IND	<i>n</i>	HH grid and/or HH interview	<i>n_psnenip_xw</i>
IND	1 through <i>n</i>	HH grid and/or HH interview	<i>n_psnenip_lw</i>
IND	4 through <i>n</i>	HH grid and/or HH interview	<i>n_psneni1_lw</i>
IND	7 through <i>n</i>	HH grid and/or HH interview	<i>n_psneni2_lw</i>
IND	10 through <i>n</i>	HH grid and/or HH interview	<i>n_psneni3_lw</i>
IND	11 through <i>n</i>	HH grid and/or HH interview	<i>n_psneni4_lw</i>
IND	1&3	HH grid and/or HH interview	<i>c_psnenip5_lw</i>
IND	1&3&4	HH grid and/or HH interview	<i>d_psnenip13_lw</i>
IND	<i>n</i>	Adult main and proxy interview	<i>n_indpxip_xw</i>
IND	1 through <i>n</i>	Adult main and proxy interview	<i>n_indpxip_lw</i>
IND	4 through <i>n</i>	Adult main and proxy interview	<i>n_indpxi1_lw</i>

¹ With just a few exceptions, which are documented below in the “Technical Details” section

IND	7 through n	Adult main and proxy interview	$n_indpxi2_lw$
IND	10 through n	Adult main and proxy interview	$n_indpxi3_lw$
IND	11 through n	Adult main and proxy interview	$n_indpxi4_lw$
IND	1&3	Adult main and proxy interview	$c_indpxip5_lw$
IND	1&3&4	Adult main and proxy interview	$d_indpxip13_lw$
IND	n	Adult main interview	$n_indinip_xw$
IND	1 through n	Adult main interview	$n_indinip_lw$
IND	4 through n	Adult main interview	$n_indini1_lw$
IND	7 through n	Adult main interview	$n_indini2_lw$
IND	10 through n	Adult main interview	$n_indini3_lw$
IND	11 through n	Adult main interview	$n_indini4_lw$
IND	1&3	Adult main interview	$c_indinip5_lw$
IND	1&3&4	Adult main interview	$d_indinip13_lw$
IND	n	Adult self-completion	$n_indscip_xw$
IND	4 through 7	Adult self-completion	$n_indsci1_lw$
IND	1&3	Adult self-completion	$c_indscip5_lw$
IND	1&3&4	Adult self-completion	$d_indscip13_lw$
IND	1&3&4&5	Adult self-completion	$e_indscip29_lw$
IND	1&3&4&5&6	Adult self-completion	$f_indscip61_lw$
IND	1 & 3 to 7	Adult self-completion	$g_indscip125_lw$
IND	1	Youth self-completion	$a_ythscip_xw$
IND	1&2	Youth self-completion	$b_ythscip_lw$
IND	n (w3 onwards)	Youth self-completion	$n_ythscip_xw$
IND		(Design weight)	$a_psnenip_xd$

HH = household; IND = individual

5.3 Naming conventions for weighting variables

Naming conventions have been adopted for the weighting variables. This will help users to select the name of the weight they need or to identify the nature of a weight. The structure is as follows, and is consistent with (a subset of) the naming conventions for the main survey:

$w_xxxyyzz_aa$, where

Wave	Target population	Instrument	Sample	Type of weight
w_	xxx	yy	zz	_aa
a_ b_ c_ ... o_ p_ q_	hhd: household psn: persons 0+ ind: persons 16+ yth: persons 10-15	en: enumeration (grid) in: interview px: interview or proxy sc: self-completion bp: blood pressure (only in IP12) bm: height and weight (only in IP12) hs: hair sample (only in IP12) bs: dry blood spot (only in IP12) fb: full blood sample (only in IP12)	ip: Innovation Panel i1: Innovation Panel since wave 4 (2011), including IP4 refreshment i2: Innovation Panel since wave 7 (2014), including IP4 and IP7 refreshments i3: Innovation Panel since wave 10 (2017) i4: Innovation Panel since wave 11 (2018) i5: Innovation Panel since wave 14 (2021)	xw: cross-sectional analysis weight li: issue weight lw: longitudinal analysis weight xd: cross-sectional design weight

W=wave
 XXX=target population
 YY=instrument
 ZZZ=sample
 AA=weight type

Target population (XXX):

HHD: household
 PSN: persons 0+
 IND: persons 16+
 YTH: persons 10-15

Instrument (YY):

EN: enumeration (grid)
 IN: interview
 PX: interview or proxy
 SC: self-completion

BP: blood pressure (only in IP12)
BM: height and weight (only in IP12)
HS: hair sample (only in IP12)
BS: dry blood spot (only in IP12)
FB: full blood sample (only in IP12)

Sample (ZZ):

IP: Innovation Panel

I1: Innovation Panel since wave 4 (2011), including IP4 refreshment

I2: Innovation Panel since wave 7 (2014), including IP4 and IP7 refreshments

I3: Innovation Panel since wave 10 (2017)

I4: Innovation pPpanel since wave 11 (2018)

I5: Innovation Panel since wave 14 (2021)

Type of weight (AA):

XW: cross-sectional analysis weight

LI: issue weight

LW: longitudinal analysis weight

XD: cross-sectional design weight

Example

a_indinip_xw is the cross-sectional analysis weight for adult main interview data from IP wave 1, representing the population of persons aged 16+.

Longitudinal weights for partial wave sets: An additional device is used to indicate weights for longitudinal analysis of combinations of waves that do not include all waves up to the current one ("partial wave sets"). This consists of a numeric indicator following the "ZZ" part of the variable name. The number is a decimal representation of the binary number that indicates the combination of waves, where 1 indicates inclusion and 0 indicates exclusion, and where the waves are in reverse order. For example, data from waves 1, 3 and 4 can be represented by the binary number 1101, which translates to the decimal number 13 (1+4+8). Thus, **d_indinip13_lw** is the weight for individual interview data from waves 1, 3 and 4 (whereas **d_indinip_lw** is the weight for the "complete wave set" of waves 1, 2, 3 and 4). For ease, the variable label indicates the combination of waves, e.g. "(acd)" to indicate waves 1, 3 and 4.

The partial wave sets presented with this release are 5 (Wave 1 and Wave 3) and 13 (Waves 1, 3, and 4), 29 (Waves 1,3,4,5), 61 (Waves 1,3,4,5,6), and 125 (Waves 1,3,4,5,6,7).

5.4 Technical details

Weights for Wave 14 onwards

Issue weight

With each refreshment sample we create a new issue weight at an enumeration level. This weight consists of two parts: a cross-sectional weight for the new boost sample, for which participation in the panel is for the first time in that wave, and a longitudinal weight for the continuing part of the sample.

Cross-sectional part

Refreshment (boost) samples are equal probability samples for Great Britain. Thus no correction for unequal probabilities is required. In order to correct for nonresponse at a household level we obtain predictors from our sampling frame, from interviewer observations in case the boost was administered fully by face-to-face interviewers, and from outside information. The outside information is obtained from official statistics resources (e.g. Census, ONS) and is linked to our dataset based on LSOA level. The predictors are checked for missingness and imputation is used where required. Response at a household level is then predicted using logistic regression. The inverse probabilities of response are then applied to everyone who was enumerated in responding households. This weight is then scaled to the mean of 1.

Longitudinal part

The continuing part of the panel requires correction for attrition between the previous boost wave and the current boost wave. Predictors from the household level questionnaire, household grid and enumeration information are used to predict this attrition. The inverse probabilities are then computed. To create a longitudinal weight these inverse probabilities are multiplied by the issue weight from the previous boost. Children born between the two boosts are given their mother's longitudinal weight. The longitudinal weight is then scaled to the mean of 1. For further information take a look at our online training material [Creating tailored weights for UKHLS](#).

Combining longitudinal and cross-sectional parts

A new weight is created by giving a longitudinal weight to the continuing panel members and cross-sectional weight to the new panel members selected through the boost. The cross-tabulation of age, gender and region is then checked against ONS mid-year equivalent estimates and post-stratification is applied to match IP data to the population statistics. The final weight is then scaled to the mean of 1.

Cross-sectional weights

All cross-sectional weights are created in the same way using the most recent issue weight as a base weight. The enumeration nonresponse between issue weight and the relevant instrument is predicted using logistic regression. Predictors come from the household grid, household interview and enumeration information in the recent refreshment wave. The inverse response probabilities are multiplied by the issue weight. The resulting weight is then shared to TSMs through a weight share procedure. After this those who still have no valid weight value are given it through a nonzero weight-share method. The final weight is then scaled to the mean of 1.

Weights before Wave 14

This section describes how the following weights were derived:

- Cross-sectional weights for each wave;
- Longitudinal weights for complete wave sets up to and including each wave from IP2 onwards;
- Longitudinal weights for partial wave sets.

Common features of all weights

Note that all models used to predict response propensities as described in the Technical Details are fitted using stepwise backward logistic regression with $p=0.05$. Each set of weights has been scaled by a constant factor to produce a mean of one amongst cases eligible to receive the weight. In consequence, weights which are defined as equal (e.g. D_HHDENIP_XW AND D_PSNENIP_XW for the refreshment sample) will not necessarily have the same numeric value, but will retain the same between-person/household relative value.

Wave 1 cross-sectional weights

a_hhdenip_xw

a_psnenip_xw

a_indinip_xw

a_indpxip_xw

a_indscip_xw

a_ythscip_xw

Each IP1 cross-sectional weight consists of a design weight, which is adjusted for non-response and post-stratified to population estimates.

The design weight is the same for a household (**a_hhdenip_xd**) and for each member of the household (**a_psnenip_xd**). Design weights are equal for the vast majority of sample members and differ only for cases that involved a) sub-sampling of dwellings because there were more than three dwellings at the address, or b) sub-sampling of households because there were more than three households in the dwelling (see section 4 on sampling).

The first stage of non-response adjustment is at the household level. The adjustment consists of the reciprocal of predicted values from a logistic regression model of household response, where this is defined as completion of at least the household grid (around 59% of households responded in IP1). The covariates in the model were a set of small area indicators and Census 2001 variables, including those used in sample stratification, as well as interviewer observation variables collected during the survey field work. The household weight consists of the design weight with this household-level adjustment. This weight (**a_hhdenip_xw**) is defined for all households that participated in IP1.

The enumerated person weight (**a_psnenip_xw**) is defined for all persons in households that participated in IP1. The weight equals the IP1 household weight, post-stratified by age, sex, and grouped Government Office Region. The post-stratification targets are taken from Office for National Statistics 2008 mid-year population statistics. The post-stratification adjustments were calculated for each cell of a 56-cell matrix, as the ratio of population count to weighted sample count. The cells were defined by seven age categories, sex, and four (groups of) regions/countries.

The adult main interview weight (**a_indinip_xw**) is defined for the 85% of IP1 enumerated persons aged 16 or over who completed the individual interview. It consists of the enumerated person weight, adjusted by the reciprocal of predicted values from a logistic regression model of adult main interview response, conditional on enumeration. The covariates in the model were as described above for household response, plus personal

characteristics from the household grid and a limited number of household characteristics from the household grid and household questionnaire. After this non-response adjustment was applied, post-stratification was implemented, based upon the same 56-cell matrix described above for the enumerated person weight.

The adult main or proxy interview weight (**a_indpxip_xw**) is defined for the 91% of IP1 enumerated persons aged 16 or over for whom either an adult main interview or a proxy interview was completed. It was derived in the same way described above for the adult main weight, the only difference being the definition of response in the logistic regression model.

The adult self-completion weight (**a_indscip_xw**) is defined for all respondents to the adult main interview who also completed the self-completion questionnaire. (A small number of persons 16 or over completed the self-completion questionnaire but not the adult main interview – these responses are not included in the data set.) The weight was derived in the same way described above for the adult main weight, but the logistic regression model is based on all enumerated persons aged 10 or over and predicts response to the self-completion (adult or youth) questionnaire.

The youth self-completion weight (**a_ythscip_xw**) is defined for all persons aged 10 to 15 who completed the youth questionnaire. The weight was derived in an identical way to that described above for the adult self-completion weight.

Wave 2 longitudinal weights

b_psenip_lw
b_indinip_lw
b_indpxip_lw
b_ythscip_lw

The IP2 longitudinal enumerated person weight (**b_psenip_lw**) is defined for all Original Sample Members (OSMs) enumerated at both IP1 and IP2 – that is, in a responding household at both waves, plus newborns (children of OSM mothers, born between IP1 and IP2). It consists of the IP1 enumerated person weight, adjusted for conditional non-response at IP2. The adjustment consists of the reciprocal of predicted values from a logistic regression model of IP2 enumeration conditional on IP1 enumeration. The covariates were the same as those described above for the IP1 adult main interview weight. Newborns were assigned the same weight as their mother.

The IP2 longitudinal adult main interview weight (**b_indinip_lw**) is defined for all OSMs who completed the adult main interview at both IP1 and IP2 and for 16 year-old OSMs who completed the interview at IP2 but were too young to be eligible for it at IP1. It consists of the IP1 adult main interview weight (or the IP1 enumerated person weight in the case of IP2 16 year-olds), adjusted for conditional non-response at IP2. The adjustment consists of the reciprocal of predicted values from a logistic regression model of IP2 adult main interview response conditional on IP1 adult main interview response. The covariates were measures from the IP1 adult main interview and the IP1 household grid and household interview, plus interviewer observations.

The IP2 longitudinal adult main or proxy interview weight (**b_indpxip_lw**) is defined for all OSMs who at both IP1 and IP2 *either* completed the adult main interview *or* had a proxy interview carried out on their behalf, and for 16 year-old OSMs who completed either instrument at wave 2 regardless of their response in IP1 as they were too young to be eligible for the adult main interview at IP1. It should therefore be used in preference to the IP2 longitudinal adult main interview weight for analysis which is restricted to variables that are available from the proxy interview. The weight was derived by adjusting the IP1 adult main or proxy interview weight (or the IP1 enumerated person weight, in the case of IP2 16 year-olds) for non-response at IP2. The adjustment came from a model of IP2 response (main or proxy) conditional on IP1 response (main or proxy). Model covariates were the same as described above for the longitudinal adult main weight, except that variables from the individual interview were restricted to those also included in the proxy interview.

The IP2 longitudinal youth self-completion weight (**b_ythscip_lw**) is defined for all persons who completed the youth self-completion questionnaire at both IP1 and IP2 and for 10-year-old OSMs who completed it in IP2 but were too young to have been eligible for it at IP1. The weight was derived by adjusting the IP1 youth weight (and the IP1 enumerated person weight, in the case of 10 year-olds at IP2) for non-response to the youth self-completion at IP2. The adjustment came from a model of IP2 youth response conditional on IP1 youth response.

Wave 2 cross-sectional weights

b_hhdenip_xw
b_psnenip_xw
b_indinip_xw
b_indpxip_xw

The IP2 cross-sectional enumerated person weight (**b_hhdenip_xw**) is defined for all persons (OSMs and Temporary Sample Members (TSMs)) enumerated at IP2. For persons in households where all household members are OSMs, it is equal to the longitudinal enumerated person weight. For persons in households with at least one TSM at wave 2, it is derived through the weight share method: each person in the household is given a weight of a/b , where a is the sum of the longitudinal enumerated person weights for all OSMs in the household and b is the total number of persons (OSMs and TSMs) in the household.

The IP2 cross-sectional household weight (**b_hhdenip_xw**) is defined for all households who responded at IP2. It consists simply of the mean² of the IP2 cross-sectional enumerated person weights for all persons (OSMs and TSMs) in the household.

The IP2 cross-sectional adult main interview weight (**b_indinip_xw**) is defined for all persons who completed the IP2 adult main interview (with one exception, noted in the following). In households containing one or more responding TSMs, each respondent in the household is given a weight of a/b , where a is the sum of the longitudinal weights (**b_indinip_lw**) in the household and b is the total number of respondents in the household who are either an OSM with a non-zero longitudinal weight or a TSM. In all other households (OSM-only households and households in which no TSMs completed the main interview), the cross-sectional weight is equal to the longitudinal weight. Note, that **b_indinip_xw** will equal zero for all persons in the household if no persons have a non-zero **b_indinip_lw**. This can happen, for example, if the only person completing the IP2 interview is a TSM or is an OSM who did not complete the IP1 interview.

The IP2 cross-sectional adult main or proxy weight (**b_indpxip_xw**) is defined for all persons who *either* completed the adult main interview *or* had a proxy interview carried out on their behalf. It is derived in a way exactly analogous to that described above for the cross-sectional adult main interview weight, but based upon **b_inpxnip_lw** instead of **b_indinip_lw**. Thus, there will be weights of zero for persons in households in which no person has a non-zero value of **b_inpxnip_lw**.

² On the mainstage survey, and for subsequent waves of IP, cross-sectional household weights were defined as the minimum of the cross-sectional enumerated person weights, rather than the mean. This was done on the grounds that the probability of a household being enumerated is equal to or greater than the probability of the highest-probability individual in the household being enumerated, so variation between households in the maximum probability may better reflect household probability than variation in the mean probability.

Wave 3 through 9 longitudinal weights for complete wave sets

n_psnenip_lw

n_indinip_lw

n_indpxip_lw

At each wave subsequent to IP2, the basic longitudinal enumerated person weight (***n_psnenip_lw***) is defined for all persons who were enumerated at all waves up to and including the current one. It consists of the enumerated person longitudinal weight from the previous wave, with an adjustment for nonresponse at the current wave. The adjustment consists of the reciprocal of the predicted probability of enumeration at the current wave conditional on enumeration at all previous waves. The probability is predicted by a model based on OSMs only in which covariates come from the previous wave household grid and household questionnaire. Newborns are assigned the enumerated person longitudinal weight of their biological mother.

At each wave subsequent to IP2, the longitudinal adult main interview weight (***n_indinip_lw***) is defined for all OSMs who at each wave up to and including the current one completed the interview (or were aged under 16 but were continuously enumerated)³. The weight is derived as follows. First, for all OSMs who had a non-zero longitudinal adult main interview weight at the previous wave, a model is fitted to predict completion of the main interview at the current wave. The reciprocal fitted values are multiplied by the longitudinal adult main interview weight from the previous wave. Second, for all OSMs who have reached the age of 16 since the previous wave, reciprocal fitted values from a model of main interview response at current wave conditional on enumeration at current wave are multiplied by the current wave longitudinal enumerated person weight. The model is based on all OSMs aged 16 or over at the current wave, but the predicted values used only for those aged exactly 16. The weights for 16 year-olds are then scaled so that when combined with those for persons aged 17 or over, the weighted proportion of 16-year olds is equal to that obtained by applying the current wave longitudinal enumerated person weight to all enumerated persons aged 16 or over.

Analogously, at each wave subsequent to IP2, the longitudinal adult main or proxy interview weight (e.g. ***c_indpxip_lw***) is defined for all persons for whom either a main or proxy interview was completed at all waves up to and including the current one at which they were eligible (i.e. aged 16 or over). The weight is derived in exactly the same way described above for main interview weights, except that the models predict response to main or proxy interview and that, for persons aged 17 or over, the reciprocal predicted values are multiplied by the previous wave longitudinal adult main or proxy interview weight.

³ Thus, an OSM who turns 16 and subsequently responds to the adult main interview will have a longitudinal adult main interview weight even though they were not eligible to have a longitudinal adult main interview weight at any previous wave. For example, a sample member who responded to the main interview at IP3 and IP4, but was aged 14 at IP1 and 15 at IP2, will have a non-zero value of D_INDINIP_LW.

Wave 3 and 4 longitudinal weights for partial wave sets

c_psnenip5_lw
d_psnenip13_lw
c_indinip5_lw
d_indinip13_lw
c_indpxip5_lw
d_indpxip13_lw
c_indscip5_lw
d_indscip13_lw

Longitudinal weights are also provided for the partial wave sets {1, 3}, which is denoted as 5 in the weight variable names, and {1, 3, 4}, which is denoted as 13 in the weight variable names. Note that wave 2 was carried out using mixed modes and, as a consequence, there are a sizeable number of wave non-respondents at that wave.

These longitudinal weights for partial wave sets were derived using methods directly analogous to those for the longitudinal weights for complete wave sets. For example, the longitudinal enumerated person weight for waves 1, 3 and 4 (**d_psnenip13_lw**) consists of the longitudinal enumerated person weight for waves 1 and 3 (**c_psnenip5_lw**), with an adjustment for nonresponse at the wave 4 conditional on response at waves 1 and 3. For weights that involve separately modelling those who were already eligible at the previous wave in the set and those who have become eligible subsequently (**indinip**, **indpxip**, **indscip**), there will be a larger proportion of sample members defined as newly eligible whenever the previous wave in the set was two or more years previously. For example, creation of **c_indinip5_lw** (waves 1 and 3) involved one model for persons already eligible and responding at wave 1 (to adjust their wave 1 main interview weight) and another for persons who became eligible for the main interview over the 2-year period between waves 1 and 3 (to adjust their wave 3 enumerated person weight).

Note that the only longitudinal weights for the adult self-completion questionnaire are for the partial wave sets {1, 3} and {1, 3, 4} as the self-completion questionnaire was not administered at wave 2.

As with the other longitudinal weights, sample members who had reached age 16 since the previous wave in the set are assigned a longitudinal weight based on their current wave enumerated person weight, adjusted for non-response to the current wave self-completion questionnaire conditional on enumeration.

Wave 3 cross-sectional weights

c_hhdenip_xw
c_psnenip_xw
c_indinip_xw
c_indpxip_xw
c_indscip_xw
c_ythscip_xw

The cross-sectional enumerated person, main interview, and main or proxy interview weights (**c_psnenip_xw**, **c_indinip_xw**, **c_indpxip_xw**) are each based on the respective longitudinal weight (**c_psnenip5_lw**, **c_indinip5_lw**, **c_indpxip5_lw**), with application of the weight share method to assign weights to TSMs. In households containing one or more responding (to the relevant instrument) TSMs, each respondent in the household is given a weight of a/b , where a is the sum of the relevant longitudinal weights in the household and b is the total number of respondents in the household who are either an OSM with a non-zero longitudinal weight or a TSM. In other households, the cross-sectional weight equals the respective longitudinal weight. Some respondents will therefore receive a zero cross-sectional weight, namely OSMs with a zero longitudinal weight and TSMs in households containing no OSMs with a non-zero longitudinal weight.

The IP3 household weight (**c_hhdenip_xw**) equals the minimum of the cross-sectional enumerated person weights (OSMs and TSMs) in the household.

The cross-sectional adult self-completion and youth self-completion weights (**c_indscip_xw**, **c_ythscip_xw**) are based on the wave 3 cross-sectional enumerated person weight, with an adjustment for non-response to the self-completion questionnaire. The adjustment is derived from a single model of response to (either) self-completion questionnaire, based on all enumerated persons aged 10 or over at wave 3 (who were therefore eligible for either the youth or adult self-completion questionnaire).

Wave 4 cross-sectional weights

d_hhdenip_xw
d_psnenip_xw
d_indinip_xw
d_indpxip_xw
d_indscip_xw
d_ythscip_xw

All IP4 cross-sectional weights combine the continuing sample with the refreshment sample, for whom IP4 was the first wave. Weights were derived in different ways for the two samples.

For the refreshment sample, weights were developed in the same way as for IP1, consisting of a design weight adjusted for non-response. The design weight corrected for multiple

dwellings or households at an address. The non-response adjustment is based on covariates from Census 2001 and small area statistics. A separate model was used for each of England, Wales and Scotland as different predictors were available for each country. Adjustments for households in England were based on a model for England only; adjustments in Wales were derived from a model for England and Wales; adjustments in Scotland came from a model for all three countries using only common predictors. (The refreshment sample sizes in Wales and Scotland were too small to support separate models.) These weights serve both as household weights (**d_hhdenip_xw**) and enumerated person weights (**d_psnenip_xw**) for the refreshment sample. The enumerated person weight was then post-stratified by gender in London and by gender and five age groups for the rest of Great Britain, based on Office for National Statistics mid-2010 population estimates (the most recent available at the time the weights were derived).

Conditional on enumeration, response to the adult main interview, adult main or proxy interview, adult self-completion and youth self-completion were each modelled using stepwise backward logistic regression. Again, separate models were fitted for England, Scotland and Wales. The covariates used were country-specific and came from neighbourhood statistics, household questionnaire and household grid. The inverse predicted response probability was multiplied by the enumerated person weight to produce the respective weight. Post-stratification was applied for each of the three weights (adult main, adult main or proxy, self-completion – where the latter includes both adults and youth) with the same categories as used for enumerated person post-stratification.

All resulting weights were scaled to a mean of one within refreshment sample.

For the continuing sample, weights were created using the same procedure as for the IP3 cross-sectional weights. The source weights were the respective wave 1, 3 and 4 longitudinal weights (e.g. **d_indinip13_lw** in the case of **d_indinip_xw**). As for IP3, self-completion weights consisted of the cross-sectional enumerated person weight with a non-response adjustment derived from a model of response to the self-completion conditional on enumeration. All resulting weights were scaled to a mean of one within the continuing sample.

Note that no adjustment is made for the fact that the refreshment sample, unlike the continuing sample, in principle includes immigrants to Great Britain since 2008. Under the assumption of ignorable immigration, each of the two weighted samples should represent the 2011 population. Using both together will maximize the statistical precision of cross-sectional population estimates. For this reason, each cross-sectional weight as described above is released as a single variable encompassing both samples.

Self-completion longitudinal weight between waves 5 and 7

A self-completion weight (**e_indscip29_lw**) is provided for the longitudinal analysis of waves 1, 3, 4 and 5. This reflects the omission of self-completion instrument in the second wave of IP. The weight is the product of two parts, the first being self-completion longitudinal weight from wave 4 (**d_indscip13_lw**) and the second being an adjustment for nonresponse to the wave 5 self-completion conditional on self-completion response in all previous waves, but wave 2. The number 29 reflects the decimal form of the binary representation of the cross-wave response pattern, i.e. 11101 (for waves 5, 4, 3, 2, and 1 respectively), as described above.

In IP6 and IP7, to reflect that IP2 did not have a self-completion component, a separate weight (**f_indscip61_lw**, **g_indscip125_lw**) is created for those panel members who completed the self-completion questionnaire in all eligible waves, i.e. waves 1, 3, 4, 5, 6 and 1, 3, 4, 5, 6, 7, respectively. These weights were created in the same way as weight **e_indscip29_lw**, described above, and use this weight as the base.

The self-completion longitudinal weight is discontinued at IP8 to reflect the variety of modes used through the time of IP. The cross-sectional self-completion weight is provided and its calculation is described below.

Longitudinal weights incorporating Wave 4 refreshment sample

Starting at wave 5 we provide additional longitudinal weights that include the IP4 refreshment sample. These weights are appropriate for any longitudinal analysis of data collected from wave 4 (2011) onwards.

The method for adjusting for attrition between waves 4 and 5 is identical to the one described in the section on longitudinal weights for complete wave sets, with the exception of the base weight. While the weights for complete wave sets do not include refreshment, the weights described here do.

The base weight for enumeration is equal to **d_psnenip13_lw**, scaled to the mean of 1, for the original part of the sample; and is equal to **d_psnenip_xw** scaled to the mean of 1 for the 2010 refreshment part of the sample. Conditional on nonzero value for the base weight, the nonresponse between wave 4 and 5 is modelled. The newborns are assigned the weight of their biological mother. The resulting weight, scaled to the mean of 1, is called **e_psneni1_lw**.

Similarly, individual response weights (**e_indpxi1_lw**, **e_indini1_lw** and **e_indsci1_lw**) are calculated. The base weight is **d_indpxip13_lw** and **d_indinip13_lw** for the original sample and **d_indpxip_xw** and **d_indinip_xw** for the 2010 refreshment sample (for proxy or main interview and only main interview weights respectively). The attrition correction is identical to the one used for the complete wave set weights, reflecting new base weights. Nonresponse correction for 16-year-olds is also identical to the one for complete wave set weights with the exception of the enumeration base weight being **e_psneni1_lw**.

For IP6, IP7, IP8, and IP9 longitudinal weights are created to enable analysis including the IP4 refreshment sample. The weights are created in an identical way to the longitudinal weights for complete wave sets with the difference in base weight. The base weight is the IP5 longitudinal weight for the combined (original plus IP4 refreshment) sample (e.g. **e_psneni1_lw**, **e_indpxi1_lw**, **e_indini1_lw** or **e_indsci1_lw**).

Longitudinal weights from Wave 7 onwards

We provide a number of different longitudinal weights with a new series of longitudinal weights starting with each refreshment. For longitudinal analysis that starts with IP1 the user should use '**n_XXXXXip_lw**' weight, for analysis that starts with IP4 (that also includes IP4 refreshment) – '**n_XXXXXi1_lw**', with IP7 – '**n_XXXXXi2_lw**', with IP10 – '**n_XXXXXi3_lw**' and so on.

When a new refreshment sample joins the panel the weight for the refreshment sample is calculated in a similar fashion to the first wave weight. The design weight is set as 1 – reflecting an equal probability sample design. The household nonresponse is then corrected using a number of geographically linked variables to the LSOA codes of the sampled

households. The information comes from Census and a number of other government statistics for England and Wales. The nonresponse correction is derived as an inverse of the predicted probabilities from a stepwise logistic regression that predicts enumeration. Because of the small sample size in Scotland we do not run a separate model for Scotland in IP, but give them an average nonresponse correction. Enumerated individuals receive the weight of their household. The household enumeration weight for the refreshment is post-stratified by country, and the person enumeration weight is poststratified by age (5 categories) and sex to match government statistics. This weight is then scaled to the mean of 1 and represents a base weight for the refreshment which is available only for OSM members.

A new refreshment that starts at wave n obtains a longitudinal weight from wave $n+1$. For this the most inclusive longitudinal weight at wave n that includes all the previous refreshments is joined to the base weight for the refreshment (both are scaled to the mean of 1 beforehand). The person enumeration response at wave $n+1$ is then modelled conditional on the non-zero value of the joined weight to obtain a new longitudinal person enumeration weight with the subscript 'im' (i1 stands for weights that include IP4 refreshment, i2 includes additionally IP7 refreshment, i3 includes IP10 refreshment, and i4 includes IP11 refreshment). The proxy and individual response weights that include the most recent refreshment are then modelled: the outcome is a proxy (or main) response in both waves n and $n+1$ (except for the 16- to 18-year-olds who need to respond only in wave $n+1$) conditional on the longitudinal enumeration weight using predictors from the current wave household grid and household questionnaire. The inverse of the predicted probabilities is then multiplied by the enumeration base weight to create longitudinal weights that include refreshment: $n+1_indpxin_lw$ and $n+1_indinin_lw$.

Cross-sectional weights for Wave 5 onwards

All IP5 and IP6 cross-sectional weights combine the original (IP1) sample with the IP4 refreshment sample. All IP7, IP8 and IP9 cross-sectional weights combine the original (IP1) sample, IP4 refreshment sample and IP7 refreshment sample. The IP10 cross-sectional weights combine all the previous samples and IP10 refreshment sample. And starting at IP11 cross-sectional weights also incorporate IP11 refreshment in addition to all other samples.

In the wave with a refreshment (e.g. IP7 and IP10) the cross-sectional weight is calculated in two parts, the first part being a 'usual' cross-sectional weight (derived from a longitudinal enumeration weight and weight shared to TSMs) for 'continuing samples' (all samples excluding refreshment part) and a cross-sectional weight for refreshment part of the sample (which is calculated in the same way as wave 1 weight, using predictors from linked Census data and other geographical linked data predicting household response, and poststratified by Government Office Region at household level and by gender and 5 age groups at individual level). Both of the resulting weights for 'continuing' samples and the refreshment sample are scaled to 1 before being combined into one variable – the cross-sectional enumerated person weight ($n_psnenip_xw$).

The cross-sectional enumeration person weight ($n_psnenip_xw$) in waves that do not have a new refreshment subsample (e.g. IP5, IP6, IP8, IP9 and so on) is based on the respective longitudinal weight ($n_psneni1_lw$ between IP5 and IP7; $n_psneni2_lw$ between IP8 and IP10, non-refreshment part of $n_psneni3_lw$ for IP11, and $n_psneni4_lw$ for IP12 onwards), with application of the weight share method to assign weights to TSMs. In households

containing one or more enumerated TSMs, each respondent in the household is given a weight of a/b , where a is the sum of the relevant longitudinal weights in the household and b is the total number of respondents in the household who are either an OSM or a TSM. In other households, the cross-sectional weight equals the respective longitudinal weight. Some respondents will therefore receive a zero cross-sectional weight, namely respondents in households containing no OSMs with a non-zero longitudinal weight.

The household weight (**$n_hhdenip_xw$**) equals the minimum of the cross-sectional enumerated person weights (OSMs and TSMs) in the household in the waves without refreshment. In the years with refreshment it consists of two parts of which the first one (for continuing subsamples) is equivalent to the household weight in non-refreshment waves, scaled to 1; and the part for refreshment is calculated in the same way as wave 1 weight, using predictors from linked Census data and other geographical linked data predicting household response, and poststratified by Government Office Region.

The cross-sectional individual response weights for adults (**$n_indpxip_xw$** , **$n_indinip_xw$** and **$n_indscip_xw$**) are calculated conditional on successful enumeration (**$n_psnenip_xw$**), and response to the relevant instrument: proxy or main questionnaire, only main questionnaire and self-completion questionnaire.

The cross-sectional weight for youth questionnaire data (**$n_ythscip_xw$**) is calculated as the wave n enumerated person weight (**$n_psnenip_xw$**) multiplied by an adjustment for nonresponse to the youth questionnaire conditional on enumeration in wave n . Due to the small sample size of youth enumerated sample members, the model uses all respondents age 10 or above and models whether they responded to the self-completion questionnaire excluding predictors that are relevant only for adults (e.g. marriage or employment status). The resulting weight is inferred only to the relevant age group, 10- to 15-year-olds.

Wave 12 biomarker weights

For the users who are interested in biomarker data collected in IP12 we provide separate weights. For interviewer or nurse measured estimates of blood pressure please use **$l_indbpip_xw$** weight, and those of height and weight please use **$l_indbmip_xw$** . If you are interested in combining measured and self-reported blood pressure and/or height and weight measures you could either use a suboptimal weight **$l_indinip_xw$** , or create your own weight tailored to your analysis (use **$l_indinip_xw$** as your base weight). The weights for measured blood pressure **$l_indbpip_xw$** and for height and weight **$l_indbmip_xw$** can be found on **$l_indresp_ip.dta$** data file. If you are interested in comparing self-reported measures to those obtained by an interviewer and/or nurse see our online training material [Creating tailored weights for UKHLS](#).

For analysing full blood samples please use **$l_indfbip_xw$** , and for analyzing dry blood spots samples please use **$l_indbsip_xw$** available on **$l_indresp_ip.dta$** . If you are interested in comparing measures obtained from full blood sample to those obtained from dry blood spots sample it is best to restrict your analysis to those who have both measures and use **$l_indfbip_xw$** weight. For analysing hair samples, please use **$l_indhspip_xw$** available on **$l_hair_ip.dta$** .

The four weights **$l_indbpip_xw$** , **$l_indbmip_xw$** , **$l_indfbip_xw$** and **$l_indbsip_xw$** were calculated using stepwise logistic regression predicting response to blood pressure/ height

and weight / full blood samples / dry blood spots samples respectively conditional on positive **I_indinip_xw**. The predictors used were obtained from IP12 household and personal questionnaires. The reciprocal of the predicted probabilities was multiplied by **I_indinip_xw** to obtain the final weight. The obtained values were weight shared to all eligible members of a household, and additionally weight shared to those with remaining zero weights. Those who were not eligible for each separate biomeasure were excluded from each model.

Because hair samples were collected from adults and youth, the **I_indhsip_xw** weight is modelled in the same way as above but using **I_psnenip_xw** as a base weight and predictors from the household questionnaire.

6. Experiments, methodological studies and non-experimental new content carried in the Innovation Panel

Submit a proposal for an experiment or other new survey content

Each wave of the Innovation Panel has experiments related to measurement or fieldwork procedures, or to address substantive social science questions. Researchers can [apply in an annual competition](#) to have their own experimental or methodological studies on the IP. Since Wave 16, the annual Innovation Panel competition has also included a specific call for proposals for studies comprising new questions without significant experimental or survey methods components.

A series of working papers is published documenting findings from experiments and other studies. For a list of all working papers, see References – summary results from IP experiments and new data.

Researchers may have research questions related to the experimental manipulation or simply need to be aware of potential effects the experiment may have on their own research questions. This information may also be useful in formulating research ideas. The descriptions identify variables allocating cases to different conditions and other variables relevant to carrying out the experiment. Allocation variables are copied into the next waves as well.

In designing the experiments and allocating sample members to randomised treatments great care is taken to avoid contamination and confounding of experiments (Lynn and Jäckle 2019). Depending on the purpose of the experiment, treatments are randomised between individuals, between households, between interviewers, or between primary sampling units. The level at which randomisations are allocated is recorded in the documentation below. For experiments that might affect the same outcomes, the randomisations are fully crossed. In addition, all randomisations are stratified.

6.1. Overview of Innovation Panel experiments and studies

The following table summarises experiments and studies carried in the Innovation Panel for all waves of the study. Each is described in greater detail in its own section of this User Guide. The table includes a brief title and description, a summary of the design, and the waves in which it was carried. Experiments and studies are categorised as:

- Procedural experiments
- Questionnaire design experiments - general issues
- Questionnaire design experiments - specific topics
- Non-experimental studies

Procedural experiments

Study	Summary	Wave
7.1 Respondent incentives to encourage participation and web response	Different values of unconditional respondent incentives, including • £5, £10 or £5 increasing to £10 if all household members participate (original sample); • £10, £20, £30 (IP4, IP7 refreshment samples); • £10 +£5 bonus if all household members complete by web (web mixed modes group) • £10 +£20 bonus if all household members complete by web (web mixed modes group)	All
7.2 Conditional and unconditional incentives	½ households issued to standard unconditional incentives, ½ issued to treatment: previous wave respondents received unconditional incentive, previous wave non-respondents received conditional incentive. Amount of incentive determined by allocation to experiment 1.	7
7.3 Mixed mode experiments: telephone and face-to-face	1/3 households issued to face-to-face, 2/3 issued to telephone with respondents followed up face-to-face. Two different strategies for issuing telephone sample to face-to-face: as soon as one person requires face-to-face visit, or only once all household members attempted by telephone.	3

Study	Summary	Wave
7.4 Mixed mode experiments: web and face-to-face	1/3 households issued to face-to-face interviewing, 2/3 invited to complete survey online. Non-respondents followed up by face-to-face interviewers. (Wave 12 had a third mode allocation: nurse-led.)	5-12, 15-17
7.5 Paper vs CASI self-completion	½ respondents completed adult self-completion on paper, ½ on CASI	4, 5, 6
7.6 Advance materials: letters vs. cards	½ respondents received advance letter, ½ in greeting card format; same content	2
7.7 Advance materials: content of advance letters	4 versions of advance letter text, varying the content to test theories about how people can be persuaded to take part in surveys. Plus self-completion questions about self-rated helpfulness, conformity, and preference for consistency	5
7.8 Targeted advance letters	½ sample received standard advance letter; ½ received targeted version – targeted groups: • 16-29 year-olds • employment-busy • have dependent children under age 15 • living in London • of pensionable age	6
7.9 Number of mailings between interviews	• ½ households received 1 mailing between interviews • ½ received 3 mailings	7
7.10 Different ways of asking respondents to register on participant website	2×2 crossed design: • Invitation to register with participant website sent in letter versus email • £5 incentive for registering versus no incentive	4

Study	Summary	Wave
7.11 Improving interviewer observations about characteristics of address	Wording of questions about likelihood that household has a car or children was improved for ½ of interviewers; standard version for other 1/2	4
7.12 Effect of content of re-issue letter on refusal conversion	½ reissued households received standard letter, for ½ letter included additional information leaflet	4
7.13 Early bird scheduling: encouraging respondents to call interviewer to schedule appointment	2/3 households were encouraged to call their interviewer in advance of fieldwork to schedule an appointment; ½ of these were offered £5 incentive if they did.	4
7.14 Targeted weekday invitation emails	½ households sent email invitation on a day determined by data from previous waves about weekdays on they completed the web survey, ½ households sent email invitation on a Monday (control)	9
7.15 Using prospect theory in wording of advance letters	Wording of advance letters: ½ households emphasis on positive outcome of participation, ½ emphasis on negative outcome of not participating	10
7.16 Invitation letters for mixed mode survey	Text of advance letters for IP11 refreshment sample allocated to Web-first, split into four groups: • Mention interviews with all HH members, mention interviewer follow up • Mention interviews with all HH members, but not interviewer follow up • Mention interviewer follow up, but not interviews with all HH members • Mention neither	11
7.17 Spending Study 2	• ½ households invited to Spending Study 2 in IP11 interview • ½ invited later by postal letter	11
7.18 Invitation to complete pre-interview blood pressure measure:	• 1/3 households given information on their nearest pharmacy to enable blood pressure measurement • 1/3 received an altruistic/pro-social appeal to encourage participants to get their blood pressure measured.	12

Study	Summary	Wave
information treatment vs pro-social appeal	• 1/3 none of the above (control)	
7.19 Fieldwork compression experiment	• 1/5 households continuous longer interview, full set of rotating modules • 1/5 potential break-off request, full set of rotating modules • 1/5 continuous longer interview, reduced set of rotating modules • 1/5 potential break-off request, reduced set of rotating modules • 1/5 control group, standard IP13 questionnaire and incentives	13
7.20 Event triggered data collection	• 70% households: invited to event-triggered data collection • 30% households: not invited	13
7.21 Consent to send survey questions by SMS	IP13: • ½ households: consent asked in Demographics module (early in questionnaire) • ½ households: consent asked in Contact Details module (at end of questionnaire) IP15: question re-asked non-experimentally of those who did not consent in IP13	13,15
7.22 Wellbeing app study	• ½ households each: invited to app early vs late in IP13 interview • ½ households each: 2 vs 10 min daily app survey • 1/3 households each: no bonus vs £10 for completing all 14 days vs £2.50 on 4 randomly selected days	13
7.23 Asking for Living Apart Together partner details	• ½ households: asked partner details in survey • ½ households: asked partner details in interwave mailing	14
7.24 Contact protocols for IP14 refreshment sample	Multiple experiments: prenotification letters and number of reminders, logos/branding on envelopes, gift incentive, explaining longitudinal nature of the survey, early bird incentive, explaining nature of household survey, collecting email addresses of all adults in the household grid, encouraging messages during the survey	14

Study	Summary	Wave
7.25 PERKs for unconditional incentives	Households allocated incentive groups: • ½ households sent unconditional Love2Shop vouchers with the advance letter • ½ households told to download their electronic voucher online Crossed with four motivational messages: • No additional message • E-incentives are more eco-friendly • E-incentives are easier to use Both messages.	16
7.26 Youth online survey	Households with children aged 10-1: • ½ households: Cover letter mentioned that if the child completed the survey online or returned the questionnaire, they would be sent an additional £5 gift card. • ½ households: no additional incentive Information leaflet: • ½ households: information leaflet was targeted to the young person ½ households: leaflet targeted to the parent	16
7.27 Test of Government logo on envelopes for continuing sample	• ½ households: envelopes with government logo ½ households: no government logo	16
7.28 Youth survey invite mailing	Variations on process for sending the invitation to the youth survey (within the adult letter or in an envelope addressed to the young person).	17

Questionnaire design experiments: General issues

Study	Summary	Wave
8.1 Subsetting the questionnaire content	• ½ respondents asked questions about environmental behaviours/height and weight every year; long partnership and fertility history in wave 1 • ½ asked every second year; short history in wave 1	1-5
8.2 Showcards vs. no showcards	½ respondents answered questions with long lists of response options using showcards, ½ without showcards	1-3
8.3 Impact of question wording and context on measuring change	Four related experiments: • ½ respondents had standard question wording, ½ had wording where definitions were less ambiguous • ½ respondents were given explicit instructions for a “select one” type question where response categories were not mutually exclusive • ½ respondents were asked about the dates of events implicitly (“when...”), ½ were asked explicitly (“in which month and year...”); all asked about strategies used to recall dates Effects on high/low frequency context on questions with vague quantifiers: ½ respondents had a high frequency context, ½ a low frequency context for a question about the frequency of behaviours using vague quantifiers	2, 3, 4
8.4 Dependent interviewing wording	• Waves 3, 4: ½ respondents asked yes/no question whether response from previous waves “still the case?”, ½ asked “has this changed?” Waves 5, 7: two additional forced choice versions asking “is this still the case or has it changed?” or “has this changed or is it still the case?”	3, 4, 5, 7
8.5 Branched vs. unbranched rating scales for measuring attitudes	½ respondents asked standard agree/disagree questions, ½ first asked whether they agree or disagree and then about the strength of their attitude	3, 4, 5

Study	Summary	Wave
8.6 Smiley faces vs. text-based scales in child self-completion	½ of youth rated satisfaction with different domains using smiley faces, ½ using text based scales	5, 6
8.7 Quality of recall data with web vs. face-to-face	Respondents asked to recall facts gathered contemporaneously at earlier waves. Half of respondents completing the survey on the web received a commitment pledge as an experimental treatment to encourage more accurate reporting of historical information.	6
8.8 Methods of reducing item non-response in web surveys	3 treatments: • Standard procedure: dk/refused not offered initially; if respondent presses “next” without answering, the options appear • As above but with additional prompt asking respondent to complete the question • Follow-up questions at end of questionnaire for items not answered	6
8.9 Separating systematic measurement error components using MTMM in longitudinal studies	Respondents asked 6 questions about immigration at start and again at end of questionnaire, varying: • Whether question emphasizes positive or negative aspects • Number of scale points	7, 8, 9
8.10 Replicating classic response order experiments across countries	• Series of cross-sectional question wording experiments	7
8.11 Impact of response scale direction on responses	For ½ respondents the direction of the answer scale was reversed	7, 8
8.12 Enhancing respondent engagement with the survey through tailored interesting questions	Two treatments: • 3 additional questions, tailored to respondents’ interests as reported in IP2, or 3 additional questions that are not tailored, e.g. questions on TV watching and favourite programs. No additional questions	7

Study	Summary	Wave
8.13 Grid design in mobile surveys	½ of respondents assigned to a standard static grid, ½ assigned to a dynamic grid	10
8.14 Don't know/prefer not to answer response presentation	Respondents randomly allocated to different ways of presenting DK/REF answer options: • Standard procedure • Instructions on DK procedure given first DK/REF response options offered in the initial presentation of question	11
8.15 Collecting mobile phone numbers	IP13 treatments: • ½ households asked about mobile phone numbers before other contact details • ½ asked standard UKHLS contact details module IP15 treatments: • ½ households asked about mobile phone numbers before other contact details ½ asked question worded to increase compliance with providing mobile phone number	12, 15

Questionnaire design experiments: Specific topics

Study	Summary	Wave
9.1 Measures of consumption and expenditure	IP1 three treatments: • Question about overall expenditure without cues • Question about overall expenditure with detailed categorical cues • Separate questions about amounts of expenditure broken out into reporting on each category rather than an overall figure. IP6 two treatments: • Asked for total benefit unit expenditure by adding up a set of expenditure categories using a showcard to trigger recall of expenditure on each category. • Asked for an amount of expenditure for each expenditure category, then reconciled the total amount spent for accuracy.	1, 6
9.2 Measuring satisfaction	Waves 2, 3, 6, several aspects of question wording varied: • Number of scale points • Showcards versus no showcards • Labelling of end points only or all scale points • Position early or late in questionnaire • CASI versus interviewer administered	1-3, 5, 6
9.3 Reference groups in measuring satisfaction	Four treatments: • Standard questions • Satisfaction rating relative to others of same gender • Others of same education • Others of same education and gender	

Study	Summary	Wave
9.4 Measuring identity	For questions about identity ½ respondents were asked about their “profession”, ½ about their “occupation”	2
9.5 Measures of wealth	Four ways of collecting information about money held in savings and investments, 2×2 crossed design varying: • Aggregate amounts versus itemized • Reports from all adults versus one financial reporter	3
9.6 Context of questions about consent to data linkage with administrative records	2×2 crossed design varying whether • consent asked in context or at end of interview • independent question or dependent reminding respondent of whether or not they gave consent previously	4
9.7 Respondent preferences about mode of data collection	½ respondents asked to rate specific modes then generalized preference; ½ asked first about generalized preference, then ratings of specific modes	4, 5, 6
9.8 Feasibility of directly measuring household energy use	2×2 crossed design varying: • which meter readings household was asked for (odometer only / gas, electricity and odometer) • whether advance letter mentioned that respondent would be asked for meter readings	5, 6
9.9 Context effects in fertility decisions	• ½ respondents asked about expected fertility before questions about friendship networks, ½ asked afterwards	4, 5
9.10 Vignettes: measuring partner satisfaction with division of household labour	Vignettes describing hypothetical scenarios of partners sharing domestic and paid work; 3 vignettes per respondent; dimensions varied: (1) paid work; (2) earnings; (3) presence of children; (4) housework allocations; and (5) use of paid help.	5, 6

Study	Summary	Wave
9.11 Subjective expectations about the returns to higher education and decisions to attend university	½ respondents and their parents shown information about average earnings by gender and subject, ½ no information	5, 8, 9
9.12 Measuring change in self-assessed disability	• ¼ asked standard questions: filter question whether long term illness or disability, if “yes” asked about difficulty with everyday activities • ½ respondents asked follow-up questions if they reported a long term illness or disability, and had not reported this at previous interview, or vice versa about reasons for change. All asked about areas of everyday life where they have difficulty due to health • ¼ only asked about difficulties with everyday life activities, not asked about long-term health problems	6, 7
9.13 Associated Study: measuring time and risk preferences	A total of 91 lottery questions about respondents risk and time preferences, including actual payouts	6, 7
9.14 Assessing how people think about environmental taxes	10 treatments varying questions about respondents’ willingness to pay increased environmental taxes. For each of 5 question pairs, one variant makes no mention of tax reductions elsewhere, while the other variant does	7
9.15 Validity of interviewer ratings of respondent health	½ interviewers asked to assess respondents health at start of interview, ½ at end	8
9.16 Social desirability bias in attitudes towards immigration	Item count list	8, 9
9.17 Measuring sexual identity using direct and indirect questioning	Item count list and direct question	8, 9

Study	Summary	Wave
9.18 What do the general population regard as “successful ageing”?	Vignettes describing different scenarios about the circumstances of a 75 year old; 3 vignettes per respondent; dimensions varied: (1) gender, (2) chronic disease, (3) disability, (4) physical functioning, (5) cognitive functioning, (6) interpersonal engagement, and (7) productive engagement.	9, 17
9.19 Benefits unit finances module	Experiment 1: ½ respondents shown a summary of all income sources reported and asked to check and correct, ½ respondents no summary screen. Experiment 2: budget reconciliation module with ½ benefit units asked about money taken from savings/new credit and money put into savings/repayment of credit (gross flows), vs. changes in savings and credit accounts (net flows).	9
9.20 Presentation of response options in satisfaction questions	1/3 households assigned to each treatment: • Grid format • One question per screen, response options vertically aligned • One question per screen, response options horizontally aligned	9
9.21 Improving consent to link to the electoral register	• Random allocation of households to opt-in vs. opt- out and two versions of wording the consent question	10, 11
9.22 Financial management within couples	• Households randomly allocated to two different versions of questions asking about financial management and perception of money ownership within couples	10
9.23 Non-resident parents and reasons for separation	• Households randomly assigned to two different versions asking about non-resident parents and reasons for separation	10
9.24 Variations of the EQ-5D questions	Respondents were randomly allocated to three equal sized groups: • ask EQ-5D-3L Late and ask EQ-5D-5L Early • ask EQ-5D-3L Early and ask EQ-5D-5L Late • ask EQ-5D-5L Late ONLY	11
9.25 HMRC data linkage consent experiment	IP11: • Respondents randomly allocated to an easy vs. difficult version of the consent question • For CAPI respondents question difficulty was crossed with early vs. late location in the questionnaire	11, 15

Study	Summary	Wave
	IP15: • FTF respondents randomly allocated to CAPI, CASI, partial CASI consent question • Web respondents randomly allocated to control or consent question with message from and photo of study director	
9.26 Does competition over public services decrease support for residency rights of immigrants?	• Conjoint experiment and vignettes	11
9.27 Reporting of height and weight	• Respondents interviewed face-to-face: • ½ households were asked to report height and weight by the interviewer • ½ households asked about height and weight in the self-completion module	12
9.28 Biomarker and sample collection	• Depending on the mode of interview (see Section 16.4), respondents were asked to provide blood samples, dried blood spots, and hair samples.	12
9.29 LinkedIn consent	Two randomised allocations: • ½ respondents asked early, ½ late in the questionnaire • ½ respondents shown a motivational statement before the consent question, ½ no motivational statement	14
9.30 Vignettes: intentions to prepare for automation	Vignettes describing different scenarios about the prospected severity of technological developments (severe threat vs. minor threat) and the prospected timeframe of such developments (short-term vs. long-term) followed by questions about attitudes and behaviours regarding re- and upskilling	14
9.31 Comparison of ReQoL-10 to other scales	Randomised order: • ½ respondents: ReQoL and EQ5D early and GHQ late • ½ respondents: ReQoL and EQ5D late and GHQ early	14
9.32 Proxy nomination	Two randomised allocations: • ½ respondents asked early, ½ late in questionnaire • ½ respondents asked wording version 1, ½ asked version 2	14

Study	Summary	Wave
9.33 Twitter consent	Two randomised allocations • ½ respondents shown 'help' links with additional information on the same page • ½ respondents shown help links on a separate page	15
9.34 Marginal propensity to consume	Question wording with 3 randomisations, crossed: • Two-part vs. direct question on marginal propensity to consume • 3 month vs. 12 month period • £500 vs. £2500 windfall gain	15
9.35 Informal care	Two randomised allocations: • Control: current UKHLS questions • Amended caring questions • Combining comparable questions	15
9.36 Alcohol consumption	Two randomised allocations: • Questions used by NHS • Questions used by UKHLS	15
9.37 Body Volume Index app and body measurements	Waist and hip measurements collected as self-measurement, interviewer observed self-measurement, and from an app that calculates measurements based on photos of the respondent Randomised allocations for the invitation to the app study: • Respondents promised feedback on their total body fat, visceral body fat, or no feedback • £5 additional unconditional incentive for completing the survey vs. £5 conditional on using the app	15
9.38 National identity	Randomised allocations: • Importance of being British first • National identity first	15

Study	Summary	Wave
9.39 Vignettes: measuring flexibility stigma - double whammy or femininity stigma	Vignettes describing different characteristics of job applicants, followed by questions about whether respondent would recommend the applicant and their perceptions of the applicant	16
9.40 Cognitive reflection and politically motivated reasoning	• 1/3 households allocated to control version • 1/3 households allocated to Brexit Remain version • 1/3 households allocated to Brexit Leave version	16
9.41 Robustness of climate change worries measurement	• ½ households allocated to 'Worry me' version • ½ households allocated to 'Worry about' version	16
9.42 Understanding of the long-term future	Households allocated to • Financial questions only • Environmental questions only • Both financial and environmental questions And different question order/position: • Environmental questions early/financial questions late • Financial questions early/environmental questions late	16
9.43 Mental health questions comparisons	Households allocated to different versions of the UKHLS questions about mental health: • Wave 8 version • Wave 10 version • Wave 14 version	16
9.44 Asking for child red book pictures	Households with children <16: • ½ asked for information from red book (personal child health record) before interview • ½ asked during the interview	16

Study	Summary	Wave
9.45 Spatial cognition mobile app game	• ½ households: £10 conditional incentive • ½ households: £30 conditional incentive	16
9.46 Consent decision process	Respondents asked variants of question seeking consent to link to data from the Department for Work and Pensions about welfare benefits. Aiming to see whether participants can be encouraged to think more reflectively about consent request. Face-to-face respondents allocated in equal 1/3s between 3 variants; web-respondents allocated in equal 1/5s between those plus 2 more options. • Standard version • Benefits / risks version • Value for science version • (Web only) Reasons for/against consenting version • (Web only) Objective understanding version	17
9.47 Identification of informal caregiving	Experiment to compare standard Understanding Society questions on informal care against 'activity-based' questions.	17
9.48 Labour market expectations	Experiment to investigate respondents' reporting of labour market expectations, depending on how options are presented to them.	17
9.49 Indoor residential environment: consent for in-home sensor	Variations on consent questions seeking consent to place in-home sensors in respondents' homes. Variation based on: • Whether information was placed in the question text or in supplemental help text. • The maximum duration they would be asked to keep the sensor in the home. • Whether they were offered feedback from the sensor.	17

Study	Summary	Wave
9.50 Domestic energy use: consent for smart meter data linkage	Variations on consent questions seeking consent to collect smart meter data varying the presentation of the information provided (whether in bullet point format; paragraph format; or abbreviated bullet point format with further information in help text).	17

Non-experimental studies

Study	Summary	Wave
10.1 Questions about twins	Non-experimental: whether respondent is a twin, type of twin, whether they would disclose twin's address and likelihood that twin would participate in a follow-up study	5
10.2 Measuring finger length ratios as indicator of prenatal testosterone exposure	- IP6: non-experimental measurement of 2nd and 4th digit length on both hands in adult interview; - IP7 measurement in adult interview repeated for new sample members (including refreshment sample), and included in youth self-completion questionnaire	6, 7
10.3 Associated Study: time-use diary	Each respondent asked to complete two paper time use diaries. Random allocation to a weekday and a weekend day.	7
10.4 Spending Study 1	A mobile app based study where respondents were asked to upload pictures of all their shopping receipts for a month, fielded autumn 2016 (after IP9 interviews)	9
10.5 Consent to link Twitter data	All respondents asked whether they use Twitter and whether willing to link their Twitter account to their survey data	10
10.6 An investigation of children's consistency in reporting their parents' occupations	Youth and young adults asked about mother's and father's occupation	11
10.7 Panel conditioning	Questions about how participating in the survey affected the way respondents think, their behaviours, their attitudes, and why they participate.	13
10.8 Living Apart Together (LATs) partner and survey	Questions about name and contact details.	13

Study	Summary	Wave
10.9 Asking for Parents Living Apart other parent details	Questions about name and contact details.	15
10.10 What constitutes an interesting size effect when measuring people's psychological state	Question comparing current life satisfaction to life satisfaction when previously interviewed.	16
10.11 People's expectations of gender discrimination related to work	Question asking respondent's expectations of facing gender discrimination in the workplace.	16
10.12 The extent and nature of the use of domestic workers	Questions about the use of paid domestic labour (cleaners, gardeners, etc.)	16
10.13 Judging the passage of time	Questions asking respondents' perceptions of how quickly time passes for them.	16
10.14 Adaptation to eco-climate emergency	Questions about emotions related to climate change and support for / opposition to various associated policies.	17
10.15 Youth online survey	New online version of the youth questionnaire.	17
10.16 Indoor residential environment and energy use: background information	Additional non-experimental data collection gathered in support of the experiments described in: • 9.49 Indoor residential environment: consent for in-home sensor • 9.50 Domestic energy use: consent for smart meter data linkage	17

7. Procedural experiments

References to publications using data from IP experiments are included in the summary table above.

7.1 Respondent incentives to encourage participation and web response

A set of randomised assignments related to payment of respondent incentives extends across waves 1-6 of the IP. At Wave 6, experimental allocation was nested within mixed mode treatments (See Mixed mode experiments). Households within primary sampling units (PSUs) were randomly allocated to treatments within PSUs. All enumerated adults within a household received the same incentive offer.

At Wave 1, each household received an initial unconditional £5 incentive sent with the advance letter, which was “topped-up” with the total incentive amount for the household after the interview. It contrasts lower and higher payments, uniform for the household, with a higher incentive if all eligible enumerated adults in the household were interviewed. The experimental treatment groups were as follows:

Group 1 – £5 per interviewed household member

Group 2 – £10 per interviewed household member

Group 3 – £5 per interviewed household member increasing to £10 per person if all eligible enumerated adults in the household were also interviewed

The variable in the data that controls allocation at the household level is **a_groupincentive** on the record **a_hhsamp_ip**.

At Wave 2, the incentive experiment had some groups where the payment amount was the same as at Wave 1 and some where the payment level was reduced. Incentives were sent in advance of fieldwork to named individuals through a postal mailing. Rising 16 year olds newly eligible for an adult interview at Wave 2 received an advance mailing with the appropriate adult incentive. The complete co-operation top-up (in Group 5) was posted to households once the final eligible person was interviewed. Within PSUs, households were randomly allocated to each of the 5 treatment groups. All adults within each household received the same treatment:

Group 1 – Receive £5 (as per IP1)

Group 2 – Receive £10 (as at IP1)

Group 3 – Receive £5 (reduction from £10 at IP1 to £5 at IP2)

Group 4 – Receive £5 rising to £10 if complete household co-operation (as per IP1)

Group 5 – Receive £5 (reduction from possible £10 at IP1 to £5)

The controlling variable is **b_ff_incentw2** on the record **b_hhsamp_ip**.

In Wave 3, five of the six treatment groups retained the same treatment as at Wave 2. One of the two groups receiving £10 at both Wave 1 and Wave 2 had the incentive decreased to £5, to test the effect of a decrease at an early stage in the panel. All incentives were sent in advance of fieldwork to named individuals through a postal mailing. Rising 16-year-olds

newly eligible for an adult interview at Wave 3 received an advance mailing with the appropriate adult incentive. The complete co-operation top-up was posted to households once the final eligible person was interviewed. Households within PSUs were allocated to treatments. There were 6 experimental groups and all adults within each household received the same treatment:

Group 1 – £5 (same at IP1 & IP2)

Group 2 – £10 (same at IP1 & IP2)

Group 3 – £5 (was £10 at IP1 & IP2)

Group 4 – £5 (was £10 at IP1)

Group 5 – £5 to £10 for complete cooperation (same at IP1 & IP2)

Group 6 – £5 (was £5 to £10 at IP1)

The controlling variable is **c_ff_incentw3** on the record **c_hhsamp_ip**.

At Wave 4, some of the continuing sample households received an increase in incentive, from £5 to £10 (See group 2 and group 6 below.) Over the history of their participation, group 2 always received £5 in the past, whereas group 6 had started off with £10. All other groups for continuing households remained as at wave 3. The refreshment sample (groups 9, 10, and 11) received larger amounts than those traditional for the UK. See Section 4 about the refreshment sample.

Group 1 – £5 (same at IP1, IP2, IP3)

Group 2 – £10 (was £5 at IP1, IP2, IP3)

Group 3 – £10 (same at IP1, IP2, IP3)

Group 4 – £5 (was £10 at IP1, IP2)

Group 5 – £5 (was £10 at IP1)

Group 6 – £10 (was £10 at IP1 and £5 at IP2 & IP3)

Group 7 – £5 to £10 for complete cooperation (same at IP1, IP2, IP3)

Group 8 – £5 (was £5 to £10 at IP1)

Group 9 – £10 per interviewed household member

Group 10 – £20 per interviewed household member

Group 11 – £30 per interviewed household member

At Wave 4, the controlling variable is **d_ff_incentw4** on the record **d_hhsamp_ip**.

At Wave 5, the Wave 4 conditions were repeated with one exception. Respondents in the experimental treatment receiving £5 unconditionally, rising to £10 if all household members completed the interview, were randomly allocation to receive either a £5 or a £10 unconditional incentive. The result is that at Wave 5, 7/12 of the sample received a £5 incentive and 5/12 received a £10 incentive. All refreshment sample experimental allocations from Wave 4 were repeated at Wave 5.

Group 1 – £5 (same at IP1, IP2, IP3, IP4)

Group 2 – £10 (was £5 at IP1, IP2, IP3, IP4)

Group 3 – £10 (same at IP1, IP2, IP3, IP4)

Group 4 – £5 (was £10 at IP1, IP2)

Group 5 – £5 (was £10 at IP1)

Group 6 – £10 (was £10 at IP1 and £5 at IP2, IP3, IP4)

Group 7 – £5 (was £5 rising to £10 for complete cooperation at IP1, IP2, IP3, IP4)

Group 8 – £10 (was £5 rising to £10 for complete cooperation at IP1, IP2, IP3, IP4)

Group 9 – £5 (was £5 to £10 at IP1)
Group 10 – £10 per interviewed household member
Group 11 – £20 per interviewed household member
Group 12 – £30 per interviewed household member

At Wave 5, the controlling variable is **e_ff_incentw5** on the record **e_hhsamp_ip**.

In addition, Wave 5 had a mixed mode survey interviewing experiment that examined whether incentives can be used to maximize take-up of the web survey. The design had two factors. The first factor was whether the day on which the email invitation to the web survey is sent. The second factor is whether conditional incentives – offered in addition to the existing unconditional incentives – can increase the take-up of the web survey.

Respondents were allocated to either receive the first email invitation for web interviewing on a Friday or a Monday. This was crossed with an incentive experiment where households were randomly allocated to receive an additional conditional web bonus or no bonus. In the web bonus group, the household reference person was offered a conditional bonus if they complete their part of the survey (household grid, household questionnaire and their individual questionnaire) within 3 days. Once the household grid was completed, all other household members were offered a conditional bonus for completing the web survey. In the control group all household members received the unconditional incentive only.

Group 1 -- Monday, £5 bonus
Group 2 -- Monday, No bonus
Group 3 -- Friday, £5 bonus
Group 4 -- Friday, No bonus

The controlling variable is **e_ff_invitew5** on record **e_hhsamp_ip**.

Certain individuals not interviewed on the first visit to the household if face-to-face were approached to do their individual interview either online or face-to-face. If online, then the offer of a bonus to go online was controlled with experimental treatments 1 & 3 versus 2 & 4 on this item.

At Wave 6, experimentation with incentives was nested within mixed mode interviewing treatment (See 7.4 Mixed mode experiments: web and face-to-face). As with previous waves, allocation to experimental treatment groups was at the household level: all individuals within the household received the same experimental treatment and any split-off households retained experimental allocation from the previous household at both the current and all previous waves.

The one-third of households allocated to a face-to-face interviewing approach at Wave 6 were all allocated to a £10 unconditional incentive. For the five of the nine Wave 5 incentive groups in the original sample, this represented an increase from £5. For two of the Wave 5 refreshment sample incentive groups this was a decrease from £20 or £30.

The two-thirds of households allocated to a web interviewing approach at Wave 6 were allocated in equal proportions to three incentive experimental treatments. These were

crossed with the Wave 5 incentive treatments. All treatments received an unconditional incentive with differences across them in the amount (£10 or £30) and whether an additional conditional incentive was offered for whole household completion within a two-week web-only interviewing period (See MIXED MODE EXPERIMENTS). The groups are characterised as follows:

Group 1 - £10 unconditional incentive

Group 2 - £10 unconditional with a £20 conditional incentive for full-household completion by web in the allotted time

Group 3 - £30 unconditional incentive

The controlling variable is **f_ff_incentw6** on record **f_hhsamp_ip**.

There were four households in the £10 treatment group who became aware of the £30 treatment group and were sent an extra £20. The variable **f_incentcomp** on the record **f_hhsamp** identifies these cases so that they can be excluded from analyses.

Tables 7.1 and 7.2 document the allocation of cases to experimental groups within mixed mode treatments. In order to reflect all historical incentive treatments, the controlling variable takes 48 unique values. However, only 3 different incentive amounts were used at Wave 6:

Groups 1-12, 13, 16, 19, 22, 25, 28, 31, 34, 37, 40, 43, 46 = £10 unconditional incentive

Groups 14, 17, 20, 23, 26, 29, 32, 35, 38, 41, 44, 47 = £10 unconditional + £20 conditioned on whole household web completion

Groups 15, 18, 21, 24, 27, 30, 33, 36, 39, 42, 45, 48 = £30 unconditional incentive

At Wave 7, the incentive experiment paralleled the experimental allocation carried at IP6 with a repeat of the elevated incentive amounts for the IP7 refreshment sample comparable to the IP4 incentive experiment the same purpose. Incentive allocation was nested within the mixed mode allocation repeated through from IP5. **Households within IP1 (or IP4) PSUs were allocated the treatments such that all individuals within households received the same incentive, and all split-off households received the same incentive as their originating household.** The continuing IP sample was allocated to the identical treatments as at IP6:

Face to Face sample (i.e., **ff_gridmodew5** = 1) = £10

Web allocated (i.e., **ff_gridmodew5** = 3) to three groups:

Group 1 - £10 unconditional incentive

Group 2 - £10 unconditional with a £20 conditional incentive for full-household completion by web in the allotted time

Group 3 - £30 unconditional incentive

The IP7 refreshment sample was allocated to the face-to-face experimental treatment (i.e., `ff_gridmodew5 = 3`), then divided evenly between three incentive experimental groups where each individual in the household received the same incentive amount:

Group 1 - £10 unconditional incentive

Group 2 - £20 unconditional incentive

Group 3 - £30 unconditional incentive

The controlling variable is `ff_incentw7` on record `hhsamp`. The first 48 unique values were carried over from `ff_incentw6` indicating a parallel to the IP6 approach. Codes 49, 50 and 51 were used to indicate the incentive amount for the IP7 refreshment sample. Tables 7.1 and 7.2 document the allocation of cases to experimental groups, though a shorthand allocation for the continuing sample is as follows:

Groups 1-12, 13, 16, 19, 22, 25, 28, 31, 34, 37, 40, 43, 46 = £10 unconditional incentive

Groups 14, 17, 20, 23, 26, 29, 32, 35, 38, 41, 44, 47 = £10 unconditional + £20 conditioned on whole HH completion

Groups 15, 18, 21, 24, 27, 30, 33, 36, 39, 42, 45, 48 = £30 unconditional incentive

Group 49 = IP7 refreshment sample £10 unconditional

Group 50 = IP7 refreshment sample £20 unconditional

Group 51 = IP7 refreshment sample £30 unconditional

At Wave 8, the incentive experiment paralleled the experimental allocation carried at previous waves with a repeat of the elevated incentive amounts for the IP7 refreshment sample comparable to the IP4 incentive experiment. Incentive allocation was nested within the mixed mode allocation at wave 8, with reference to the sample status vis., whether IP1/IP4 or IP7 refreshment such that the IP7 refreshment will maintain the incentive treatment it received at IP7. Households within IP1 (or IP4) PSUs were allocated the treatments such that all individuals within households received the same incentive, and all split-off households received the same incentive as their originating household.

IP1/IP4 continuing Face to Face sample (`ff_gridmodew5 = 1` & `ff_gridmodew8 = 1` & `ff_hhorig = 7, 10`) = £10 unconditional incentive.

IP1/IP4 former mixed-mode sample shifting to F2F (`ff_gridmodew5 = 3` & `ff_gridmodew8 = 1` & `ff_hhorig = 7, 10`) = £10 unconditional incentive.

IP7 refreshment sample issued Face-to-Face (ff_gridmodew8 = 1 and ff_hhorig = 11):

Group 1 - £10 unconditional incentive

Group 2 - £20 unconditional incentive

Group 3 - £30 unconditional incentive

Web allocated (i.e., ff_gridmodew8 = 3):

Group 1 - £10 unconditional incentive

Group 2 - £10 unconditional with a £20 conditional incentive for full-household completion by web in the allotted time

Group 3 - £30 unconditional incentive

The controlling variable is **ff_incentw8** on record **h_hhsamp_ip**. It takes 87 unique values to reflect the historical incentive treatments for each case as determined by past incentive experiments and the current Wave 8 experimental allocation to incentive treatment. Tables 7.1-7.3 detail the meaning of each of these 87 codes which can be summarised as follows:

Groups 1-49, 52, 55, 58, 61, 64, 67, 70, 73, 76, 79, 82, 85 = £10 unconditional incentive

Groups 51, 54, 57, 60, 63, 66, 69, 72, 75, 78, 81, 84, 87 = £30 unconditional incentive

Group 50 = £20 unconditional incentive

Groups 53, 56, 59, 62, 65, 68, 71, 74, 77, 80, 83, 86 = £10 rising to £20 with full-household completion

At Wave 9, the incentive experiment has identical allocations to the incentive experiment at IP8. Households within IP1 (or IP4) PSUs were allocated to treatments such that all individuals within households received the same incentive, and all split-off households received the same incentive as their originating household.

IP1/IP4 continuing Face to Face sample (ff_gridmodew5 = 1 & ff_gridmodew8 = 1 & ff_hhorig = 7, 10) = £10 unconditional incentive.

IP1/IP4 former mixed-mode sample shifting to F2F (ff_gridmodew5 = 3 & ff_gridmodew8 = 1 & ff_hhorig = 7, 10) = £10 unconditional incentive.

IP7 refreshment sample issued Face-to-Face (ff_gridmodew8 = 1 and ff_hhorig = 11):

Group 1 - £10 unconditional incentive

Group 2 - £20 unconditional incentive

Group 3 - £30 unconditional incentive

Web allocated (i.e., ff_gridmodew8 = 3):

Group 1 - £10 unconditional incentive

Group 2 - £10 unconditional with a £20 conditional incentive for full-household completion by web in the allotted time

Group 3 - £30 unconditional incentive

The controlling variable is **ff_incentw9**. It takes 87 unique values to reflect the historical incentive treatments for each case as determined by past incentive experiments and the current IP9 experimental allocation to incentive treatment. Tables 1-3 (including amount of incentive for IP8 and IP9 only for space purposes) detail the meaning of each of these 87 codes which can be summarised as follows:

Groups 1-49, 52, 55, 58, 61, 64, 67, 70, 73, 76, 79, 82, 85 = £10 unconditional incentive

Groups 51, 54, 57, 60, 63, 66, 69, 72, 75, 78, 81, 84, 87 = £30 unconditional incentive

Group 50 = £20 unconditional incentive

Groups 53, 56, 59, 62, 65, 68, 71, 74, 77, 80, 83, 86 = £10 rising to £20 with full-household completion

At Wave 10, the incentive experiment had identical allocations as in wave 9. Households within IP1 (or IP4) PSUs were allocated to treatments such that all individuals within households received the same incentive, and all split-off households received the same incentive as their originating household.

IP1/IP4 continuing Face to Face sample (ff_gridmodew5 = 1 & ff_gridmodew8 = 1 & ff_hhorig = 11, 10) = £10 unconditional incentive.

IP1/IP4 former mixed-mode sample shifting to F2F (ff_gridmodew5 = 3 & ff_gridmodew8 = 1 & ff_hhorig = 11,12) = £10 unconditional incentive.

IP7 refreshment sample issued Face-to-Face (ff_gridmodew8 = 1 and ff_hhorig = 13):

Group 1 - £10 unconditional incentive

Group 2 - £20 unconditional incentive

Group 3 - £30 unconditional incentive

IP10 refreshment sample issued Face-to-Face (ff_gridmodew10 = 1 and ff_hhorig = 14):

Group 1 - £10 unconditional incentive

Web allocated (i.e., ff_gridmodew8 = 3):

Group 1 - £10 unconditional incentive

Group 2 - £10 unconditional with a £20 conditional incentive for full-household completion by web in the allotted time

Group 3 - £30 unconditional incentive

The controlling variable is **ff_incentw10** (in record HHSAMP). It takes 88 unique values to reflect the historical incentive treatments for each case as determined by past incentive experiments and the current experimental allocation to incentive treatment. Tables 1-3 (including amount of incentive for IP8 and IP9 only for space purposes) detail the meaning of each of these codes which can be summarised as follows:

Groups 1-49, 52, 55, 58, 61, 64, 67, 70, 73, 76, 79, 82, 85, 88 = £10 unconditional incentive

Groups 51, 54, 57, 60, 63, 66, 69, 72, 75, 78, 81, 84, 87 = £30 unconditional incentive

Group 50 = £20 unconditional incentive

Groups 53, 56, 59, 62, 65, 68, 71, 74, 77, 80, 83, 86 = £10 rising to £20 with full-household completion

Additionally, all previous wave non-responding households which were issued to field received an incentive conditional on their survey participation at the value of their previously assigned group value.

At Wave 11, the incentive experiment was identical to the incentive experiment at IP10 for the IP1, IP4 and IP7 samples, with the exception that now the £20 conditional incentives were conditional on the individual, not household completing the survey. In previous waves, respondents received the £20 extra conditional on entire household completion. IP10

households, allocated to the mixed-mode design, follow the same allocations as IP1/IP4/IP7 mixed-mode households. The IP11 refreshment sample was also allocated to the mixed-mode experiment. All IP11 sample members received £10 unconditionally, with individuals in the mixed-mode receiving an extra £15 conditional on their completion of the web survey within the first three weeks. Households within PSUs were allocated to treatments such that all individuals within households will received the same incentive, and all split-off households received the same incentive as their originating household.

IP1/IP4 continuing Face to Face sample (ff_gridmodew5 = 1 & ff_gridmodew11 = 1 & ff_hhorig = 11, 12) = £10 unconditional incentive.

IP1/IP4 former mixed-mode sample shifting to F2F (ff_gridmodew5 = 3 & ff_gridmodew11 = 1 & ff_hhorig = 11,12) = £10 unconditional incentive.

IP7 refreshment sample issued Face-to-Face (ff_gridmodew11 = 1 and ff_hhorig = 13):

Group 1 - £10 unconditional incentive

Group 2 - £20 unconditional incentive

Group 3 - £30 unconditional incentive

IP10 refreshment sample issued Face-to-Face (ff_gridmodew11 = 1 and ff_hhorig = 14):

Group 1 - £10 unconditional incentive

Web allocated for IP1/IP4/IP7/IP10 samples (i.e., ff_gridmodew11 = 3 and ff_hhorig = 11, 12, 13, 14):

Group 1 - £10 unconditional incentive

Group 2 - £10 unconditional with a £20 conditional incentive for **individual** completion by web in the allotted time

Group 3 - £30 unconditional incentive

IP11 refreshment sample issued Face-to-Face (ff_gridmodew11 = 1 and ff_hhorig = 18):

Group 1 - £10 unconditional incentive

Web allocated for IP11 (i.e., ff_gridmodew11 = 3 and ff_hhorig = 18):

Group 1 - £10 unconditional incentive with a £15 conditional incentive for each individual completion on web within three weeks

The controlling variable is **ff_incentw11 on record K_HHSAMP_IP**. It takes 93 unique values to reflect the historical incentive treatments for each case as determined by past incentive experiments and the current IP11 experimental allocation to incentive treatment. Tables 7.1-7.3 (including amount of incentive for IP8-IP11 only for space purposes) detail the meaning of each of these 93 codes which can be summarised as follows:

Groups 1-49, 52, 55, 58, 61, 64, 67, 70, 73, 76, 79, 82, 85, 88, 89, 92 = £10 unconditional incentive

Groups 51, 54, 57, 60, 63, 66, 69, 72, 75, 78, 81, 84, 87, 91 = £30 unconditional incentive

Group 50 = £20 unconditional incentive

Groups 53, 56, 59, 62, 65, 68, 71, 74, 77, 80, 83, 86, 90 = £10 rising to £20 with **individual** completion

Group 93 = £10 rising to £15 with **individual** completion

Additionally, all previous wave non-responding households which are issued to field received an incentive conditional on their survey participation at the value of their previously assigned group value.

At Wave 12, the initial unconditional incentives sent to households were identical to those sent at IP11, regardless of mode allocation at IP12. This allocation was simplified at IP12 by using a four condition controlling variable which captures almost every of the above 93 conditions: £10, £20, £30, and £10+£20. Allocation to these groups mirror the allocation to these values in ff_incentw11. The final allocation variable is **ff_incentw12** on record **I_hhsamp_ip**:

Group 1 - £10

Group 2 - £20

Group 3 - £30

Group 4 - £10 + £20

In addition, respondents received conditional incentives for completion of several of the health-based measures collected at IP12. In particular, respondents received £5 conditional on completion of each of the biomarkers needing outside of the interview completion. These additional measures include participation in:

1. Completing a blood pressure measure on their own prior to the interview

2. Mailing back dried blood samples
3. Mailing back a hair sample

There is no experimental aspect to these conditional incentives; all respondents received an incentive for completion of each of these.

At Wave 13, the initial unconditional incentives sent to households were identical to what was sent at IP12, which was a simplified version of what was done at IP11, regardless of mode allocation at IP13. Households within PSUs were allocated the treatments such that all individuals within households received the same incentive, and all split-off households received the same incentive as their originating household.

Additionally, all previous wave non-responding households which were issued to field received an incentive conditional on their survey participation at the value of their previously assigned group value.

In addition to the ongoing incentive experiment, as part of the Fieldwork Compression experiment at IP13 (see Section 7.19), there was an additional incentive given depending on experimental condition. 80% of respondents were offered a longer survey, with ½ of this (40%) told of the longer interview in the advance letter, and ½ (40%) asked to answer additional questions within the survey. Respondents asked in the advance letter were sent an additional unconditional incentive of £5; those asked in the survey to answer more questions were offered a conditional incentive of £5.

At Wave 14, the incentive experiment was largely ended, as much information has been gleaned from incentive experiments and in order to match the incentive structure on the main *Understanding Society* survey. Most respondents received a £20 unconditional voucher to complete the survey, matching what is done in the main survey. The exception to this amount was for IP respondents who in past waves have received £30; their incentive remained at £30 pounds, as reduction to their incentive may adversely impact response rates. This leaves two distinct incentive amounts to compare.

Respondents were allocated based on their previous wave allocation, last adjusted (through simplification) at IP12. The allocation at IP12 followed the controlling variable `ff_incentw12` which had households allocated to the incentive levels £10, £20, £30 and £10 + £20. The three groups allocated to incentives of £10, £20, and £10 + £20 were allocated to receive an incentive of £20, with households previously receiving £30 continuing to receive £30. Given this is entirely based on previously allocation, no new allocation is needed, and those in about 90% of households received £20, with the remaining 10% receiving £30. The controlling variable is set to Inapplicable (-8) for households from the wave 14 refreshment sample.

n_ff_incentw14 on record **n_hhsamp_ip**:

- 1 £20
- 2 £30

The wave 14 refreshment sample included an incentive experiment. Half the sample were offered £30 if they completed the survey online within a five-week deadline. The other half were offered £20 for completing the survey, plus £10 if they completed it within two weeks. The 'early bird' deadline was mentioned in all mailings. This experiment is controlled by the variable **N_FF_EARLYBIRD** on record **N_HHSAMP_IP**:

- 1 Two week time frame to complete (£20 plus £10 if early)
- 2 Five week time frame to complete (£30 at end)

At Wave 15, the incentive experiment comprised only the retention of the £30 condition for those who had previously received £30 incentives. Given this is entirely based on previous allocation, no new allocation was needed. Participants in about 90% of households received £20, with the remaining 10% receiving £30. The controlling variable is on record **o_hhsamp_ip**:

- 1 = £20
- 2 = £30

At Wave 16, the incentive treatments from wave 15 were carried forward. The controlling variable is **ff_incentw15** on record **p_hhsamp_ip**:

- 1 = £20
- 2 = £30

At Wave 17, the incentive treatments from wave 16 were carried forward. The controlling variable is **ff_incentw17** on record **q_hhsamp_ip**:

- 1 = £20
- 2 = £30

Table 7.1 Incentive experimental allocation FOR FF_GRIDMODEW11 = 1 & FF_LOWWEBW8 = -9 & FF_HHORIG == 11, 12, 13, 14, 18 (Continuous face-to-face respondents)

CODED VARIABLE											AMOUNT		
ff_group3	ff_incent w2	ff_incentw 3	ff_incentw 4	ff_incentw 5	ff_incentw 6	ff_incent w7	ff_incent w8	ff_incent w9	ff_incent w10	ff_incent w11	IP9	IP10	IP11
1	1	1	1	1	1	1	1	1	1	1	£10	£10	£10
			2	2	2	2	2	2	2	2	£10	£10	£10
2	2	2	3	3	3	3	3	3	3	3	£10	£10	£10
		3	4	4	4	4	4	4	4	4	£10	£10	£10
	3	4	5	5	5	5	5	5	5	5	£10	£10	£10
			6	6	6	6	6	6	6	6	£10	£10	£10
3	4	5	7	7	7	7	7	7	7	7	£10	£10	£10
			8	8	8	8	8	8	8	8	£10	£10	£10
	5	6	8	9	9	9	9	9	9	9	£10	£10	£10
<i>IP4 Refreshment sample</i>			9	10	10	10	10	10	10	10	£10	£10	£10
			10	11	11	11	11	11	11	11	£10	£10	£10
			11	12	12	12	12	12	12	12	£10	£10	£10
<i>IP7 Refreshment sample</i>						49	49	49	49	49	£10	£10	£10
						50	50	50	50	50	£20	£20	£20
						51	51	51	51	51	£30	£30	£30
<i>IP10 Refreshment sample</i>									88	88	£10		£10
<i>IP11 Refreshment sample</i>										92	£10		£10

Table 7.2 Incentive experimental allocation FOR FF_LOWWEBW8 = 1 & FF_GRIDMODEW11 = 1 & FF_HHORIZ = 11, 12 (low propensity web respondents now face to face)

CODED VARIABLE										AMOUNT		
ff_gro up3	ff_incen tw2	ff_incen tw3	ff_incen tw4	ff_incent w5	ff_incen tw6	ff_incentw 7	ff_incentw 8	ff_incentw 9	ff_incentw1 0	IP9	IP10	IP11
1	1	1	1	1	13	13	13	13	13	£10	£10	£10
					14	14	14	14	14	£10	£10	£10
					15	15	15	15	15	£10	£10	£10
			2	2	16	16	16	16	16	£10	£10	£10
					17	17	17	17	17	£10	£10	£10
					18	18	18	18	18	£10	£10	£10
2	2	2	3	3	19	19	19	19	19	£10	£10	£10
					20	20	20	20	20	£10	£10	£10
					21	21	21	21	21	£10	£10	£10
		3	4	4	22	22	22	22	22	£10	£10	£10
					23	23	23	23	23	£10	£10	£10
					24	24	24	24	24	£10	£10	£10
	3	4	5	5	25	25	25	25	25	£10	£10	£10
					26	26	26	26	26	£10	£10	£10
					27	27	27	27	27	£10	£10	£10
		4	6	6	28	28	28	28	28	£10	£10	£10
					29	29	29	29	29	£10	£10	£10
					30	30	30	30	30	£10	£10	£10

Continued ...

Table 7.2 continued

CODED VARIABLE											AMOUNT		
ff_grou p3	ff_incent w2	ff_incent w3	ff_incent w4	ff_incentw 5	ff_incent w6	ff_incent w7	ff_incentw 8	ff_incent w9	ff_incen tw10	ff_incen tw11	IP9	IP10	IP11
3	4	5	7	7	31	31	31	31	31	31	£10	£10	£10
					32	32	32	32	32	32	£10	£10	£10
					33	33	33	33	33	33	£10	£10	£10
	5	6	8	9	34	34	34	34	34	34	£10	£10	£10
					35	35	35	35	35	35	£10	£10	£10
					36	36	36	36	36	36	£10	£10	£10
	5	6	8	9	37	37	37	37	37	37	£10	£10	£10
					38	38	38	38	38	38	£10	£10	£10
					39	39	39	39	39	39	£10	£10	£10
IP4 Refreshment sample	9	10	11	12	40	40	40	40	40	40	£10	£10	£10
					41	41	41	41	41	41	£10	£10	£10
					42	42	42	42	42	42	£10	£10	£10
	10	11	12	13	43	43	43	43	43	43	£10	£10	£10
					44	44	44	44	44	44	£10	£10	£10
					45	45	45	45	45	45	£10	£10	£10
	11	12	13	14	46	46	46	46	46	46	£10	£10	£10
					47	47	47	47	47	47	£10	£10	£10
					48	48	48	48	48	48	£10	£10	£10

Table 7.3 Incentive experimental allocation FOR FF_GRIDMODEW11 = 3 & FF_LOWWEBW8 = 0 (Continuing WEB respondents)

CODED VARIABLE											AMOUNT		
ff_grou p3	ff_incent w2	ff_incent w3	ff_incent w4	ff_incent w5	ff_incent w6	ff_incen tw7	ff_incen tw8	ff_incent w9	ff_incent w10	ff_incen tw10	IP9	IP10	IP11
1	1	1	1	1	13	13	52	52	52	52	£10	£10	£10
					14	14	53	53	53	53	£10+£20	£10+£20	£10+£20
					15	15	54	54	54	54	£30	£30	£30
			2	2	16	16	55	55	55	55	£10	£10	£10
					17	17	56	56	56	56	£10+£20	£10+£20	£10+£20
					18	18	57	57	57	57	£30	£30	£30
2	2	2	3	3	19	19	58	58	58	58	£10	£10	£10
					20	20	59	59	59	59	£10+£20	£10+£20	£10+£20
					21	21	60	60	60	60	£30	£30	£30
		3	4	4	22	22	61	61	61	61	£10	£10	£10
					23	23	62	62	62	62	£10+£20	£10+£20	£10+£20
					24	24	63	63	63	63	£30	£30	£30
	3	4	5	5	25	25	64	64	64	64	£10	£10	£10
					26	26	65	65	65	65	£10+£20	£10+£20	£10+£20
					27	27	66	66	66	66	£30	£30	£30
		6	6	6	28	28	67	67	67	67	£10	£10	£10
					29	29	68	68	68	68	£10+£20	£10+£20	£10+£20
					30	30	69	69	69	69	£30	£30	£30

Continued...

Table 7.3 continued

CODED VARIABLE											AMOUNT			
ff_grou p3	ff_incent w2	ff_incent w3	ff_incent w4	ff_incent w5	ff_incent w6	ff_incen tw7	ff_incen tw8	ff_incen tw9	ff_incen tw10	ff_incen tw11	IP9	IP10	IP11	
3	4	5	7	7	31	31	70	70	70	70	£10	£10	£10	
					32	32	71	71	71	71	£10+£20	£10+£20	£10+£20	
					33	33	72	72	72	72	£30	£30	£30	
				8	34	34	73	73	73	73	£10	£10	£10	
					35	35	74	74	74	74	£10+£20	£10+£20	£10+£20	
					36	36	75	75	75	75	£30	£30	£30	
	5	6	8	9	37	37	76	76	76	76	£10	£10	£10	
					38	38	77	77	77	77	£10+£20	£10+£20	£10+£20	
					39	39	78	78	78	78	£30	£30	£30	
IP4 Refreshment sample				9	10	40	40	79	79	79	79	£10	£10	£10
						41	41	80	80	80	80	£10+£20	£10+£20	£10+£20
						42	42	81	81	81	81	£30	£30	£30
				10	11	43	43	82	82	82	82	£10	£10	£10
						44	44	83	83	83	83	£10+£20	£10+£20	£10+£20
						45	45	84	84	84	84	£30	£30	£30
				11	12	46	46	85	85	85	85	£10	£10	£10
						47	47	86	86	86	86	£10+£20	£10+£20	£10+£20
						48	48	87	87	87	87	£30	£30	£30
IP10 Refreshment sample							88	89	£10	£10				
							88	90	£10	£10+£20				
							88	91	£10	£30				
IP11								93	£10+£15					

7.2 Conditional and unconditional incentives

Research on the use of incentives generally finds that unconditional incentives are more effective than conditional incentives. However, on a longitudinal study, after the first couple of waves, it may be that sending unconditional incentives to previous non-responding individuals within responding households is a waste of resources, which could be more effectively re-directed elsewhere. This experiment randomly allocated households to two groups. In one the incentive was sent unconditionally to every adult in the advance letter as usual. In the other, only those adults who participated in the previous wave were sent an unconditional incentive in advance. That is, if **ff_condincw7** = 2, conditional incentives were given to all respondents with **ff_ivlowlw** = 2, 3 and any rising 16 year olds with **ff_ivlowlw** = 5. Unconditional incentives were sent in advance to all respondents with **ff_ivlowlw** = 1 and all 16 year olds with **ff_ivlowlw** = 4.

Since the range of incentives is larger in the Innovation Panel than the main-stage, incentive cards were used instead of paper vouchers. Previous wave non-responding adults in the treatment group (**ff_condincw7**=2 & **ff_ivlowlw**=2/3/5) who did respond at IP7 were given or sent their incentive post-interview. Where the adult was interviewed in person by an interviewer, this required the interviewer to write the incentive amount on the front of a blank incentive card, hand it to the respondent and make a note of the serial number and amount required in CAPI. When the interviewer dialled in, this triggered the Operations Department to activate that card with the appropriate amount. Where the individual was interviewed online or by telephone, this triggered the sending of a pre-activated incentive card with the correct amount.

Households within PSUs were randomly allocated to treatments.

Controlling variable(s): **ff_condincw7** on record **hhsamp**

Group 1 – Unconditional Incentive control group

Group 2 – Unconditional/Conditional Incentive Treatment

NOTE that the condition/unconditional incentive experiment interacts with the incentive amount experiment and as the actual amount of incentive was determined by the value of **ff_incentw7**.

7.3 Mixed mode experiments: telephone and face-to-face

There is considerable interest in using mixed mode strategies for conducting social surveys, in particular to enhance response rates and fieldwork efficiency while at the same time reducing costs. There is also interest in mode differences in survey measures. In Wave 2, Computer Assisted Personal Interviewing (CAPI) and Computer Assisted Telephone Interviewing (CATI) were examined. In Waves 5 and 6, we contrasted CAPI and Web interviewing.

In the Wave 2 mixed mode experiment, the IP2 sample was divided into three equal sized experimental groups, and each group received a different treatment in terms of questionnaire mode and sequence of modes. Within PSUs, households were randomly

assigned to experimental treatment – and all individuals within households were treated the same way. There were three experimental groups:

Group 1, CAPI: Households in this group were only eligible for face-to-face interviews.

Group 2, CATI “Move one, move all”: All households in this group were issued for telephone interviewing; if one person could not be interviewed by telephone all remaining members were transferred to CAPI. CATI was the mode of first contact, if that mode failed; the case was transferred to face-to-face. If enumeration was completed in CATI, individual interviews would be attempted within CATI until any household member indicated that they were unable to complete the interview by telephone. This may be because they refused, were classified as a noncontact or were unwilling/unable to complete the interview by phone. As soon as one individual interview could not be obtained, all outstanding household members were allocated to field and attempts were made to interview the remaining sample members face-to-face.

Group 3, CATI “Try all”: All households in this group were issued to telephone interviewing and more attempts were made to complete interviews by telephone. CATI was the mode used to contact the household. If that mode failed; the case was transferred to face-to-face for follow-up. If enumeration was completed in CATI, individual interviews were attempted within CATI. Households in this group were only eligible to be transferred once attempts had been made to contact and interview all household members by telephone. Attempts to contact each household member by telephone continued even if one household member was unable to be interviewed by telephone. The case was transferred to field for face-to-face interviews only after all eligible adults had either refused, were classified as non-contacts or were unwilling/unable to complete the interview by phone.

Experimental allocation is given by the variable **b_ff_modew2** in the data file **b_hhsamp_ip**. It is important to realise that the household enumeration and household questionnaire could be completed in one mode, with the individual questionnaires being completed in different modes from these or from other individuals interviewed in the household. The variable **b_trtocapi** on the record **b_hhsamp_ip** indicates that the case was transferred at some point from CATI to CAPI. The variable **b_modetype** on the record **b_hhsamp_ip** indicates the mode the household enumeration was completed in, the variable **b_hhmodetype** in the data file **b_hhresp_ip** indicates the mode of administration for the household questionnaire. Finally, the variable **b_indmode** in the data files **b_indall_ip** and **b_indresp_ip** indicates the mode of administration for the individual questionnaire.

The same questionnaires were used for CAPI and CATI, with only some necessary adaptations for telephone, such as dropping references to showcards.

Twenty households and 37 individuals were issued face-to-face and interviewed face-to-face but used the telephone instrument. The variable **b_modeallerr** on the datafile **b_indresp_ip** flags the cases with this error in mode allocation.

7.4 Mixed mode experiments: web and face-to-face

At Waves 5, 6, 7, 8, 9, 10, 11, 12, 15, 16, 17 the Understanding Society IP was used to investigate the use of web interviewing. The incorporation of web into a mixed mode design has potential both to reduce survey costs and improve quality. Details of the handling of these mixed-mode allocations at waves 13 and 14, which differed due to the COVID-19 pandemic, are also provided below.

Please see the end of this section for a documentation of the various outcome variables for these waves.

The Wave 5 sample had two components: the original sample, for which this was the 5th wave, and the refreshment sample, for which this was the 2nd wave. Households in both samples were randomly assigned within PSUs to one of two treatment groups. The controlling variable is `w_ff_gridmodew5` on record `w_hhsamp`:

Group 1 - Face-to-face (one-third of each sample);

Group 2 - Mixed mode (two-thirds of each sample).

The distribution of the Wave 5 issued sample of households across samples and mode treatments is summarised in the table below. The randomisation was implemented across PSUs, so that each sampling point contained a mix of households in each treatment group.

The face-to-face treatment involved standard *Understanding Society* procedures. Each adult sample member (aged 16 or over) was sent an advance letter with an unconditional incentive, after which interviewers called to attempt computer-assisted personal interviewing (CAPI) interviews. The value of the incentive (in both samples) was subject to experimental allocation. In each household one person was asked to complete the household grid and household questionnaire. All household members aged 16 or over were asked for an individual interview and to complete a self-completion questionnaire, which was randomly allocated to be either a computer assisted self-interview (CASI) or a paper questionnaire booklet. Young people (in this group) aged 10-15 were administered a paper self-completion questionnaire.

Table: Number of households allocated to experimental groups				
	Original Sample		Refreshment sample	Total
	Responded at Wave 4	Did not response at Wave 4		
Face-to-face	321	43	168	532
Mixed modes	618	110	315	1043
Total	939	153	483	1575

Note: Numbers shown are the numbers of households issued to the field, based on information held prior to the start of field work. During the course of field work, additional (split) households were identified. In the Refreshment sample, only responding households from their first wave (Wave 4 of the panel) were issued at Wave 5.

Fieldwork procedures for the web mode treatment were as follows: sample members aged 16 or over were sent a letter with the unconditional incentive, inviting them to take part by web. The letter included the URL and a unique user ID, which was to be entered on the welcome screen. A version of the letter was additionally sent by email to all sample members for whom we had an email address. For people who had indicated at previous waves that they do not use the internet regularly for personal use, the letter said that they would also have the opportunity to do the survey with an interviewer. Up to three email reminders were sent at 3-day intervals. Sample members who had not completed the web interview after two weeks were sent a reminder by post and interviewers started visiting them to carry out CAPI interviews. **Note** that this was the two-week web-only period in which households allocated to the additional £20 incentive conditional on whole household treatment at Wave 6 could qualify for the additional incentive (See PROCEDURAL EXPERIMENTS: INCENTIVES AND RESPONSE). The web survey remained open throughout the fieldwork period.

The first household member to log on to do the web survey was asked to complete the household grid, which collects information on who is currently living in the household. The web grid included an additional question to identify who is responsible for paying bills. The household questionnaire could be completed by either this person or their spouse/partner. For these sample members the household questionnaire was displayed first, then leading on to the individual questionnaire. (The household questionnaire is relatively short – around 10 minutes – and collects household-level information such as housing tenure, rent/mortgage payments, expenditure, utility bills, household consumer durables and some measures of material deprivation.) Once one partner had completed the household questionnaire, it would not appear for the other partner.

At Wave 5, the youth survey was administered either on paper or by web, depending on the mode used by the parent(s). If the parent(s) had responded by web and we had their email address, an invitation was sent to the parent by email with a request to forward it to their child. If the parent had been interviewed in CAPI, the interviewer gave the youth the paper self-completion questionnaire. Otherwise, a questionnaire was sent by post.

The adult web questionnaire was based on the CAPI one, with some adaptations, e.g. incorporating interviewer instructions into question wording, removing references to showcards, and making “help” screens more respondent-appropriate. There were differences in the visual display of items between the web survey and the computer-assisted self-completion portion of the CAPI administered survey. Notably, self-completion components as part of CAPI were self-administered using the standard Blaise visual presentation ordinarily seen by interviewers. On the web survey, where a series of items utilized the same set of response options, the items were formatted in a grid rendering a difference in the visual presentation of these items across modes.

At Wave 5, the web survey was not suitable for completion using a small mobile device (e.g. smart phone). If a mobile device was used to access the log-on page, the respondent was automatically directed to a page requesting that they log on from a computer.

The Wave 5 mixed mode treatment also included two (crossed) experiments to test ways of increasing web response rates:

- Half the households were offered an additional conditional incentive: if all eligible household members completed the web survey within two weeks, they each received an additional £5. This was mentioned in the advance letters to all household members in this treatment group.
- Half the households were sent the advance letter and first email to arrive on a Friday. The other half were sent them to arrive on a Monday.

At Waves 6, 7, 8 and 9, households who were not completed at the end of the standard face-to-face fieldwork period, and were not adamant refusals, were contacted again in a 'mop-up' stage of fieldwork. This included non-responding individuals in partially responding households. The nature of the mop-up contact was differentiated, however, by mixed mode allocation:

- In the mixed mode group, the 'mop-up' contact was made by telephone. The telephone interviewer reminded the sample member that they could participate on the web, but was also able to administer the Wave 6 interview by telephone (CATI). Cases for which a telephone number was not known were not contacted again at the mop-up stage.
- The face-to-face group was contacted to offer a web interview during the 'mop-up' stage. Individuals were sent a letter with the URL of the web instrument and their unique log-on code. Those for whom we had email addresses, this invitation was sent by email. A few days later, a telephone interviewer contacted all those for whom phone number was known in order to remind them of the web questionnaire, and to administer a telephone interview if possible.
- At Wave 7, The IP7 refreshment sample was not included in the web or telephone mop-up. Outstanding refreshment sample households and individuals continued to be attempted face-to-face during this period. At wave 8 and 9, the IP7 refreshment sample was also included in the mop-up phase.

At Wave 8, allocation to mode remained broadly similar to past waves such that the IP6 and IP7 "Face-to-Face first" sample remained "Face-to-Face first" at wave 8, and the IP6 and IP7 "web first" sample remained "web first". However, a subgroup of households previously allocated to the "web first" group were deemed to have very low web propensity and so moved to the "Face-to-Face first" group. Web propensity was determined through modelling observed characteristics, including mode of completion for previous waves. Web propensity has been calculated for the whole sample, including cases that at IP6 and IP7 were allocated to the "Face-to-Face first" group. The IP7 refreshment sample remained Face-to-Face first at IP8.

At Wave 9, 2/3 of households in the IP7 refreshment sample were randomly allocated to the "web first group" and the remainder to the "face-to-face first group". Besides this change, the mixed mode allocation remained the same as IP8.

In all waves and subsamples, allocation to mode was made in advance of fieldwork at the household level and (in some cases) depended on prior mode allocation and Web mode

response propensity as outlined above. The fed-forward variable **ff_gridmodew8** controlled allocation to mode. Controlling variables are on record **w_hhsamp_ip**:

w_ff_gridmodew8

Group 1	F2F
Group 3	WEB

w_ff_lowwebw8 (Flag for low web-propensity)

-9	Inapplicable, previous wave F2F cases
0	Web allocation
1	Low Web propensity

At Waves 6-9, the mixed mode treatment included an incentive experiment to test an incentive plan that might increase whole household web response rates:

- One-third of mixed mode households were offered a £30 pound unconditional individual incentive to participate
- One-third of mixed mode households were offered a £10 unconditional individual incentive with an additional £20 per individual if the whole household completed within the two-week web-only period.

Please see Experiment 7.1 for details on these incentive experiments.

At Wave 6 to 9, the youth survey was administered only on paper. Interviewers distributed paper self-completion questionnaires to youth whose parent(s) were interviewed in CAPI. Youth whose parents were interviewed by web were sent a questionnaire by post.

At Wave 10, the allocation to mode remained broadly similar to past waves such that IP1/IP4/IP7 Face-to-Face sample remained Face-to-Face and the IP1/IP4/IP7 Mixed-Mode sample remained Mixed-Mode at IP9. The wave 10 refreshment sample was allocated to Face-to-Face. The 'mop-up' stage at the end of the standard fieldwork period was as in prior waves. The controlling variable randomised at the household level is:

ff_gridmodew10 in record **J_HHSAMP_IP**:

1	F2F
3	WEB

At Wave 11, the allocation to mode remained the same as in previous waves for the IP1/IP4/IP7 samples, with Face-to-Face allocated sample members remaining Face-to-Face at IP11, and the IP1/IP4/IP7 Mixed-Mode sample remaining Mixed-Mode at IP11.

As with all previous refreshment samples, the IP10 refreshment sample was allocated to Face-to-Face only in the wave in which it was recruited (IP10). For IP11, 1/3 of this sample was randomly allocated to the Face-to-Face only design, with the other 2/3 allocated the Mixed-Mode design.

The IP11 refreshment sample was randomly allocated to modes from the start: 1/2 of households were allocated to Face-to-Face only and 1/2 to Mixed-Mode. The IP11 reserve

refreshment sample that was issued towards the end of fieldwork was however allocated to Face-to-Face only. The contact letter for households allocated to the Mixed-Mode design included a URL and sign-in information to complete the survey. The variable controlling allocation for this experiment at IP11 is:

ff_gridmodew11 in record **k_hhsamp_ip**:

- | | |
|---|-----|
| 1 | F2F |
| 3 | WEB |

The variable that identifies whether the IP11 refreshment sample members were part of the initial batch or the reserve sample is **ff_samplestatusw11** in record **k_hhsamp_ip**:

- | | |
|---|---------------------|
| 0 | Existing household |
| 1 | Main refreshment |
| 2 | Reserve refreshment |

The IP11 refreshment sample included an addition experiment with the invitation to the mixed mode survey (see 7.16 below).

At Wave 12, the goal was to compare standard design modes, i.e. CAPI and Web to a standard of nurse collected biomarkers. Three strands of data collection were therefore conducted: Nurse-led, Interviewer-led (standard CAPI), and mixed-mode (sequential web-CAPI). Both interviewer- and nurse-led data collections included measurements of blood pressure and height and weight. Nurse-led interviews included an additional collection of full blood, dried blood spots and a hair sample. Interviewer-led and web respondents were given a kit containing materials to enable dried blood spots and hair samples to be taken independently and returned at a later time (see Section 9.28). All participants were asked to collect a blood pressure measurement prior to completion of visit or web data collection, incorporating an experiment to examine prosocial versus informational content on response and quality of measurement (see Section 7.18).

One third of households were allocated to each of these three modes of data collection. Allocation to mode at IP12 was independent of any allocations at prior waves or sample status. As in past waves, respondents in the web-first condition were first invited to complete an interview on the web; those that did not complete the web survey were then assigned to an interviewer for completion.

The variable controlling allocation for this experiment at IP12 is **ff_gridmodew12** in record **l_hhsamp_ip**:

- | | |
|---|--------------|
| 1 | Face-to-face |
| 3 | Web-first |
| 4 | Nurse |

At Wave 13, respondents were allocated to one of two modes: Face-to-Face first and Web-first. However, due to the outbreak of COVID-19 and the measures put into place to limit its spread through personal contact, all households were invited to complete the survey in the web-first design. All follow-ups of web non-respondents by an interviewer were through telephone. Overall, the change in modes had little impact on content asked to respondents.

The one exception was that those responding to the telephone follow-up were not asked most of the self-completion modules, which is standard practice in past years of telephone mop-up phases.

While not implemented in practice due to the outbreak, below are details of the IP13 mixed-mode experiment. Given allocation will likely carry over waves, as in the past, this design will likely reflect what will be done in IP14. The allocation to Face-to-Face first or Web-first followed what was done from IP5 through IP11, however, allocation was new, so respondents were not necessarily in the same mode condition as they were in past waves. As with these previous allocations, about 1/3 of the households were allocated to the Face-to-Face first design, with the other 2/3 allocated to the Web first design.

The variable controlling allocation for this experiment at IP13 is **ff_gridmodew13** in record **m_hhsamp_ip**:

- 1 Face-to-face first
- 3 Web-first

At Wave 14, all sample members were issued to Web-first and followed up by telephone, due to the ongoing COVID-19 restrictions.

At Wave 15, the prior mixed mode allocations were resumed. The IP14 refreshment sample households were allocated 2/3 to web-first and 1/3 to face-to-face first.

The variable controlling allocation for this experiment at IP15 is **ff_gridmodew15** in record **o_hhsamp_ip**:

- 1 Face-to-face first
- 3 Web-first

At Wave 16, the prior mixed mode allocations were again maintained. The variable controlling allocation for this experiment at IP16 is **ff_gridmodew16** in record **p_hhsamp_ip**:

- 1 Face-to-face first
- 3 Web first

At Wave 17, the mixed mode allocations were maintained. The variable controlling allocation for this experiment at IP17 is **ff_gridmodew17** in record **q_hhsamp_ip**:

- 1 Face-to-face first
- 3 Web first

Indicators of Mode Allocations and Outcomes in Waves 5, 6, 7, 8, 9, 10, 11, 12, 15, 16, 17

The controlling variables in the data which control mode allocation and related variables containing information on interview outcomes are outlined below. The letter “w” indicates the wave, taking the letters “e”, “f”, etc.

On record **w_hhsamp_ip**:

- **w_ff_gridmodew*** contains the treatment allocation, where * corresponds to the wave (so e_ff_gridmodew5 is the wave 5 allocation variable)
- **w_ivfhqo** indicates whether household questionnaire was completed, **w_hhgridmode** indicates the mode in which the household grid was completed
- **w_hhintmode** indicates the mode in which the household questionnaire was completed
- **w_ivfho** provides a summary of which instruments were completed by the household
- **w_hhmodes** indicates whether the instruments a household completed were done by face-to-face, web or telephone only, or by using a mix of modes. **Note**, a cross-tab against w_ivfho can be used to identify whether the household completed all instruments or whether some are missing at a particular wave.
- **w_web_outcome** indicates the household outcome at end of web-only period (available for waves 5, 6, 7 only)
- **w_f2f_outcome** indicates the household outcome at end of face-to-face interviewing period (available for waves 5, 6, 7 only)
- **f_tel_outcome** indicates household outcome at end of telephone mop-up period (available for waves 5, 6, 7 only)
- **w_ff_lowwebw8** flag for low web-propensity (from wave 8 onwards)

on record **w_indall_ip**

- **w_ivfio** indicates the individual interview outcome
- **w_indmode** indicates the mode in which the individual interview was completed

7.5 Paper vs CASI self-completion

This study was relevant to changes toward self-completion being via CASI (computer-assisted self interviewing) in the main Understanding Society survey in Wave 3. It tests the effects of paper vs. CASI self-completion on substantive measures and attrition. Attrition cannot be evaluated until later waves. The experiment was initiated at Wave 4 and continued until Wave 6.

At Wave 4, half of the Innovation Panel sample received the CASI instrument whereas the other half received the paper instrument. The refreshment sample was included in this experiment.

The controlling variable **d_ff_casiw4** is on record **d_hhsamp_ip** and is coded as follows:

Group 1 = CASI

Group 2 = Paper.

The variable **d_scac** on record **d_indresp_ip** records the actual use of the self-completion. It is coded as follows:

- 1 Accepted as self-completion
- 2 Accepted as self-completion but interviewer to complete due to reading or sight problems
- 3 Accepted as self-completion but interviewer or someone else to help translate
- 4 Refused self-completion
- 5 Not able to do self-completion

At Wave 5, households were randomly allocated to either receive the same self-completion mode as at Wave 4, or the other mode. At Wave 6, only those groups that switched mode at Wave 5 were randomly allocated to the opposite condition. This design will give the opportunity of looking at the effects of (i) different and (ii) changing modes of the self-completion instrument have on the reliability of longitudinal measures. The table below outlines the experimental allocation at Waves 4, 5, and 6.

It should be noted that all respondents were allocated to an experimental treatment, but only face-to-face respondents received the actual implementation.

Table: Experimental allocation to self-completion instruments		
Wave 4	Wave 5	Wave 6
1 – CASI	1 – CASI	1 – CASI
	2 – Paper	2 – Paper
2 – Paper	1 – CASI	1 – CASI
		2 – Paper
	2 – Paper	1 – CASI

The Wave 5 controlling variable **e_ff_casiw5** is on record **e_hhsamp_ip**, while **f_casiw6** on record **f_hhsamp_ip** controls the Wave 6 allocation. Both variables are coded as follows:

- 1 CASI
- 2 Paper

As with Wave 4, the Wave 5 variable **e_scac** on record **e_indresp_ip** records the actual use of the self-completion. in wave 6, the variable is **f_scac** on record **f_indresp_ip**. Both items are coded as follows:

- 1 Accepted as self-completion
- 2 Accepted as self-completion but interviewer to complete due to reading or sight problems

- 3 Accepted as self-completion but interviewer or someone else to help translate
- 4 Refused self-completion
- 5 Not able to do self-completion

Note, due to a programming error at Wave 5 around 50 per cent of those eligible to receive the questions in face-to-face CASI mode did not get asked the experimental questions (313 people, based on unedited data). It should be noted that this does not confound the experiment (i.e. no respondents were asked questions in the wrong mode), but this error does reduce its power to detect mode differences. Cases affected by this programming error are flagged with the item **e_casiflagger** on record **e_indresp_ip**. Please see Section Known Data Issues for details concerning programming errors in the Wave 5 questionnaire.

7.6 Advance materials: letters vs. cards

The Wave 2 experiment compared a formal letter on printed stationery with a greeting card type format. The messages in cards and letters were equal in length, text, incentives, and information related to legitimacy of the study and privacy concerns. Thus, differences were only in appearance and format. The envelopes for both cards and letters were personally addressed. The letter was also personally addressed internally. The experimental treatments were as follows:

- 1 A formal letter on printed letterhead stationery with the survey logo
- 2 A card in a greeting card format

Within PSUs, households were randomly allocated to one of two treatment groups and all eligible respondents in the household were treated the same. The controlling variable is **b_ff_advance** on the record **b_hhsamp_ip**.

7.7 Advance materials: content of advance letters

In Wave 5, the content of the advance letters was varied to test theories about how people can be persuaded to take part in surveys. A two factor manipulation tested the effects of persuasion ideas, being a helpful person and being similar to other respondents. In the advance letters, one half of sample members received an additional sentence “your responses in previous survey show that you are a helpful person”; the other half of the sample had no such sentence. Second, one-half the sample received the sentence “almost everyone like you responded in the last wave of the survey” and the other half received no such sentence:

- 1 Helpful person
- 2 Respondents like you
- 3 Helpful person, Respondent like you
- 4 Control group

Within PSUs, households were randomly allocated to treatment groups and all eligible respondents in the household were treated the same. The controlling variable is **e_ff_persuasionw5** on the record **e_hhsamp_ip**.

In addition, the self-completion questionnaire has items to assess self-rated helpfulness, conformity, and preference for consistency as potentially useful predictors of response status in future waves.

7.8 Targeted advance letters

At Wave 6, multiple versions of the advance letter were used to test whether letters targeted at particular sample subgroups and referenced issues of likely importance to them could positively affect participation rates.

Half the sample received a standard advance letter. The other half of the sample received a 'targeted' letter. The 'targeted letter' enlisted respondents by saying that Understanding Society "helps researchers and policy makers understand the changes in the needs of the country across diverse subjects..."

- For 16-29 year olds the letter referenced "...subjects like the impact of the economic climate on employment prospects and the influence of mobile technology on life"
- For the employment-busy (working 39+ hours, or working 30+ hours and commuting 60+ minutes one way) the letter referenced "...subjects like your work-life balance, your position on your employment and your retirement"
- For those who had dependent children under age 15 the letter referenced "...diverse subjects like the provision of child care, schooling and education"
- For those living in London the letter referenced "...subjects like the cost of living and the provision of schools, housing and public transportation"
- For those of pensionable age the letter referenced "...subjects like the provision of social care and the cost of energy and fuel"

Random allocation to treatment of standard versus tailored letter was of households within PSUs. The controlling variable is **f_ff_advancew6** on record **f_hhsamp_ip** which takes the values:

- | | |
|---|----------------------------|
| 1 | 'Standard' advanced letter |
| 2 | "Tailored" advanced letter |

The specific type of advance letter with respect to his experiment is indicated by the variable **f_ff_letterw6** on record **f_indsamp_ip** which takes the values:

- | | |
|---|-------------------|
| 1 | "Standard Letter" |
| 2 | "Employment busy" |
| 3 | "Children" |
| 4 | "Age 16-29" |
| 5 | "London or SE" |
| 6 | "Pensionable age" |

7.9 Number of mailings between interviews

Do multiple contacts between waves increase response, by reducing the proportion of untraced movers and increasing the sense of ‘belonging’ the sample member has with the study? Or could it have a negative effect, by annoying sample members and/or giving more chances to refuse? At the time of this experiment, Understanding Society had recently moved from one between-wave mailing per year to 3-4 mailings. This experiment attempted to gauge the effect this had on response at IP7 by allocating households at random to receiving one vs 2-3 mailings between IP6 and IP7. The standard procedure at the time was for the IP sample to be sent one between-wave mailing each year – usually around November. This experiment proposed that one half of households were sent three mailings in this period (September, November, February). The mailing content was the same that was produced for the main-stage mailings. This required that the random allocation for the mailings occurred earlier in the process for IP7 than scheduled (usually early-December), and had to be based on the IP6 sample rather than the fed-forward IP7 sample. Households within PSUs were randomly allocated to a control group and a frequent mailing treatment group.

Controlling variable(s) is **ff_mailingsw7** on record **hhsamp**:

- 1 Control Group
- 2 Frequent mailings

7.10 Different ways of asking respondents to register on participant website

At Wave 4, this experiment tested different ways of getting respondents to register on the Understanding Society web-site using the delivery of the between wave mailing as a vehicle for achieving this via either e-mail for those who have given us an e-mail address or by traditional postal mailing. It also tested whether the use of incentives helps in this process; half of respondents were incentivized with a choice of incentive including a traditional high street voucher or vouchers for Amazon or iTunes, or a donation to one of three charities including Help for Heroes, Oxfam, or the NSPCC. Respondents without e-mail were re-allocated to the equivalent non-email treatment: “**ff_alliwmw4**” contains the experimental allocation whilst “**ff_actiwmw4**” contains the actual allocation given these rules. The IP4 refreshment sample was excluded from this experiment and has a blank value for both variables. The table below provided the control variables used in the experiment.

Table: Controlling variable for registration experiment

	ff_actiwmw4	
ff_alliwmw4	Has E-mail?	No E-mail?
No-incentive, paper	ff_actiwmw4 = 1	ff_actiwmw4 = 1
No-incentive, e-mail	ff_actiwmw4 = 2	ff_actiwmw4 = 1
Incentive, paper	ff_actiwmw4 = 3	ff_actiwmw4 = 3
Incentive, e-mail	ff_actiwmw4 = 4	ff_actiwmw4 = 3

	ff_actiwmw4 = 5 (Already registered online)	
--	---	--

7.11 Improving interviewer observations about characteristics of address

This Wave 4 experiment examines the utility of observation data provided by interviewers. It is limited to the refreshment sample cases. The focus was on questions of evidence for children in the household or for access to personal transport such as a car or van. Version A of these questions used the original interviewer observations, which read as follows, “Based on your observation, is it likely that this address has a **car or van**?” with response options “Definitely has a car / van”, “Likely”, “Unlikely”, “Definitely does not have a car/van”, and “Cannot tell from observation”. The items continue, “Based on your observation, is it likely that this address contains one or more **children aged under 10** (including babies)?” with response options “Definitely has a child/children aged under 10”, “Likely”, “Unlikely”, “Definitely does not have a child/children aged under 10”, and “Cannot tell from observation”.

Version B of these items was written more objectively. First, “Standing outside, can you observe any signs of a car or a van belonging to this address?” with response options “Yes, probably belonging to this address”, “Yes, unsure whether belonging to this address”, and “No”. Next, “Standing outside, can you observe any signs of children under 10 (including babies) at this address?” with response options “Yes” and “No”.

Households were randomly allocated within PSUs. The controlling variable is **d_ff_arfexpw4** on the record **d_hhsamp_ip**. The experimental allocation is as follows:

- 1 “Version A” original wording
- 2 “Version B” alternative wording.

The substantive information is contained in **d_children_b**, **d_children_a**, **d_carvan**, and **d_carvan2**.

7.12 Effect of content of re-issue letter on refusal conversion

This experiment carried at **Wave 4** examined whether additional information included in the re-issue letter would help boost response at this stage. All households were allocated within PSU to a treatment group, but only re-issued households were included in the experiment. Version A received the standard re-issue letter, and Version B received the re-issue letter plus an additional leaflet explaining the importance of participating in the study.

The controlling variable is **d_ff_reissuesw4** on record **d_hhsamp_ip**. The experimental allocation is as follows:

- 1 Standard re-issue letter

7.13 Early bird scheduling: encouraging respondents to call interviewer to schedule appointment

This experiment carried at **Wave 4** tested whether response rates could be improved if respondents are encouraged to telephone their interviewer in advance of fieldwork to set an appointment to be interviewed. A portion of the sample was given an incentive for setting and keeping an appointment. The incentive was £5 if the respondent called to make an appointment and was interviewed.

All addresses were allocated to treatment groups. However, only prior wave productive households were included in the offer and the refreshment sample and suspected split-households were excluded.

The controlling variable is **_ip d_ff_apptsw4** on record **d_hhsamp** whereas **d_ff_samplecat** on record **d_hhsamp_ip** contains the actual manipulation as used in field at ip4 after prior unproductive and splits were excluded. **d_ff_apptsw4** is coded as follows:

- 1 Early-bird offer, with incentive
- 2 Early-bird offer, no incentive
- 3 No offer

d_ff_samplecat is coded as follows:

- 1 Early-bird offer, with incentive
- 2 Early-bird offer, no incentive
- 3 No offer
- 4 Refreshment sample

In addition, information about the appointment made and whether it was kept was recorded for all sample members. Variables containing this information are **d_conmth**, **d_ebmover**, **d_conoc**, **d_apptoc**, **d_apptday2**, **d_ebend**, **d_issue_num**, **d_conday**, **d_ebmovtype**, **d_apptmth**, **d_firstappt**, **d_appttime2**, **d_ebintro**, **d_conmethod**, **d_ebmovconf**, **d_apptday**, **d_conperson2**, **d_apptoc2**, **d_ebcontact**, **d_conperson**, **d_ebaddupdate**, **d_appttime**, **d_apptmth2**, and **d_ebothinf** on record **d_admineb_ip**.

7.14 Targeted weekday invitation emails

This experiment examined whether targeting respondents by sending email invitations to complete the survey on different days affects response outcomes and data quality. The targeted component reflects the idea that people may be more likely to respond on particular days, possibly due to preference or time availability.

This experiment involved only households in the mixed-mode sample of the IP who had been in the mixed-mode in previous waves. Half of these households were assigned to the control condition, where normal contact procedures were followed. The other half of the mixed-mode sample were assigned to the experimental condition and everyone in the household was sent an email invitation on the day predicted to be more likely to lead to response. This prediction sought to take advantage of the longitudinal context of the survey, where there is a wide range of information on panel household members and their response behaviour. Knowledge of the response behaviour was based on the paradata collected in the data collection process. Paradata from IP waves 5-8 was used to identify likelihood of respondent preference for a given day of the week depending on when the household interview was completed during the week in the past.

It was expected that an invitation email sent when the panel household was more likely to respond (have time to do so) could increase the chance of participation. It was also expected that targeted timings could improve response speed, as sample households were expected to respond more promptly. Individuals within these experimental households who did not initially respond were sent a reminder email based on their individually expressed preferred day, estimated based on response to the individual questionnaire.

The allocation of respondents for this experiment is controlled by the following variables:

i_ff_maildayhwh9 in file i_hhsamp_ip: Day of email invitation to the entire household

0	Sunday
1	Monday
2	Tuesday
3	Wednesday
4	Thursday
5	Friday
6	Saturday
7	Control
8	Unassigned

i_ff_maildayindw9 in file i_indsamp_ip: Day of email reminder to nonresponding individuals

0	Sunday
1	Monday
2	Tuesday
3	Wednesday
4	Thursday
5	Friday
6	Saturday
7	Control
8	Unassigned

7.15 Using prospect theory in wording of advance letters

The objective of this experiment is to understand whether a rephrasing of the appeal to altruism in advance letters and invitation letters could enhance participation rates. Specifically, Kahneman and Tversky's prospect theory states that requests based on avoiding a negative outcome are more likely to achieve compliance than requests based on achieving a positive outcome. In the longitudinal survey context, prospect theory implies that it should be more effective to emphasise the negative consequences of not participating than the positive consequences of participating. However, to date all advance and invitation letters on Understanding Society, including IP, have relied on appeals to altruism that solely emphasise positive consequences of participating. The objective, therefore, is to test whether co-operation rates depend on whether the appeal is worded positively or negatively, and whether this in turn depends on moderating factors such as length of time in the panel and previous participation behaviour.

There is one controlling variable for this experiment, for whether a household is assigned to the positive outcome or negative outcome wording of the advanced letter. Households were randomly assigned to one of these two conditions with an equal 50/50 split. The allocation of respondents for this experiment is controlled by the following variable:

ff_mailwordw10 in file j_hhsamp_ip: Outcome wording of advance letter at IP10

- 1 Positive outcome wording letter
- 2 Negative outcome wording letter

7.16 Invitation letters for mixed mode survey

This is the first time the IP has approached a new sample for a web survey. This experiment is only within this new sample, taken at IP11. It focuses on different ways of asking for the participation of other household members and with different ways of introducing the CAPI follow-up phase. Thus, the 840 web-first addresses were randomly allocated to four groups of 210 addresses each: two ways of asking for the participation of other household members were crossed with two different ways of introducing the CAPI follow-up phase.

For households being informed of other household member participation in the advance letter, there was a £10 unconditional incentive and a promise of an additional £15 for each person in the household (16+) who completed the questionnaire online within three weeks. There was a reminder letter after 7 working days, again mentioning the £15 conditional incentive and the deadline. Second reminder letter after 15 working days.

Upon completion of the household grid, if there is more than one adult in the household a screen conveys a message along the lines of “We would like to invite <name> to take part in the survey too. They too will receive £15 for doing so by <date>. Please either enter their email address (we will email them their own personal invite) or click here to print an invitation letter with their own unique entry code.”

For households not being informed of other household member participation in the advance letter, there was a £10 unconditional incentive and a promise of an additional £15. Reminder letters were as above. Upon completion of the household grid, the screen regarding other household members appeared, as above.

For households informed of the interviewer follow-up aspect in the invitation letter, the advance letter also stated that if they were unable to participate online, there would be an opportunity to be visited by an interviewer instead. There was a reminder letter after 7 working days, again mentioning the £15 conditional incentive and the deadline additionally mentioning the interviewer visit option. Households not being informed of this interviewer follow-up option did not have this information in the advance or 7-day reminder letters. For all groups, the second reminder letter sent after 15 working days announced that an interviewer would call and that each person who took part, either face-to-face or online, would receive £10.

Further, for all groups: Reminder letters were sent to the household, if no response yet. If the grid had been completed and at least one individual had not yet responded, the additional persons were sent an invitation letter appropriate to their group. Individuals who received the invitation letter at the first reminder stage received a personal second reminder if they had not yet responded. Individuals for whom the grid was completed subsequent to the first reminder stage received an invitation letter at the second reminder

stage. The allocation of respondents for this experiment is controlled by the following variable:

ff_invitew11 on record **k_hhsamp_ip**:

- 1 Inform of other HH members' participation, not of interviewer follow-up option
- 2 Do not inform of other HH members' participation, not of interviewer follow-up option
- 3 Inform of other HH members' participation, inform of interviewer follow-up option
- 4 Do not inform of other HH members' participation, inform of interviewer follow-up option

7.17 Spending Study 2

This is a follow up to Spending Study 1, which was conducted between IP waves 9 and 10 (see Section 10.4). The data from Spending Study 2 have been deposited with the UK Data Service (SN 8909, [10.5255/UKDA-SN-8909-1](https://www.ukdataservice.ac.uk/datacatalog/studies/study?id=8909)). The data from Spending Study 2 can be linked to the data from the annual IP interviews using the individual identifier "pidp".

As part of the study, respondents were asked to download and install an application for their smartphone and to use it to record their daily spending. This included an experiment to examine the impact of when the invitation to download the app is made. Half of the households were asked to download the app during the course of the IP11 interview. For face-to-face respondents in this treatment group, the interviewer was able to assist as needed. The remaining half of households were invited to download the app for the study in an interwave postal mailing.

The experimental control variable indicated below was allocated equally among households, within strata of combined mode allocation and sample origin. Samples originated at IP1, IP4, IP7, IP10, and IP11. The mode allocation refers to the mixed-mode design described above (7.1). For samples that are part of the mixed-mode allocation at IP11, allocation of when the invitation was made occurred for each sample with modes equally. For samples not included in the mixed-mode experiment at IP11, allocation of invitation timing occurred equally only within that sample. The variable controlling this experiment is:

ff_ininterview on record **k_hhsamp_ip**:

- 1 Invitation to download app made in-interview
- 2 Invitation to download app made interwave (postal letter)

Variables used for this experiment are on record **k_indresp_ip**:
appoutc, **appwhy1_code**, **appwhy2_code**

7.18 Invitation to complete pre-interview blood pressure measure: information treatment vs pro-social appeal

Wave 12 included an advance letter experiment whereby one-third of the sample was provided with information on their nearest pharmacy to enable blood pressure measurement. One-third included an altruistic/pro-social appeal text to the letter to encourage participants to get their blood pressure measured. The remaining sample did not receive any of these treatments – the control group.

Allocations to treatment were at household level so everyone in the household was treated the same.

All sample members received a conditional £5 if blood pressure was measured.

Group 1 – Information treatment

This group was given specific information about a pharmacy local to them that provides free blood pressure checking.

For sample members outside London, we had information about a nearby pharmacy that provides free blood pressure checks for around 95% of sample members. For those living in London, this fell to around 75%. Thus, the allocation to experimental group was stratified by region (London/outside London) to ensure that the sample was balanced between the three groups within each area. Within the information treatment group there were some sample members for whom we did not have the required information (around 5% of those outside London, 25% of those within London). These sample members were re-allocated after the initial allocation to the control group.

Group 2 – pro-social message

This group was sent an advance letter which included a sentence or two about the social benefits of getting one's blood pressure measured.

The variable controlling this experiment is **I_ff_bpinfo** on record **I_hhsamp_ip**:

- | | |
|---|-----------------------------|
| 1 | Information Treatment |
| 2 | Pro-social appeal Treatment |
| 3 | Control |

Variables used for this experiment are on record **I_indresp_ip**:

I_slfbpchk, I_slfbpday, I_slfbpmnt, I_slfbptim, I_slfbploc, I_slfbplocoth_code, I_slfbpdatasys, I_slfbpdatadia, I_slfbpdatapul, I_noslfbp1, I_noslfbp2, I_noslfbp3, I_noslfbp4, I_noslfbp5, I_noslfbp6, I_noslfbp97, I_noslfbpoth_code, I_slfbpprob, I_slfbpprres_code, I_debslfbp

7.19 Fieldwork compression experiment

Currently the main Understanding Society survey collects each wave of data over the course of two years. That is, in any given calendar year, half of the sample are completing the survey for wave t . The other half already completed wave t the year before and are now completing wave $t+1$. However, it is proposed that in a future wave, Understanding Society will compress fieldwork into a single year. This compression will mean that half the sample will skip the rotating content of one survey wave. As an alternative, these respondents could be asked to complete a longer survey that asks some or all of the rotating modules they would otherwise miss.

This experiment tests two ways in which participants are asked to complete these additional modules, and two sets of additional questions, a longer or shorter version. This leads to four experimental groups, as well a control group. This experiment also includes the use of additional incentives, either unconditionally or conditionally offered, depending on the design allocated to.

With the continuous, longer interview, the advance letter noted that the interview would be longer than usual, but during the interview no indication was given as to when the respondent was answering the additional modules. The letter noted that because the interview was longer, the incentive would include an additional unconditional £5.

With the potential 'break-off' point, there was a request towards the end of the standard IP13 interview to complete an extra set of questions. The participant was informed at this point that if they completed the rest of the interview, they would be given (in CAPI) or sent (in Web) an additional £5 voucher (conditional).

The control group received the standard IP13 questionnaire and no additional incentive.

The longer interview contained all of the modules planned for wave16, except one, and took about 15-25 minutes depending on respondent routing. The shorter interview contained five modules, all except one being rotating modules that would normally be asked of the whole sample, i.e. containing no routing, and took about 10 minutes. The one exception in both the longer and shorter version is that the Wealth and Assets module was dropped and replaced by a module on Panel Conditioning (see Section 10.7).

The allocation was randomised at the household level, with 1/5 of households allocated to each condition. The controlling variable is **ff_compressionw13** on record **m_hhsamp_ip**:

- 1 Continuous longer interview, full set of rotating modules
- 2 Potential break-off request, full set of rotating modules
- 3 Continuous longer interview, reduced set of rotating modules
- 4 Potential break-off request, reduced set of rotating modules
- 5 Control

The variables from the additional modules are:

fwcintro, jbperfp, jbonus, jbrise, tujbpl, tuin1, jbpen, jbpenm, jbpeny4, penmcn, penmpy, penmtp, penspb, wktime, wkends, jbflex, jbfxuse, jbfxfinf, wktech, wktech, wktechinet,

wkhome, wkaut1, wkaut2, wkaut3, wkaut4, wkaut5, depenth1, depenth2, depenth3, depenth4, depenth5, depenth6, jblkcha, jbxpcha, jblkchb, jbxpchb, jblkchc, jbxpchc, jblkchd, jbxpchd, jblkche, jbxpche, jbsec, envhabit1, envhabit2, envhabit3, envhabit4, envhabit5, envhabit6, envhabit7, envhabit8, envhabit9, envhabit10, envhabit11, carmiles, flyes, nflyin, nflyeu, nflyos, hrs_slph, hrs_slpm, tslp_30m, tslp_wak, tslp_cgh, med_slp, tsta_awk, slp_qual, trcarfq, trbusfq, trtrnfq, trbikefq, needcar, cyclepath, hubuys, hubuyshhm, hufrys, hufryshhm, humops, humopshhm, huiron, huironhhm, hupots, hupotshhm, hudi, hudiyyhm, hupayhswrk, hucdress, hucdresshhm, hucbed, hucbedhhm, hucunwell, hucunwellhhm, hucplay, hucplayhhm, huchomework, huchomeworkhhm, hucferry, hucferryhhm, huboss, howling, chargv, charfreq, charam, save, saved, savreg, savlt, workdis, journeysat, workpark, workparkpay, ppen, ppent, ppyrs, ppreg, ppam, ppampc, volun, volfreq, volhrs, scacfwc, scrffwc, scunfwc, scenv_ftst, scenv_crlf, scenv_grn, scenv_bccc, scenv_pmep, scenv_meds, scenv_crex, scenv_tlat, scenv_nowo, scenv_fitl, scenv_noot, scenv_canc, scopecl30, scopecl200, scwemwba, scwemwbb, scwemwbc, scwemwbd, scwemwbe, scwemwbf, scwemwbg, scopfama, scopfamb, scopfamd, scopfamf, scopfamh, condthink, conddo, conddo2, condatt, whypart, whypart2.

7.20 Event triggered data collection

In order to expand the data collected by the survey, and to collect data on life events closer to the time at which they occur, Understanding Society is testing methods to collect data on a more continual basis rather than just annually. The first round of ‘Event Triggered Data Collection’ was implemented in the IP from February 2020 until January 2021. At the beginning of each month, sample members were sent an invitation by email and/or SMS, inviting them to a web survey containing a single question about whether they had experienced any of a list of life events in the previous calendar month. If yes, they were asked a series of follow-up questions about each event reported. The data collected in these monthly surveys are deposited separately with the UK Data Service (SN: 8990, [10.5255/UKDA-SN-8990-1](https://www.ukdataservice.ac.uk/datacatalog/studies/study?id=10.5255/UKDA-SN-8990-1)) and can be linked with the data from the annual IP interviews using the individual identifier “pidp”.

The variable **ff_eventeligw12** on record **m_indsamp_ip** identifies sample members who were eligible for the event triggered data collection experiment. This included sample members who were in an IP11 respondent household ($k_{ivfho}=10$ | $k_{ivfho}=11$ | $k_{ivfho}=12$) and where at least one household member uses the internet at least several times a week:

ff_eventeligw12 “Eligible for event triggered data collection experiment”

- 1 Eligible
- 2 Not eligible

Not all eligible sample members were invited to the Event Triggered Data Collection: individuals in a random 70% of households were invited, individuals in the remaining 30% of

households were not invited and serve as the control group. This experiment is controlled by the variable **ff_eventtrigw12** on record **m_indsamp_ip**:

ff_eventtrigw12 "Allocation to event-triggered data collection"

- 1 Allocated to event triggered data collection
- 2 Not allocated to event triggered data collection

The variable **ff_eventissuw12** on record **m_indsamp_ip** identifies sample members who were issued to ETDC experiment. That is, they were eligible (**ff_eventeligw12**=1) and in the randomised treatment group invited to event-triggered data collection (**ff_eventtrigw12**=1):

ff_eventissuw12 "Issued to event-triggered data collection"

- 1 Issued
- 2 Not issued

The variables **ff_eventeligw12**, **ff_eventtrigw12** and **ff_eventissuw12** are set to Inapplicable (-8) for sample members who were not eligible for the annual interviews, for example because they were younger than 16, adamant refusers, or their address was unknown. In addition, there are 333 individuals who were to be issued to ETDC, but were returned from the wave 12 fieldwork as adamant refusals and were therefore dropped from the ETDC sample file.

The event triggered data collection included two experiments. The allocation variables for these are included in the sample file deposited with the event triggered data. The first is an incentive experiment (controlled by variable **ff_eventincetw13**): half of sample members were offered £1 for completing the monthly survey, half were offered £1 for completing the event question plus £2 if they reported any events. The second experiment concerned the timing of the two reminders that were sent after the initial invitation to the survey (controlled by variable **ff_eventremindersw13**): half of sample members were sent reminders daily, the other half were sent reminders every two days.

To acknowledge the event triggered data collection, the IP13 interview included some additional wording in the first question of the Annual Events History module ("calintro") and a debrief question about the event triggered data collection ("eventdebrief"). These were routed on the variable **ff_eventtrig_w12**. However an error in the sample file for IP13 meant that this variable was set to missing for all sample members. Therefore this additional wording and the debrief question were erroneously not shown to any sample members.

7.21 Consent to send survey questions by SMS

All IP13 respondents who said they use a mobile phone were asked for consent to be sent occasional survey questions via SMS. The location of this consent was experimental manipulated, being either early on in the survey (in the Demographics module) or at the end of the survey (in the Contact Details module). The controlling variable allocating respondents equally to the two conditions was generated within the questionnaire, with

equal allocation occurring at the respondent level. The variable controlling this experiment is **smsconsent** on record **m_indresp_ip**:

- 1 Consent asked in Demographics module
- 2 Consent asked in Contact Details module

The variables affected by this experiment are on the record **m_indresp_ip**: **consentsms_early**, **consentsms_late**.

At Wave 15, questions about SMS consent were asked again, but were not manipulated experimentally. Updates were made to the previous procedure by: repeating the consent for SMS questions for everyone who did not consent in IP13; the consent question was only asked in the Contact Details module; one follow-up question was added about reasons for not consenting; and a subset of questions from the IP12/IP13 Mobile Device Use module were repeated to understand respondents' mobile use.

7.22 Wellbeing app study

The Wellbeing app collected information on emotion and self-regulation, external stressors, attachment, and interactions with loved ones. During the Wave 13 interview all respondents who had completed at least one previous interview were invited to download the app and asked to use it every evening for 14 days. Respondents were offered £1 for every day on which they completed the app survey. The data collected in the app are deposited separately with the UK Data Service (SN: 9065, [10.5255/UKDA-SN-9065-1](https://ukdataservice.ac.uk/datacatalog/studies/study?id=10.5255/UKDA-SN-9065-1)) and can be linked to data from the annual IP interviews using the individual identifier "pidp".

The invitation to the Wellbeing app study included a number of experiments. These are: 1) the invitation to the study placed early vs. late in the Wave 13 survey; 2) the length of the daily app survey, 2 vs. 10 minutes; 3) the bonus incentives offered for completing all 14 days of the study, Control (none) vs. £10 for all 14 days completed vs. £2.50 each on four randomly selected days if they completed the survey on those days.

Login details for the app were created prior to the Wave 14 fieldwork for all sample members who had completed at least one previous interview. New sample members who completed the Wave 14 survey were routed around the invitation to the Wellbeing app study.

The experiments are controlled by several variables. All allocations were done at the household level. All allocations were equal, i.e. $\frac{1}{2}$ for two conditions, $\frac{1}{3}$ for three conditions, etc. All treatment allocations were crossed or stratified to ensure that the allocations for each experiment were balanced across the treatments to the other experiments.

The controlling variables are on the record **m_hhsamp_ip**:

ff_appposw13 (1/2 each)

- 1 Wellbeing App invitation early in IP13 interview
- 2 Wellbeing App invitation late in IP13 interview

ff_applengthw13 (1/2 each)

- 1 2 minute daily app survey
- 2 10 minute daily app survey

ff_appincentw13 (1/3 each)

- 1 No additional incentive
- 2 £10 if all 14 days completed
- 3 £2.50 each on four randomly selected days if app survey completed on those days

The sample file for the Wellbeing app study includes a final set of randomised allocation variables (***ff_rndaw13***, ***ff_rndbw13***, ***ff_rndcw13***, ***ff_rnddw13***) that control the random day on which bonuses were given for the group ***ff_appincentw13***=3.

The variables affected by this experiment are in the record ***m_indresp_ip***:

mobtext, ***appoutc1***, ***approb1***, ***welldebrief1***, ***welldebriefoth1***, ***casiintnofwc***, ***appoutc2***, ***approb2***, ***welldebrief2***, ***welldebriefoth2***.

7.23 Asking for Living Apart Together partner details

In Wave 13, respondents with partners that live apart were asked for the contact details of these partners non-experimentally (see Section 10.8). The contact details were collected in a manner similar to questions asked in the Stable Contact Details module. The request made clear that the information would potentially be used to invite their partner to complete the survey, but that no information about the respondent would be given to their partner, other than to say they gave the contact details.

The experiment in Wave 14 was conducted to improve respondents' willingness to provide partner contact details, employing two potential improvements. In the first condition (sample 2), respondents were again asked to provide contact details for their partner, but with more reassuring wording than what was used in Wave 13. This included wording saying the additional survey would be a one-off short survey. In addition, initial analysis of the Wave 13 data suggested respondents were more likely to provide a physical address for their partner than email, both of which were more likely to be given than a telephone number. The order of these questions was therefore changed in, such that physical address was asked first. Respondents declining to provide details were asked reasons why in an open-text question.

In the second condition, respondents indicating they have a partner living apart were sent information in the interwave mailing about how their partner could participate in the one-off short survey and asked to provide contact details via a webform or by post (sample 3).

For both conditions, respondents who had provided their partners' contact details in Wave 13 were not asked again.

The randomised allocation for this experiment was allocated at the household level, with half the sample allocated to each group. The variable controlling this experiment is **n_ff_ncrrexp14** on record **n_hhsamp_ip** (allocation stratified by **sampleorig**, **ff_incentw14**, **ff_gridmodew13**, **l_gor_dv**, with the latter being recoded to group together values 7&8, 9&10, 2&3, 4,5&6, 1&11):

- 1 Asked partner details in survey letter
- 2 Asked partner participation in interwave mailing

The variable affected by this experiment is in the file **n_indresp_ip: ncrresp**. the variables containing the contact details are not released with the data (questions **ncrrequest**, **ncrtitle**, **ncrfname**, **ncrsname**, **ncrctadd1**, **ncrctadd2**, **ncrcttown**, **ncrctcnty**, **ncrctpcode**, **ncremail**, **ncremail2**, **ncrcttel1**, **ncrcttel2**, **ncrreason**).

See also: 10.8 Living Apart Together (LATs) partner and survey

7.24 Contact protocols for IP14 refreshment sample

The wave 14 refreshment sample was part of a larger pilot in preparation for recruiting a refreshment sample for the main Understanding Society study. The pilot included a clustered sample that was added to the Innovation Panel data as a new refreshment sample (identified by **hhorig=19**), and an unclustered sample whom we do not plan to contact for future waves. The pilot was conducted as a web survey and included a series of experiments with the protocols for recruiting a new sample for a web survey. The experimental allocation variables are included in the wave 14 Innovation Panel data, although the data exclude the unclustered part of the pilot sample. The randomised allocation variables for these experiments are set to Inapplicable (-8) for the continuing samples, as they only applied to the wave 14 refreshment sample. For more details and results of the experiments from the full pilot sample see (Williams et al 2022).

Prenotification letters and reminders

This experiment varied whether sample households were sent a prenotification letter before the letter inviting them to the survey, and whether they were sent two or three reminders if they did not complete the survey. The variable **n_ff_prenotcon** on record **n_hhsamp_ip** controls this experiment:

- 1 Prenotification, invitation letter, first reminder, second reminder
- 2 Prenotification, invitation letter, first reminder, second reminder, third reminder
- 3 Invitation letter, first reminder, second reminder, third reminder

Logo/branding on outside of envelope

This experiment tested the use of different logos on the envelopes sent to the refreshment sample households. A third of households received envelopes with the logo of the Institute for Social and Economic Research (ISER) as well as the logo of the University of Essex; a third received the University of Essex logo only; and the control group received a blank envelope without logos. This experiment is controlled by the variable **n_ff_logocon** on record **n_hhsamp_ip**:

- 1 Logo is ISER AND University of Essex
- 2 Logo is University of Essex only
- 3 Blank (No logo)

Include a gift incentive

This experiment tested the effect of including a gift incentive in the invitation letter. The gift consisted of branded Post-it notes that featured the Understanding Society logo as well as branded Post-it notes that included the Understanding Society logo and an encouraging message ("Help shape the decisions that affect people like you... take part in the study today."). The control group received no gift incentive. 1/3 of households were allocated to each group. This experiment is controlled by the variable **n_ff_giftcon** on record **n_hhsamp_ip**:

- 1 Receive Post-its with invitation letter with message and logo
- 2 Receive Post-its with invitation letter with logo only
- 3 No Post-its

Explain the longitudinal nature of the survey

Half the sample were informed of the longitudinal nature of the study in the leaflet sent with the invitation letter ("The study involves one survey every 12 months. Your information is most valuable when we learn how it changes over time, so we will ask you to complete the study every year until you choose to leave."). The other half were told at the end of their interview. This experiment is controlled by the variable **n_ff_hhoriglong** on record **n_hhsamp_ip**:

- 1 Informed it is a longitudinal survey in information leaflet
- 2 Not informed it is a longitudinal survey in information leaflet

'Early-bird' incentive and deadline

Half the sample were offered £30 if they completed the survey online within a five-week deadline. The other half were offered £20 for completing the survey, plus £10 if they completed it within two weeks. The 'early bird' deadline was mentioned in all mailings. This experiment is controlled by the variable **n_ff_earlybird** on record **n_hhsamp_ip**:

- 1 Two week time frame to complete (£20 plus £10 if early)
- 2 Five week time frame to complete (£30 at end)

Explain household survey

This experiment tested different ways of informing the refreshment sample that the survey requires involvement from the whole household. Half of the sample were told up front in the letter that the study is a household survey. The other half were not told in the letter but informed after the first person in the household logged into the survey and completed the household grid. This experiment is controlled by the variable **n_ff_toldallad** on the record **n_hhsamp_ip**:

- 1 Not told in letter (only in HH grid)
- 2 Told in letter

Collecting email addresses of household members

This experiment tested asking the first person who completed the survey, and was therefore asked to complete the household grid, for the email addresses of all other adult household members. This would enable us to contact the other household members by email as well as by letter. The emails collected in this way were, however, not used; this was to test the feasibility of requesting these contact details. This experiment is controlled by the variable **n_ff_ohhemail** on the record **n_hhsamp_ip**:

- 1 Do not ask for email addresses of other HH members
- 2 Do ask for email addresses of other HH members

Include encouraging messages during the web survey to signal progress

Half the sample were shown encouraging messages part-way through the survey, the other half were not shown these messages. The following messages were used for all respondents in the treatment group and inserted in different locations in the questionnaire: "We're going to ask you some important questions about money. We know it's a sensitive subject so we want to remind you that all your answers will be kept safe and anonymous." "Thank you for all the information you've shared so far. We will keep it safe." "We're going to ask some sensitive questions, so we just want to remind you that your responses will be anonymised, and no one will know it's you." "We appreciate some of those questions may have been difficult to answer, so thank you for sharing. Now keep going - you're not too far from the end of the survey." This experiment is controlled by the variable **n_ff_encourmes** on record **n_hhsamp_ip**:

- 1 Exposed to the encouraging messages
- 2 Not exposed to the encouraging messages

7.25 PERKs for unconditional incentives

This experiment compared different methods of sending out the unconditional incentives that are sent to respondents with the prenotification letter. Using e-vouchers gives participants a greater choice in the type of incentive they receive. Rather than being restricted to physical Love2Shop gift-cards which must be used in physical stores, PERKs

would let participants choose vouchers that can be used online or a donation to charity, alongside the Love2Shop gift-card.

Adult sample members eligible for unconditional incentive were randomly allocated to two groups. One received the Love2Shop gift-card as standard. The second group was invited to go online and claim their unconditional PERKS voucher. This experiment affected the participant advance materials. There were also some 'motivational' messages added to the letter experimentally: (i) no additional message, (ii) e-incentives are more eco-friendly, (iii) e-incentives are easier to use, and (iv) both messages. Households were allocated to treatment groups in equal proportions. The allocation variables are in file **p_hhsamp_ip**:

ff_perksw16 (1/2 each, allocation stratified by **sampleorig ff_gridmodew16 ff_incentw16 n_gor_dv**)

1 = Receive incentive as usual
2 = Online option

ff_perksmsgw16 (1/4 each, allocation stratified by **sampleorig ff_gridmodew16 ff_incentw16 n_gor_dv**)

1 = No message
2 = Eco-friendly
3 = Easier
4 = Both

For initial findings, see [Parutis and Burton \(2024a\)](#).

7.26 Youth online survey

More households are completing the survey online, but the youth questionnaire has remained paper-based. Even after adding another reminder and copy of the questionnaire, the youth response is still much lower than when the study was interviewer-administered only. This experiment tests offering the youth survey online to counter this decline.

Children in a household that completed the grid online or by telephone were sent the invite to the youth survey by post. This included a paper questionnaire, a covering letter with a QR code and URL and access code to complete online, unconditional incentive, and a leaflet. In half of households, the covering letter mentioned that if the child completed the survey online or returned the questionnaire, they would be sent an additional £5 gift card. The leaflet was either one targeted to the young person, or targeted towards the parent.

When the interviewer completed the grid, they handed over the youth self-completion questionnaire, covering letter (with QR code/URL and access code to complete online), and unconditional voucher. For children in these households the offer of an additional conditional offer and the leaflet were included in the first reminder (sent from the office).

All of this was implemented outside the questionnaire, either by the fieldwork agency or by the interviewer in the household. Treatment groups were allocated at the household level. The controlling variables are in file **p_hhsamp_ip**:

ff_youthincentw16 (1/2 each, allocations stratified by **ff_incentw16 ff_gridmodew16 sampleorig**)

1 = £5 conditional
2 = None

ff_youthleafletw16 (1/2 each, allocations stratified by **ff_incentw16 ff_gridmodew16 sampleorig**)

1 = Child targeted
2 = Parent targeted

For initial findings from the experiment, see [Parutis and Burton \(2024b\)](#).

See also: 10.15 Youth online survey

7.27 Test of Government logo on envelopes for continuing sample

For the Wave 14 boost (year 2) we experimented with the use of the HM Government logo on the envelopes for mailings to half of addresses. Early indications are that this has significantly increased the proportion of households where a household grid is completed, and where the household completes online. This was implemented on the boost, and so the only information we have is the address – thus the letters are addressed to “The Occupier”. It is suspected that this increases the perception that the mailing is ‘junk mail’ and so not opened. The hypothesis is that the Government logo encourages someone to open the envelope, and then read the letter.

Originally, we had considered rolling out any successful interventions on the boost to the continuing sample. However, whereas the Government logo might encourage a fresh household to open the envelope, there may be a different effect for letters addressed to a named participant. Someone who has previously participated in the survey is more likely to be aware of the Understanding Society logo, and is more likely to open an envelope addressed to them personally. Previous participants in the study will also be aware that the envelope will contain their invitation to the next wave of the study, and a gift-card. The addition of a Government logo will be new, and may potentially discourage some sample members from participating.

We experimented with the use of the Government logo on the IP mailings in IP16. Households were allocated to one of two treatment groups in equal proportions. The variable controlling allocations is in the file **p_hhsamp_ip**:

ff_govtlogow16 (1/2 each, allocations stratified by **ff_gridmodew16** (reverse coded) **ff_incentw16 ff_perksw16 sampleorig**)

1 = Government logo on envelope

2 = No Government logo

For initial findings, see [Burton \(2024\)](#).

7.28 Youth survey invite mailing

Some literature shows that children like opening their own letters. This experiment aimed to investigate whether using letters addressed to children affects their propensity to complete the youth survey.

Understanding Society's standard approach at the time of this experiment was to send invites to the youth survey addressed to responsible adults (parents, carers, or guardians) in households where there is a child aged 10 to 15. These invites, addressed to the responsible adult, include the login details their child should use if they want to complete the survey online. If there are multiple eligible children in the household, the adult is sent a separate letter for each, each containing the login details for one child plus a copy of the youth questionnaire. For interviews conducted face-to-face, interviewers will write children's names on their letters and envelopes, and seal the envelopes, before handing them over during the interview; for web respondents the letters will be mailed.

In this experiment we allocated half of the households to get special youth invites (if there are any youths in the household): the main outer envelope and the cover letter were still addressed to the responsible adult but inside the adult's envelope there was another sealed envelope with the child's name on it and the letter inside addressed to the child. This inner letter contained the login details to complete the youth survey. As with the standard approach, if there are multiple eligible children within the household, for each child a separate letter was sent to the adult, and each of these contained a copy of the paper version of the youth questionnaire.

Allocations

The variable controlling allocation to treatments, randomised at the household level:

ff_ythletw17

(1/2 allocated to each condition, allocation stratified by **sampleorig**, whether any children aged 12-15 in the household last wave {recoded version of **p_nch1215_dv** grouping all non-zero values}, **ff_careexpw17**, **ff_incentw17**, **ff_gridmodew17**)

1 Standard youth invite within adult letter

2 Youth invite in envelope addressed to youth

8. Questionnaire design experiments: general issues

8.1 Subsetting the questionnaire content

In Wave 1, the Innovation Panel tested the feasibility of using random subsets of questionnaire content. The purpose was to explore the possibility of maximising survey content while minimising overall questionnaire burden for respondents. Three areas were covered by random subsets asked of half the sample--the Partnership and Fertility Histories, Environmental Behaviour and Attitudes, and Self-reported Height and Weight.

In Wave 2, the other half of the sample was asked the Partnership and Fertility History, whilst the full sample received questions on environmental behaviour and attitudes and also self-reported height and weight.

In Waves 3 and 5 and Waves 5 and 6, this two-wave pattern for the environmental, height and weight topics was repeated to provide evidence on the effect of rotation schedule in content on the response process.

Within PSUs, households were randomly assigned to subsets. In each household, all responding adults received content consistent with their treatment.

The variable in the data which controls which subset of survey content the respondent received is **A_GROUP2** on the record **a_hhsamp_ip**, then **w_ff_group2** on **w_hhsamp_ip** where “w” indicates waves with values “B” through “F” for Waves 2 through 6 respectively.

The specific details of the sub-sets are as follows:

Environmental Behaviour and Attitudes

Group 1	Not asked the set of questions
Group 2	Asked this set of questions

The relevant items are **a_opepplch**, **a_opeich**, **a_opecbn**, **a_opeflood**, **a_opeflduk**, **a_opefood**, **a_opefduk**, and **a_opecl30** on records **a_indresp_ip**. At Waves 2 and 3, the effected questions were **w_envhabit1_a** through **w_envhabit11_b**, **w_opecl30** and **w_opecl200** on record **w_indresp_ip** where “w” indicates waves and takes values “b” and “c” for waves 2 and 3 respectively. At Waves 4, 5, and 6 the effected questions were **w_envhabit1** through **w_envhabit11**, **w_opecl30** and **w_opecl200** on record **w_indresp_ip** where “w” takes values “D”, “E” and “F”.

Height and Weight

Group 1	Not asked the set of questions
Group 2	Asked this set of questions

The relevant items are **w_hlht**, **w_hlhtf**, **w_hlhti**, **w_hlhtc**, **w_hlwt**, **w_hlwts**, **w_hlwtp**, **w_hlwtk**, **w_hlwte**, **w_hlwtl**, and **w_hlpreg** on record **w_indresp_ip** where “w” indicates waves and takes values “A”, “B”, “C”, “D”, “E”, or “F”..

Partnership and Fertility History

Group 1	A “short-form” history
Group 2	The full life history

The list below shows the affected questions in the partnership and fertility history subsets:

Table: Partnership and fertility history items	
Group 1 (Short form history)	Group 2 (Full History)
A_COH1BM on record A_INDRESP_IP A_COH1BY on record A_INDRESP_IP A_CH1BM on record A_INDRESP_IP A_CH1BY4 on record A_INDRESP_IP	A_LCMCOH on record A_INDRESP_IP A_LCMCBM on record A_INDRESP_IP A_LCMCBY4 on record A_INDRESP_IP A_LCMSPM on record A_INDRESP_IP A_LCMSPY4 on record A_INDRESP_IP A_PMARINT on record A_INDRESP_IP A_LMARM on record A_MARRIAGE_IP A_LMARY4 on record A_MARRIAGE_IP A_LMCOH on record A_MARRIAGE_IP A_LMCBM on record A_MARRIAGE_IP A_LMCBY4 on record A_MARRIAGE_IP A_LMEND on record A_MARRIAGE_IP A_LMWWM on record A_MARRIAGE_IP A_LMWY4 on record A_MARRIAGE_IP A_LSPWWD on record A_MARRIAGE_IP A_LMDVM on record A_MARRIAGE_IP A_LMSPY4 on record A_MARRIAGE_IP A_LNCOH on record A_INDRESP_IP A_LCSBM on record A_COHAB_IP A_LCSBY4 on record A_COHAB_IP A_LCSEM on record A_COHAB_IP A_LCSEY4 on record A_COHAB_IP A_LADOPT on record A_INDRESP_IP A_LNADOPT on record A_INDRESP_IP A_LACBD which is not released A_LACBM on record A_ADOPT_IP A_LACBY4 on record A_ADOPT_IP A_LACSX on record A_ADOPT_IP A_LASCST on record A_ADOPT_IP A_LACYB4 on record A_ADOPT_IP A_LACLV on record A_ADOPT_IP A_LACNO on record A_ADOPT_IP A_LACAL on record A_ADOPT_IP A_LACYD4 on record A_ADOPT_IP A_LCHLV on record A_NATCHILD_IP A_LCHYD4 on record A_NATCHILD_IP A_LCHSX on record A_NATCHILD_IP A_LCHDOBY on record A_NATCHILD_IP A_LCHDOBM on record A_NATCHILD_IP A_LCHBD which is not released A_LCHAL on record A_NATCHILD_IP A_LCHNO on record A_NATCHILD_IP A_LCHCHK on record A_NATCHILD_IP

	A_BWTPX on record A_NATCHILD_IP A_BWTEL on record A_NATCHILD_IP A_BWTWK on record A_NATCHILD_IP A_BWT on record A_NATCHILD_IP A_BWTLB on record A_NATCHILD_IP A_BWTOZ on record A_NATCHILD_IP A_BWTK on record A_NATCHILD_IP A_BWTG5 on record A_NATCHILD_IP A_LCHMOR on record A_INDRESP_IP A_LCHMORN on record A_INDRESP_IP
--	--

In Wave 2, both groups were asked the “long-form” partnership and fertility history items. These items include:

B_LCMCOH on record B_INDRESP_IP	List continued
B_LCMCBM on record B_INDRESP_IP	B_LACNO on record B_ADOPT_IP
B_LCMCBY4 on record B_INDRESP_IP	B_LACAL on record B_ADOPT_IP
B_LCMSPM on record B_INDRESP_IP	B_LACYD4 on record B_ADOPT_IP
B_LCMSPY4 on record B_INDRESP_IP	B_LCHLV on record B_NATCHILD_IP
B_PMARINT on record B_INDRESP_IP	B_LCHYD4 on record B_NATCHILD_IP
B_LMARM on record B_MARRIAGE_IP	B_LCHSX on record B_NATCHILD_IP
B_LMARY4 on record B_MARRIAGE_IP	B_LCHDOBY on record B_NATCHILD_IP
B_LMCOH on record B_MARRIAGE_IP	B_LCHDOBM on record B_NATCHILD_IP
B_LMCBM on record B_MARRIAGE_IP	B_LCHBD which is not released
B_LMCBY4 on record B_MARRIAGE_IP	B_LCHAL on record B_NATCHILD_IP
B_LMEND on record B_MARRIAGE_IP	B_LCHNO on record B_NATCHILD_IP
B_LMWWM on record B_MARRIAGE_IP	B_LCHCHK on record B_NATCHILD_IP
B_LMWY4 on record B_MARRIAGE_IP	B_BWTPX on record B_NATCHILD_IP
B_LSPWWD on record B_MARRIAGE_IP	B_BWTEL on record B_NATCHILD_IP
B_LMDVM on record B_MARRIAGE_IP	B_BWTWK on record B_NATCHILD_IP
B_LMSPY4 on record B_MARRIAGE_IP	B_BWT on record B_NATCHILD_IP
B_LNCOH on record B_INDRESP_IP	B_BWTLB on record B_NATCHILD_IP
B_LCSBM on record B_COHAB_IP	B_BWTOZ on record B_NATCHILD_IP
B_LCSBY4 on record B_COHAB_IP	B_BWTK on record B_NATCHILD_IP
B_LCSEM on record B_COHAB_IP	B_BWTG5 on record B_NATCHILD_IP
B_LCSEY4 on record B_COHAB_IP	B_BRFED on record B_NATCHILD_IP
B_LADOPT on record B_INDRESP_IP	B_BRFEDEND on record B_NATCHILD_IP
B_LNADOPT on record B_INDRESP_IP	B_BRFEDEND2 on record B_NATCHILD_IP
B_LACBD which is not released	B_LCHMOR on record B_INDRESP_IP
B_LACBM on record B_ADOPT_IP	B_LCHMORN on record B_INDRESP_IP
B_LACBY4 on record B_ADOPT_IP	
B_LACSX on record B_ADOPT_IP	
B_LASCST on record B_ADOPT_IP	
B_LACYB4 on record B_ADOPT_IP	
B_LACLV on record B_ADOPT_IP	

8.2 Showcards vs. no showcards

There were experiments involving showcards in Waves 1-3. Face-to-face interviews often rely heavily on the use of showcards to enhance measurement. However, showcards are not necessarily available in a mixed mode approach to data collection particularly when the use of telephone interviewing is included in the mode mix.

Experimentation in Wave 1 examined whether there may be a primacy effect (selecting the first listed choice) of visual cues when a showcard is used or a recency effect (selecting the last listed choice) if the list is read. Also there was experimentation with different methods of obtaining information on unearned income sources without showcards. The question topics are about labour force status and unearned income.

For the labour force status experiment, there was random allocation of households within PSUs. All interviewed adults within households received the same experimental treatment. The groups are:

Group 1 – Question asked using a showcard

Group 2 – Question asked as a “read out” without a showcard

The variable for allocation to treatments is **a_group2** on the record **a_hhsamp_ip**. The variable **a_jbstat** on the record **a_indresp_ip** contains the substantive information.

The experiment also compares three methods of obtaining measures of unearned income sources. These sources include benefits paid by the government as well as money from loans, rents, private grants, and money transfers from private individuals. The British Household Panel Survey (BHPS) protocol for enumerating unearned income sources has relied heavily on showcards to remind respondents of all potential sources in order to obtain more accurate reporting. Such reliance on showcards may not be feasible in a mixed modes approach. The UK Labour Force Survey uses an enumeration protocol that does not rely on showcards but a complex array of filter questions. A three-way split-ballot (different wordings) experiment contrasts the BHPS approach to unearned income enumeration with alternative “no showcard” designs.

Within **PSUs**, households were randomly allocated to treatments and all adult respondents within households were treated the same:

Group 1 – Original British Household Panel Study version with showcard enumeration

Group 2 – Adapted Labour Force Survey approach without showcards, no filter

Group 3 – Adapted Labour Force Survey without showcards, with two initial filter questions

The variable in the data that controls allocation to treatments is **a_group3** on the record **a_hhsamp_ip**. A list of items about substantive benefit and payment sources for the three experimental groups follows:

Table: Variables measuring unearned income sources

Group 1 – showcard	Group 2 – no showcard, no filter question	Group 3 – no showcard, filter question
a_benpeng11 to a_benpeng196	a_btypeg21 to a_btypeg296	2 initial filter questions: a_benefit_g3
a_bendisg11 to a_bendisg196	a_benunempg21 to a_benunempg296	a_payment_g3a_btypeg31 to a_btypeg396
a_bensupg1 to a_bensupg196	a_bendisg21 to a_bendisg296	a_benunempg31 to a_benunempg396
a_benpayg11 to a_benpayg196	a_bendlag21 to a_bendlag296	a_bendisg31 to a_bendisg396
a_nfa_g1	a_benpeng21 to a_benpeng296	a_bendlag31 to a_bendlag396
a_nfb_g1	a_niserps2	a_benpeng31 to a_benpeng396
a_nfc_g1	a_benctcg2	a_niserps3
a_benalg11 to a_benalg196	a_benfamg21 to a_benfamg296	a_benctcg3
a_nfe_g11 to a_nfe_g13	a_bentaxg21 to a_bentaxg296	a_benfamg31 to a_benfamg396
a_nff_g1	a_benhoug21 to a_benhoug296	a_bentaxg31 to a_bentaxg396
a_nfg_g1	a_benstag21 to a_benstag296	a_benhoug31 to a_benhoug396
a_f2_g1		a_benstag31 to a_benstag396
a_niserps		

At Wave 2, the showcards study was expanded to incorporate the entire interview, though the experiment to measure unearned income sources was discontinued. CAPI respondents were randomly allocated into groups interviewed with showcards and those interviewed without showcards. The showcard experiment was independent of all other experiments, including the mixed modes experiment carried at Wave 2. Thus, all households were allocated to experimental treatments. If a telephone case was transferred to a face-to-face interviewer for follow-up, then they received their allocated showcard treatment.

Allocation was at the **PSU** level so that interviewers would be either with or without showcards for all of their interviews to avoid contamination through the inadvertent use of visuals within face-to-face interviews. The experimental allocation is as follows:

Group 1 – Showcards

Group 2 – No showcards

The controlling variable on the data is **b_ff_showcardw2** on the record **b_hhsamp_ip**. The experiment applies to all items in the questionnaire which are indicated as having a showcard for use with face-to-face interviewing.

At Wave 3, the showcard experiment was repeated. As in Wave 2, manipulations were done at the **PSU** level such that each interviewer either used or did not use showcards. Note, at Wave 2, the showcard experiment was confounded with the showcard treatments of other experiments. At Wave 3, these are not confounded. However, the use of showcards was still consistent across all other experiments except for the life satisfaction and job satisfaction experiment where a special showcard was required for all interviewers. There is further information about this implementation problem below. See the satisfaction experiment implementation notes (below) for further details. Showcards were used extensively throughout all portions of the questionnaire. The questionnaire indicates which questions use a showcard.

The controlling variables are **c_ff_showcardsw2** and **c_ff_showcardsw3** on record **c_hhsamp_ip**. note that **c_ff_showcardsw3** represents a rotation from the Wave 2 allocation. The rotation pattern is as follows:

Table: Rotation pattern for showcard experiment, waves 2 and 3

Values for c_ff_showcardsw2 and c_ff_showcardsw3 showing the allocation rotation		
Wave 2	Wave 3	
1	1	Showcards both waves
1	2	Showcards at Wave 2 & no showcards at Wave 3
2	1	No showcards at Wave 2 & showcards at Wave 3
2	2	No showcards at Wave 2 and Wave 3

Note that there was an error in the implementation of the IP3 showcard experiment, which meant that the treatments were not necessarily implemented as allocated. For details see the section detailing known data issues (see 3.6 Fieldwork paradata

Additional data collected during the interview process (paradata) are available. These consist of call records, timings data and other information collected by the interviewers during the interview.

Call records

Call record files have information on the number of calls made as well as the issue number, interviewer identifier (scrambled), time and date and the outcome of each call. This is available in the dataset **w_callrec_ip**.

Address response form

Information collected in the address response form (ARF) by interviewers while contacting each household and requesting household members to participate in the survey is available in **w_hhsamp_ip**. This includes data on the area surrounding the address, the type of accommodation and other information that the interviewer can observe for both responding and non-responding households. Reasons for refusal are also available. Interviewers also record some information about the quality of the interview and persons present during the interview process. This is available along with substantive data collected during adult individual interviews (including proxy interviews) in **w_indresp_ip**. From Wave 7 onwards the ARF was no longer used.

Timings data files

Timings data files (**w_ptimings** and **w_h timings**) include data on the time taken to complete each question and module in the individual and household questionnaires. In IP1, the start and end times are given for blocks of questions, where blocks are one or more question modules. The times are given in seconds. From IP2 onwards the times are given in seconds for individual questions. If the variables are asked in a loop or multi-choice format, the variable name is suffixed with the multi-choice item number or loop iteration count. In Waves 5 to 9 the timings data for interviews completed by web are per screen rather than per question, although most screens contain only a single question. Where there are multiple questions per screen this is documented in the pdf questionnaire. Waves 7 onwards are released in CSV format because the variable names are long strings that are truncated when imported into Stata. From Waves 7 to 16 the timings files are **w_hhgrid_timings**, **w_hhint_timings**, and **w_indint_timings**. From Waves 17 onwards the timing files are **w_hh_module_timings**, **w_hh_question_timings**, **w_iv_moudle_timings** and **w_iv_question_timings**.

The IP11 timings data included an error which has been corrected (see the example 3.8 Stata code for matching files3.8 Stata code for matching files).

Interviewer characteristics

The interviewer id **w_intnum** can be linked to the mainstage cross-wave file **xivdata** which contains interviewer characteristics. This file is available from the UK Data Service as a separate dataset (SN 8579), under Special Licence agreement.

Keystroke paradata

For IP11 there is an additional paradata file (**k_keystroke_paradata**), which contains information automatically recorded from CAPI and web respondents, while they answered the questions in the modules “HMRC consent” and “HMRC consent follow-up” (early and late versions). For each question in these modules the strings in the variables **k_keystrokes1** and **k_keystrokes2** record the question name, the response category selected, and the timestamp when the interviewer or respondent clicked ‘next’. The variable **k_keystrokes1** is truncated for some cases and the remainder of the string can be found in **k_keystrokes2**.

3.7 Known data issues).

8.3 Impact of question wording and context on measuring change

Measuring how people's social and economic circumstances change over time is a key purpose of household panel surveys. Levels of change are often overestimated in panel surveys. That is, responses to a question are often not consistent across interviews, even if the respondent's situation has not changed. Various methods have been proposed to address these issues, but little is known about the mechanisms giving rise to the observation of spurious change. For this reason, a series of split-ballot experiments was incorporated into several waves of the Innovation Panel. The ultimate aim is to understand the processes that lead a respondent, whose situation has not changed from one interview to the next, to give a different response to a survey question in different interviews.

For the experiments, allocation was of households within **PSUs**. The variable that controls allocation to versions of Experiments 1-3 in Wave 2 is **b_ff_changew2** on record **b_indresp_ip**:

Group 1	Version A of questions
Group 2	Version B of questions

Wave 2 experiment 1: Ambiguous wording. This contrasts questions where terms or definitions in the question were potentially ambiguous – contrasted with less ambiguous question wording. Respondents were randomly allocated to either version for the following questions:

Identifying disability status: items **b_health_a** and **b_health_b** on record **b_indresp_ip**

Identifying whether the respondent is in work: items **b_jbhas_a1** through **b_jbhas_b** on record **b_indresp_ip**

Identifying whether the respondent saves regularly: item **b_save_a1** through **b_save_b** on record **b_indresp_ip**

Wave 2 Experiment 2: Ambiguous instructions. This contrasts a question where the instructions were ambiguous in that response options were not mutually exclusive and there were no clear instructions about how to select the “main” category if more than one applied, with a “select all that apply” version of the same question whose instructions are less ambiguous:

Obtaining main labour market status: items **b_jbstat_a** through **b_jbstat_d2** on record **b_indresp_ip**

Wave 2 experiment 3: Implicit or explicit questions about dates. This examines the effect of the clarity of instructions for questions about the dates of events. Questions in which the request for a date was implicit (e.g. “Since when have you...?”) were contrasted with

questions where the request for the date was explicit (e.g. “In which month and year did you...?”):

Obtaining dates of moves: items **b_mvmnth_a** through **b_mvyr_b3**, and **b_plnowm_a** through **b_plnowy4_b3**, on record **b_indresp_ip**

obtaining dates of the onset of health conditions: items **b_hconda_a** through **b_hconda_b3y** on record **b_indresp_ip**

obtaining dates of joining private pension schemes: items **b_ppyrs_a** through **b_ppyrs_b3y** on record **b_indresp_ip**

Repetition of experiments 1-3 at wave 3

In **Wave 3** the experiments were repeated with the same wordings, treatments and allocations. The controlling variable is **c_ff_changew2** on record **c_hhsamp_ip**. Relevant variables are: **c_health_a**, **c_health_b**, **c_jbhas_a1** to **c_jbhas_b**, **c_save_a1**, **c_save_b**, **c_jbstat_a** to **c_jbstat_d2**, **c_mvmnth_a** to **c_mvyr_b3**, **c_plnowm_a** to **c_plnowy4_b3**, **c_hconda_a** to **c_hconda_b3y**, **c_ppyrs_a** to **c_ppyrs_b3y** on record **c_indresp_ip**

Wave 3 experiment 4: Context and frequency. This experiment tests whether changes in the context of a question across waves can affect measures of change. The experiment uses a question about frequency of events, where the preceding question is either about high or low frequency events. The controlling variable is the IP2 variable allocating treatments to the measures of change experiments: **c_ff_changew2** on record **c_hhsamp_ip**. Items about the frequency of political discussions follow either a high frequency item (**c_mdafrq** – media watching) or low frequency item (**c_vtefrq** frequency of voting). These items are on record **c_indresp_ip**. The political discussion items are **c_poldisc1a** through **c_pldisc6a** or **c_poldisc1b** through **c_pldisc6b**.

Repetition of experiments 1-4 at Wave 4

At Wave 4, Experiments 1 through 4 were repeated with identical wording and format. However, allocation to experimental group was rotated for some of the items in experiments 1, 3 and 4 (listed below), as compared to Wave 3. The controlling variable for the Wave 4 unchanged allocation is **d_ff_changew2** on record **d_hhsamp_ip**. The controlling variable for the Wave 4 rotated allocation is **d_ff_changew4** on record **d_hhsamp_ip**.

Table: Change in questions in wave 4 compared to waves 2 and 3

	Wave 4	
Wave 2 & Wave 3	Version A	Version B
Version A	Same questions	Rotated questions
Version B	Rotated questions	Same questions

Specifically, allocation at Wave 4 was unchanged (same questions) over prior waves for the following items: **d_mvmnth_a** to **d_mvyr_b3**, **d_jbstat_a** to **d_jbstat_d2**, **d_plnowm_a** to **d_plnowy4_b3**, **d_jbhas_a1** to **d_jbhas_b**, **d_save_a1**, **d_save_b** on record **d_indresp_ip**.

Allocation at Wave 4 was rotated as compared to prior waves for the following items: **d_mdafrq**, **d_vtefrq**, **d_health_a**, **d_health_b**, **d_hconda_a** to **d_hconda_b3y**, **d_ppyrs_a** to **d_ppyrs_b3y** on record **d_indresp_ip**.

Note, the refreshment sample was allocated to all controlling variables for Experiments 1-4 and included in all treatments.

8.4 Dependent interviewing wording

At Wave 3, an experiment was included to examine the potential effects of different wording of dependent interview questions on responses. A split-ballot experiment contrasted two versions of phrasing. In both cases respondents were first reminded of the answer they had given in the previous interview. With Version A they were asked “Is that still the case?”, with Version B they were asked “Has that changed?” The experiment was audio-recorded.

Allocation to treatment groups was of households within PSUs.

Group 1	Version A: “Is that still the case?”
Group 2	Version B: “Has that changed?”

The variable that controls allocation is **c_ff_changew2** on record **c_hhsamp_ip**. Relevant variables are **c_sf1_a**, **c_sf1_b**, **c_jbterm1_a**, **c_jbterm1_b**, **c_jbhrs_a**, **c_jbhrs_b**, **c_jshrs_a**, **c_jshrs_b** on record **c_indresp_ip**.

At Wave 4 the experiment was repeated with the same allocation to treatment groups. The refreshment sample was not asked dependent interviewing questions, since we did not have any wave 3 responses for this sample. Again, the experiment was audio-recorded.

The variable that controls allocation is **d_ff_changew2** on record **d_hhsamp_ip**. Relevant variables are **d_sf1_a**, **d_sf1_b**, **d_jbterm1_a**, **d_jbterm1_b**, **d_jbhrs_a**, **d_jbhrs_b**, **d_jshrs_a**, **d_jshrs_b** on record **d_indresp_ip**.

In Wave 5 and 7, the experiment contrasts the two question formats with similar questions not answered by “yes” or “no”. Furthermore, this work exploits the mixed mode experiment to study potential response order effects, and whether these differ in CAPI and web, by varying the order of response options in the new format. Respondents were assigned to one of four experimental question variants: (1) “Still the case? Yes/No”, (2) “Has this changed? Yes/No” (3) “Still the case or has this changed?” and (4) “Has this changed or is it still the case?”. The variable that controls allocation is **w_ff_diw5** on record **w_hhsamp_ip**. This variable is coded as:

1	Still the case
2	Has this changed
3	Balanced, still first
4	Balanced, changed first

The variables used for this experiment are – on record **hhresp**:

hsroomchk_a, **hsroomchk_b**, **hsroomchk_c**, **hsroomchk_d**, **hsowndchk_a**, **hsowndchk_b**, **hsowndchk_c**, **hsowndchk_d**, **xpmg_a**, **xpmg_b**, **xpmg_c**, **xpmg_d**, **rentchk_a**, **rentchk_b**, **rentchk_c**, **rentchk_d**

On record **INDRESP**:

lkmove_a, lkmove_b, lkmove_c, lkmove_d, edtype_a, edtype_b, edtype_c, edtype_d, jbterm1_a, jbterm1_b, jbterm1_c, jbterm1_d, jbsic07_a, jbsic07_b, jbsic07_c, jbsic07_d, jbsoc00_a, jbsoc00_b, jbsoc00_c, jbsoc00_d, jbsemp_a, jbsemp_b, jbsemp_c, jbsemp_d, jbsizechk_a, jbsizechk_b, jbsizechk_c, jbsizechk_d, jbherschck_a, jbherschck_b, jbherschck_c, jbherschck_d, paygl_a, paygl_b, paygl_c, paygl_d, paynl_a, paynl_b, paynl_c, paynl_d, paytypchk_a, paytypchk_b, paytypchk_c, paytypchk_d, wktravchk_a, wktravchk_b, wktravchk_c, wktravchk_d, jshrschk_a, jshrschk_b, jshrschk_c, jshrschk_d, jspartchk_a, jspartchk_b, jspartchk_c, jspartchk_d, jstravchk_a, jstravchk_b, jstravchk_c, jstravchk_d

Note that the Wave 5 implementation of this experiment was corrupted. (Further details can be found in the Section Known Data Issues.) The experiment was therefore repeated in Wave 7.

8.5 Branched vs. unbranched rating scales for measuring attitudes

The most common method for eliciting attitudes and beliefs in surveys is to employ rating scales where respondents are asked to choose the alternative that best describes their belief in, attitude towards or agreement with a statement. Such response scales typically contain between three and seven alternatives and can be bipolar (strongly agree, agree, neither agree nor disagree, disagree, strongly disagree) or unipolar (always, often, sometimes, occasionally, never).

Attitude items that use a single bipolar rating scale are ubiquitous in surveys. Evidence suggests that a two-step branched method may be more advantageous in terms of response processes and statistical properties. In the branched or unfolding bipolar format, respondents are first asked about the overall direction of their belief or attitude (e.g. satisfied or not satisfied). Then they are asked about degree (very satisfied, somewhat satisfied, slightly satisfied). A composite score can then be computed for the pair of items.

This experiment was first conducted in Wave 3. Households within PSUs were assigned to one of two experimental treatments, with special attention to being independent of the any showcard experiments.

The controlling variable is for Wave 3 is **c_ff_branchingw3** on record **c_hhsamp_ip**. The two groups are:

Group 1 – Branched

Group 2 – Unbranched

The affected questionnaire items are **c_nbrcoh1_a** to **c_nbrcoh4_d2**, and **c_poleff1_a** to **c_poleff_d2** on record **c_indresp_ip**. The questions are about neighbourhood cohesion and political efficacy.

At Waves 4 and 5, the experiment was repeated, with the same allocation to treatments as at Wave 3. The controlling variable is **w_ff_branchingw3** on record **w_hhsamp_ip**, with relevant questionnaire items **w_nbrcoh1_a** to **w_nbrcoh4_d2**, and **w_poleff1_a** to **w_poleff_d2** on record **w_indresp_ip** where “w” indicates “D” and “E” for Waves 4 and 5.

8.6 Smiley faces vs. text-based scales in child self-completion

At Wave 5 and Wave 6, the youth questionnaire examines how to adapt questions on satisfaction for children, focusing on the use of pictorial evaluations of feelings. A split-ballot design was incorporated in the Waves 5 and 6 youth questionnaire where half of the child self-completion uses smiley faces for the questions on satisfaction in different domains, and the other half use a scale with a textual description but no smiley faces. Within PSUs, households were randomly allocated to experimental treatment all youth aged 10-15 within households received the same treatment:

Group 1	Smiley faces
Group 2	Text descriptions.

The experiment is controlled in Wave 5 by the variable **e_ff_smilesw5** on record **e_hhsamp_ip** and in wave 6 by the variable **f_ff_smilesw5** on record **f_hhsamp_ip**. The affected variables in Wave 5 are **e_yphsw**, **e_yphap**, **e_yphfm**, **e_yphfr**, **e_yphsc**, and **e_yphlf** on record **e_youth_ip**. Comparable affected variables in Wave 6 are **f_yphsw**, **f_yphap**, **f_yphfm**, **f_yphfr**, and **f_yphsc**.

8.7 Quality of recall data with web vs. face-to-face

This experiment aimed (a) to investigate differences in data quality arising from switching Innovation Panel members from face-to-face to web mode of survey administration, (b) to test methods for the mitigation of a hypothesized decline in quality associated with moving to web data collection and (c) to contribute to a general understanding of how web data collection methods can affect data quality. The experiment asks respondents to recall facts gathered contemporaneously at earlier waves and uses earlier wave data as a validation check. Half of respondents completing the survey on the web received a commitment pledge as an experimental treatment to encourage more accurate reporting of historical information. Randomisation was of households across the complete sample, i.e., regardless of PSU, into one of two experimental treatments.

The controlling variable is **F_FF_ITEMW6** on record **F_HHSAMP_IP**. It is coded as:

- 1 No commitment pledge
- 2 Commitment pledge

Affected variables in the questionnaire are **f_empv1** **f_empv2** **f_sf1recall** **f_sf5recall** **f_sf6crecall** **f_healthrecall** **f_hlwtrecall** **f_hlwtsrecall** **f_hlwtprecall** **f_hlwtkrecall** **f_recallease** **f_recalleffort** **f_webrecall** **f_webinterrupt** on record **f_indresp_ip**.

8.8 Methods of reducing item non-response in web surveys

Elevated levels of item non-response in web surveys are a concern. To see whether the incidence of item non-response can be reduced, Wave 6 contained an experiment on 6 questions which showed relatively high levels of item non-response in past waves. Wave 6 web respondents were assigned to one of two experimental treatments, or to a control

group. The control group received the standard non-response protocol. The routine approach for accessing non-response codes is to not present the codes on the screen presenting the question to respondents but then if the respondent attempts to skip the question without answering codes for “Don’t know” and “Rather not answer” appear in blue with a message stating “You forgot to answer this question”. Respondents must select an answer at this point before moving to the next screen.

The first experimental treatment received a re-active prompt for item non-response which altered the non-response message. The altered message read, “If possible, please provide an answer to this question as this is one of the key questions in this study. Please be assured that the information you give us will be treated confidentially.” The same set of response categories appeared, but without specific item non-response codes. If the respondent attempted to skip the question, the item non-response codes appeared, and respondents were forced to pick an answer.

The second experimental treatment replicated the standard non-response protocol. However, respondents who provided a “Don’t know” or “Refused” code at the tested items received a set of follow-up questions at the end of the questionnaire. These questions enquired about any difficulties in answering the questionnaire, confirming who answered the questionnaire, and re-presented the questions to which the respondent answered “Don’t know” or “Refused”. At this re-presentation of items, respondents were thanked for their participation, informed that they had not answered certain key questions, told the importance of these questions, told that these are the last questions and asked to complete them before finishing.

Households within **PSUs** were randomly allocated to the two treatments and control in equal proportion, one-third in each group. However, the experiment only applied to respondents interviewed via web and this mode control was implemented in question filtering within the survey instrument.

The controlling variable is item **f_ff_itemnonw6** on record **f_hhsamp_ip**. The groups are:

- 1 Standard procedure
- 2 Prompt for item non-response
- 3 Follow-up questions at end of questionnaire

Affected questionnaire items are **f_plbornuk**, **f_mstatsam**, **f_paygl**, **f_basrate_cawi_1** to **f_basrate_cawi_3**, **f_jsprf**, **f_fiyrdia**, **f_paygl**, **f_webexp1_code** to **f_webexp3_code**, **f_webwho1** to **f_webwho16**, **f_webnrnum**, **f_paygl_fu**, **f_payglwc_fu**, **f_fiyrdia_fu**, **f_jsprf_fu**, **f_plbornuk_fu**, **f_basrate_fu**, **f_mstatsam_fu**, and **f_webincwhy** on record **f_indresp_ip**.

The variables **F_variable_CAWI_1**, **F_variable_CAWI_2**, **F_variable_CAWI_3** and **F_variable_FU** correspond to the serial order of attempts asking the question, such that **F_variable_CAWI_3** is the final time for the initial query, and the only attempt with opt-out options available for each question. These variables correspond to the responses in the three conditions in the following way:

f_variable_cawi_1= The first time the question was asked in the motivational statement condition.

f_variable_cawi_2= The second time the question was asked in the motivational statement condition and the first time the question was asked in the control and follow-up conditions.

f_variable_cawi_3 =The third time the question was asked in the motivational statement condition and the second time the question was asked in the control and follow-up conditions. The only time the opt-out options are available.

The only way to identify this is to control using **f_ff_itemnonw6** on file **f_hhsamp**.

If the respondent does not answer or selects an opt-out option, this response is coded in an entirely different variable (i.e. not **f_variable_cawi_3**). These are all prefixed with DKRF, but have numbers that follow that identify it for specific variables. These are as follows

Basrate=	DKRF34
FiYrDIA=	DKRF41
JsPrf=	DKRF8
Mstatsam=	DKRFY1
Paygl=	DKRF31
Paygwc=	DKRF31
Plbornuk=	DKRF13

Note that these DKRF variables are coded as the positive values 1 and 2 instead of -1 and -2 for “Don’t Know” and “Refused”, respectively.

Respondents in the follow-up condition who have a value in DKRF variables are asked the question in the follow-up section, coded in **f_variable_fu**. However, unlike **f_variable_cawi_3**, missing values can be recorded in this variable.

Using the experimental variables the main variable is coded. Important to note is that the default for all variables is -2 (“Refused”) and only changed if some other answer is given. The backcoding to the main variable is done following this logic:

```
f_variable = -2
IF DKRF[N]=1 THEN f_variable=-1
ELSE IF f_variable_CAWI_3>0 THEN f_variable= f_variable_CAWI_3
ELSE IF f_variable_CAWI_2>0 THEN f_variable= f_variable_CAWI_2
ELSE IF f_variable_CAWI_1>0 THEN f_variable= f_variable_CAWI_1
```

f_variable_fu is not backcoded to the main variable, so the main variable only constitutes the first set of queries.

Error in `f_fiyrdia` and `f_fiyrdia_fu`

An error occurred in coding of `f_fiyrdia` and the implementation of `f_fiyrdia_fu`. This occurred because of the above logic. In the above logic the experimental variables are backcoded to the main variable only if the response is greater than 0.

However, the majority of respondents has and gave a zero value to this question. Hence these answers were not backcoded and responses remained incorrectly coded as -2, the default.

Due to this, everyone in the follow-up condition who gave an answer of zero were also coded as a refusal, and asked `f_fiyrdia_fu` even though they provided a substantive answer.

These variables can be corrected by using the provided experimental variables, correcting the above logic to greater or equal to 0.

8.9 Separating systematic measurement error components using MTMM in longitudinal studies

This experiment extended multi-trait multi-method (MTMM) designs beyond allowing for separation of random and common method variance. It identified three additional systematic variance components: acquiescence, social desirability, and extreme response. The research targeted opinions towards immigration with a repetition of six questions early in the questionnaire and late in the questionnaire, though of varying forms.

This experiment was carried in Waves 7, 8 and 9, with a fresh randomisation to treatment in each wave. Households within **PSUs** were randomly allocated to one of 56 different treatment groups which varied question wording, response options and the ordering in the questionnaire of these components.

Template for the questionnaire

Now some questions about people from other countries coming to live in the UK.

<First set of 6 questions>

<Other interview questions>

To help us improve our questions in the future, here are some final questions on a range of different topics which are similar to previous ones. Please don't try to remember what you answered before but treat them as if they were completely new questions.

Now some questions about people from other countries coming to live in the UK.

<second set of 6 questions>

For each of the 6 items there are "positive" and "negative" formulations. "Positive" means that with a disagree-agree scale, the socially desirable direction will be towards the higher end of the scale.

Table: Traits and social desirability direction for MTMM experiment

Trait number	SD direction	Item formulation
T1	Positive	The UK should allow more people of the same race or ethnic group as most British people to come and live here
T2	Positive	UK should allow more people of a different race or ethnic group from most British people to come and live here
T3	Positive	UK should allow more people from the poorer countries outside Europe to come and live here
T4	Positive	It is generally good for UK's economy that people come to live here from other countries
T5	Positive	UK's cultural life is generally enriched by people coming to live here from other countries
T6	Positive	UK is made a better place to live by people coming to live here from other countries
T1	Negative	The UK should allow fewer people of the same race or ethnic group as most British people to come and live here
T2	Negative	UK should allow fewer people of a different race or ethnic group from most British people to come and live here
T3	Negative	UK should allow fewer people from the poorer countries outside Europe to come and live here
T4	Negative	It is generally bad for UK's economy that people come to live here from other countries
T5	Negative	UK's cultural life is generally undermined by people coming to live here from other countries
T6	Negative	UK is made a worse place to live by people coming to live here from other countries

There are 8 different wordings of each item, corresponding to combinations of three factors: the higher- or lower-end being the socially desirable direction, the number of scale points, and whether agree-disagree or disagree-agree questions are used. These lead to 8 wordings W1-W8; an example formulation for trait one is given in the last column.

Table: Item wordings for MTMM experiment

Wording number	Social desirability	Number of scale points	Agree or Disagree	Required direction	Item formulation (using trait 1 as an example)
W1	Higher	2	AD	Negative	The UK should allow fewer people of the same race or ethnic group as most British people to come and live here
W2	Lower	2	AD	Positive	The UK should allow more people of the same race or ethnic group as most British people to come and live here
W3	Higher	11	AD	Negative	The UK should allow fewer people of the same race or ethnic group as most British people to come and live here
W4	Lower	11	AD	Positive	The UK should allow more people of the same race or ethnic group as most British people to come and live here
W5	Higher	2	DA	Positive	The UK should allow more people of the same race or ethnic group as most British people to come and live here
W6	Lower	2	DA	Negative	The UK should allow fewer people of the same race or ethnic group as most British people to come and live here
W7	Higher	11	DA	Positive	The UK should allow more people of the same race or ethnic group as most British people to come and live here
W8	Lower	11	DA	Negative	The UK should allow fewer people of the same race or ethnic group as most British people to come and live here

Instead of presenting each respondent with all 8 different wordings of the same items, any one respondent was asked only 2 different wordings of the items, one at the beginning and the second at the end of the questionnaire. There were $(8,2) = 28$ different questionnaire versions corresponding to the combinations of wordings. Moreover the ordering of the wordings was randomised, leading to 56 different conditions to be randomised in total.

Controlling variable on record **hhsamp**:

ff_mtmmw7 – Wave 7 treatment allocation: Takes the values 1-56 to allocate unique combinations of wording, response option and questionnaire ordering. Coding is as in the table below.

ff_mtmmw8 – Wave 8 treatment allocation, coding as for Wave 7.

ff_mtmw9 – Wave 9 treatment allocation, coding as for Wave 7

Table: Control variable values for MTMM experiment

	Randomised ordering 1 2			Randomised ordering 2 1	
<i>ff_mtm</i> <i>mw7</i> =	Question set early in q're	Question set late in q're	<i>ff_mtm</i> <i>mw7</i> =	Question set early in q're	Question set late in q're
1	W1	W2	29	W2	W1
2	W1	W3	30	W3	W1
3	W1	W4	31	W4	W1
4	W1	W5	32	W5	W1
5	W1	W6	33	W6	W1
6	W1	W7	34	W7	W1
7	W1	W8	35	W8	W1
8	W2	W3	36	W3	W2
9	W2	W4	37	W4	W2
10	W2	W5	38	W5	W2
11	W2	W6	39	W6	W2
12	W2	W7	40	W7	W2
13	W2	W8	41	W8	W2
14	W3	W4	42	W4	W3
15	W3	W5	43	W5	W3
16	W3	W6	44	W6	W3
17	W3	W7	45	W7	W3
18	W3	W8	46	W8	W3
19	W4	W5	47	W5	W4
20	W4	W6	48	W6	W4
21	W4	W7	49	W7	W4
22	W4	W8	50	W8	W4
23	W5	W6	51	W6	W5
24	W5	W7	52	W7	W5
25	W5	W8	53	W8	W5
26	W6	W7	54	W7	W6
27	W6	W8	55	W8	W6
28	W7	W8	56	W8	W7

Variables affected on record **indresp** in Waves 7,8 and 9:

mtmm1q1e, mtmm1q2e, mtmm1q3e, mtmm1q4e, mtmm1q5e, mtmm1q6e, mtmm2q1e, mtmm2q2e, mtmm2q3e, mtmm2q4e, mtmm2q5e, mtmm2q6e, mtmm3q1e, mtmm3q2e, mtmm3q3e, mtmm3q4e, mtmm3q5e, mtmm3q6e, mtmm4q1e, mtmm4q2e, mtmm4q3e, mtmm4q4e, mtmm4q5e, mtmm4q6e, mtmm5q1e, mtmm5q2e, mtmm5q3e, mtmm5q4e, mtmm5q5e, mtmm5q6e, mtmm6q1e, mtmm6q2e, mtmm6q3e, mtmm6q4e, mtmm6q5e, mtmm6q6e, mtmm7q1e, mtmm7q2e, mtmm7q3e, mtmm7q4e, mtmm7q5e, mtmm7q6e, mtmm8q1e, mtmm8q2e, mtmm8q3e, mtmm8q4e, mtmm8q5e, mtmm8q6e, mtmm1q1l, mtmm1q2l, mtmm1q3l, mtmm1q4l, mtmm1q5l, mtmm1q6l, mtmm2q1l, mtmm2q2l,

mtmm2q3l, mtmm2q4l, mtmm2q5l, mtmm2q6l, mtmm3q1l, mtmm3q2l, mtmm3q3l, mtmm3q4l, mtmm3q5l, mtmm3q6l, mtmm4q1l, mtmm4q2l, mtmm4q3l, mtmm4q4l, mtmm4q5l, mtmm4q6l, mtmm5q1l, mtmm5q2l, mtmm5q3l, mtmm5q4l, mtmm5q5l, mtmm5q6l, mtmm6q1l, , mtmm6q2l, mtmm6q3l, mtmm6q4l, mtmm6q5l, mtmm6q6l, mtmm7q1l, mtmm7q2l, mtmm7q3l, mtmm7q4l, mtmm7q5l, mtmm7q6l, mtmm8q1l, mtmm8q2l, mtmm8q3l, mtmm8q4l, mtmm8q5l, mtmm8q6l.

Variables affected on record **indresp** in Waves 7 and 8:

mtmmehh, mtmmemm, mtmmess, mtmmlhh, mtmmlmm, mtmmlss.

Variables affected on record **indresp** in Wave 8:

mtmml_temp, mtmme_temp.

8.10 Replicating classic response order experiments across countries

This experiment contained a number of questions originally described and analyzed by Schuman and Presser (1981), as well as some new variants. The purpose of these questions was to examine whether the experiments can be replicated decades later and across countries. The differences across countries may include varying cultural and conversational norms, which may lead to different response distributions. Respondents were asked a total of 20 questions in this experiment. Each question has between 2 – 4 versions, controlled by the variables below. Each question set has its own controlling variable which is independent of all other experiments including each other.

Each individual was allocated to unique treatment groups for each set of response order question experiments with controlling variables listed below under “**Controlling variables**”. Unlike other experiments — where allocation to treatment normally occurs in advance of fieldwork — allocation to experimental treatment occurred post enumeration via “SPSS Data Collection”, the software used by TNS BMRB to do computer assisted interviewing. All controlling variables were returned on the data as ordinary survey variables.

Table: Variables used in response order experiment

Controlling variables (on record indresp):	Variables affected in the IP7 questionnaire: all variables in the “Response Option” module on record indresp :
oilsupw7 – Coded 1-4 : Version A through D	oilsup_a, oilsup_b, oilsup_c, oilsup_d
oilcompw7 – Coded 1-4 : Version A through D	oilcomp_a, oilcomp_b, oilcomp_c, oilcomp_d
adqhousw7 – Coded 1-4 : Version A through D	adqhous_a, adqhous_b, adqhous_c, adqhous_d
insocondw7 – Coded 1-4 : Version A through D	insocond_a, insocond_b, insocond_c, insocond_d
jobsw7 – Coded 1-4 : Version A through D	jobs_a, jobs_b, jobs_c, jobs_d

Controlling variables (on record indresp):	Variables affected in the IP7 questionnaire: all variables in the “Response Option” module on record indresp:
womenpolw7 – Coded 1-4 : Version A through D	womenpol_a, womenpol_b, womenpol_c, womenpol_d
saygovw7 – Coded 1-2 : Version A and B	govcomp_a, govcomp_b
churchw7 – Coded 1-2 : Version A and B	churches_a, churches_b
avtempw7 – Coded 1-2 : Version A and B	avtemp_a, avtemp_b
noopinw7 – Coded 1-4 : Version A through D	courts_a, courts_b, courts_c, courts_d, ldsmt_a, ldsmt_b, ldsmt_c, ldsmt_d, ldcrk_d_a, ldcrk_d_b, ldcrk_d_c, ldcrk_d_d
fuelshw7 – Coded 1-4 : Version A through D	fuelsh_a, fuelsh_b, fuelsh_c, fuelsh_d
unionsw7 – Coded 1-4 : Version A through D	unions_a, unions_b, unions_c, unions_d
abrtw7 – Coded 1-2 : Version A and B	abrtgen_a, abrtgen_b, abrtdef_a, abrtdef_b
spendw7 – Coded 1-2 : Version A and B	unispnd_a, unispnd_b, busspend_a, busspend_b
inequalw7 – Coded 1-2 : Version A and B	inequal_a, inequal_b

8.11 Impact of response scale direction on responses

In Wave 7, this set of experiments reflects two different, though related, research questions. The first, most general, research question is whether and how the direction of a response scale affects survey responses. If a scale runs from positive to negative (e.g., “strongly agree” to “strongly disagree” or “excellent” to “poor”) versus from negative to positive (e.g., “strongly disagree” to “strongly agree” or “poor” to “excellent”), should we expect different responses from survey respondents? The empirical evidence is mixed. A second research question builds on this first and is more specific about the theoretical underpinnings. That is, on social surveys researchers routinely ask questions on sensitive topics: e.g. health and wellbeing, drug and alcohol consumption, sleeping problems, to name a few. Survey questions of this kind are prone to social desirability bias. In a recent study, in which respondents’ eye-movements were traced and respondents were found to not read response options presented at the bottom of the scale (Kaminska and Foulsham 2016). The second research question is therefore: can we improve survey answers to sensitive questions if the scale starts with socially unpleasant response options which might otherwise appear at the bottom of response option lists?

Households within IP7 **PSUs** were allocated to treatment such that all individuals in the household received the same version of questions, as did all split-off households.

In Wave 8, both motivational message and response option reversal experiments are repeated. However for the response option experiment in Wave 8, half of households will be randomly re-allocated the opposite treatment.

Controlling variables on record **hhsamp**:

Table: Control variable for scale ordering experiment

ff_reversew7	ff_reversew8
1 = Version A	1 = Version A
	2 = Version B
2 = Version B	1 = Version A
	2 = Version B

The variables affected in the Wave 7 and Wave 8 questionnaires are: job satisfaction (**jbsat_a7**, **jbsat_b7**), general health (**scsf1_a**, **scsf1_b**), all items in the self-completion GHQ general health module (**scghqa_a**, **scghqa_b**, ..., **scghql_a**, **scghql_b**), and all items in the self-completion satisfaction module (**sclfsat1_a**, **sclfsat1_b**, ..., **sclfsato_a**, **sclfsato_b**).

This experiment was combined with a motivational message read or displayed to a random half of respondents. The motivational message read

“In order for your answers to be most helpful to us, it is important that you try to be as thoughtful as you can. Since we need complete and accurate information from this research, we hope you will think hard to provide the information we need.”

The variable controlling who received the motivational message is **ff_motivw7** on record **hhsamp**:

Group 1 – Receives message

Group 2 – Does not receive message

8.12 Enhancing respondent engagement with the survey through tailored interesting questions

This experiment focuses on whether including extra questions that are of interest to the respondent improves their perception of the current survey and participation in the following wave. Target outcomes include response at subsequent waves as well as interviewer observations related to respondent interest in the questionnaire and likelihood to participate at a subsequent wave.

In all three samples (original, IP4 and IP7 refreshment samples), households within PSUs were allocated to treatment such that all individuals within households received the same experimental allocation. Households were divided randomly into two treatment groups – a group receiving content comprised of up to 3 tailored or ‘interesting’ questions and a control group. For the original IP sample (**ff_hhorig** = 7), the treatment questions were tailored to respondents' interests as reported at IP2 in questions concerning leisure, culture and sports participation. For the two refreshment samples (**ff_hhorig** = 10, 11), the treatment questions concerned TV watching. The control group received no additional questions. The controlling variables were **ff_contintw7** and **ff_hhorig**, with the specific version of questions for those receiving tailored questions being controlled by **ff_intqw7**:
Controlling variables:

ff_contintw7 on record hhsamp

- 1 Tailored/interesting questions
- 2 Control

Table: Variables used in interesting questions experiment

<i>ff_intqw7 on record indsamp</i>		Interesting questions asked (in indresp)
100	Group 0	teampart, teamcompete, teamcomplvl, teamread, teamfut,
101	Group 1	sportpart, sportalone, sporthealth, sportfut,
102	Group 2	actpart, actaud, actperf, actaudevr, actfut
103	Group 3	artcomp, arttype, arttime
104	Group 4	actpart, actaud, actperf, actaudevr, actfut
105	Group 5	evntatt, evnttype, vntenjy, evntfut
106	Group 6	
107	Group 7	
108	Group 8	
109	Group 9	
.	Missing	tvprogregr, tvprog, tvmostenjy

8.13 Grid design in mobile surveys

Mobile phones are increasingly being used as devices to respond to web survey invitations. There are a number of commonly held beliefs of best design for mobile phone surveys, but much less empirical evidence. One frequent example where a series of questions with the same response options are aligned in grid format, with the common response options arrayed at the top of the grid with the question stems aligned to the left of the grid. Radio buttons generally are set in the grid coordinating one response option with one question stem. The held belief is that this format, common in PC web surveys, are more problematic for mobile devices based on understood principles of visual design and usability (e.g. selecting a radio button). However, little research has explored how this design compares to other options, and if differences do arise, what design produces the best data quality.

The objective of this experiment is to compare the standard static grid design, with an alternative dynamic grid design. Dynamic grids present the response options as a constant display aligned below a question of the set normally in a grid. When a response is chosen, the response options remain unmoving, but the question dynamically changes to the next question of the set and so on. The dynamic grid is intended to better control for the possible visibility and usability issues. This experiment is applied to two self-completion modules already carried in the IP that use several grids, SF-12 and Mobile Device Use. As these are both self-completion, respondents in both the web and face-to-face conditions were included.

Respondents were assigned to the experiment individually within the course of the questionnaire. There was a random assignment of grid-type at the start of each module included in the experiment, independent of the other assignment. The controlling variables for this assignment are:

gridsf12w10 in record **j_indresp_ip**: Grid type assignment for the SF-12 module

- 1 Static grid
- 2 Dynamic grid

gridmobdevw10 in record **j_indresp_ip**: Grid type assignment for the Mobile Device Use module

- 1 Static grid
- 2 Dynamic grid

Variables used for this experiment are on record **j_indresp_ip**:

scsf1, scsf2a, scsf2b, scsf3a, scsf3b, scsf4a, scsf4b, scsf6a, scsf6b, scsf6c, scwrrymod1, scwrrymod2, scwrrymod3, scwrrymod4, scwrrymod5, scwrrymod6

8.14 Don't know/prefer not to answer response presentation

This experiment explores the impact of the presentation of Don't Know/Prefer Not to Answer response options in electronic self-completion questions. Currently in Understanding Society, the initial presentation of a question in a self-completion mode presents the question with neither the Don't Know/Prefer Not to Answer options showing. When a respondent tries to go to the next question without providing a response to the initial asking of the question, the question is represented with the Don't Know/Prefer Not to Answer options available in a different colour (blue) with a prompt to select a response. There are no initial instructions in the survey as how questions will be asked in relation to giving Don't Know/Prefer Not to Answer responses.

To explore the impact of presentation variations, three conditions are used in face-to-face CASI modules or the analogous web version. First is a control condition, presenting Don't Know/Prefer Not to Answer options as is currently done in Understanding Society. Second is as is the control condition, but with a specific prompt on the initial screen to inform respondents how they can view additional options. Third is inclusion of Don't Know/Prefer Not to Answer options as part of main response lists.

The experiment covers several modules contained in the face-to-face CASI module toward the end of the survey: the SF-12, GHQ, Young Adults, Child Development, and Non-resident relations. Additionally, there were two additional attitude questions on issues of nuclear energy and the United Nations that were also be included in the response presentation experiment. For these new additional questions, respondents were also asked follow-up questions about their self-rated level of knowledge on these topics. For those who provided a Don't Know response but indicated little knowledge on the topic, a follow-up question on reasoning was asked. Similarly, those saying they have a lot of knowledge but provided a Don't Know response on the attitude question were also asked a follow-up question about reasoning.

Individual respondents were allocated equally to these three conditions (i.e. 1/3 chance for each condition). The controlling variable for this experiment is as follows:

dkcond on record **k_indresp_ip**:

- 1 {Control, standard procedure}
- 2 {Instructions on DK procedure given first}
- 3 {DK/REF response options offered in the initial presentation of question}

The following variables used for this experiment are on record **k_indresp_ip**:
nuclen, untrust, knownuclen, knowun, whynucans1_code, whynucans2_code, whyunans1_code, whyunans2_code, whorupro, whoru edu, whorurac, whorupol, whorufam, whorusex, whoru age, scsf1, scsf2a, scsf2b, scsf3a, scsf3b, scsf4a, scsf4b, scsf5, scsf6a, scsf6b, scsf6c, scsf7, scghqa, scghqb, scghqc, scghqd, scghqe, scghqf, scghqg, scghqh, scghqi, scghqj, scghqk, scghql, nccr1, nccrm, nccry4, nccr3, nccr4, nccr5, nccr6, nccr8, , nccr9, nccr11, nccr12

The following variables used for this experiment are on record **K_chdev_ip**:
cdvla, cdcond, cddis, cdlmt, cd3pera, cd3perb, cd3perc, cd3perd, cd3pere, cd3perf, cd3perg, cdvll, cdvllb, cdvllc, cdvll d, cdvll e, cdvllf, cdvllg, cdvllh, cdvlli, cdvllj, cdvllk, chsdqc, cdvllm, cdvlln, cdvll o, cdvllp, cdvllq, cdvllr, cdvlls, cdvllt, chsdqa, chsdqb, chsdqn, chsdqd, chsdqe, chsdqf, chsdqg, chsdqh, chsdqi, chsdqj, chsdqk, chsdql, chsdqm, chsdqy, chsdqo, chsdqp, chsdqq, chsdqr, chsdqs, chsdqt, chsdqu, chsdqv, chsdqw, chsdqx, cdtvvidhrs, cdcread, cdoread, cdwread1, cdwread2, cdwread3, cdwread4, cdwread5, cderead, mealsreg, bedreg, cdtvvidhrw, cdcomp, cdpchrs, cdconsol, cdconstm, cdphsc, cdphscy1, cdphscy2, cdphscy3, cdphscy4, cdphscy5, cdphscy6, cdphscy7, chrisk, chpat, chdelay

8.15 Collecting mobile phone numbers

At the end of the interview at each wave, contact details for the respondent are collected, making sure that the ones on file are up to date and correct. In past waves, a list of several contact details is asked about simultaneously: home landline, mobile number, work number, and email address. Given the potential importance of contacting respondents on their mobile phone in upcoming waves, these experiments explored alternative designs to better ensure mobile numbers are collected.

At Wave 12, in the control group, participants were asked the questions as these have been in past IP waves. The alternative condition asked respondents specifically about mobile phone numbers separately from all the other contact details normally asked about. After asking for mobile phone number, these respondents were asked for all the remaining contact details normally asked about (home landline, work number, and email address) in a grouped manner similar to the normal design.

Allocation to this experiment was done at the household level; all respondents within a household received the same set of contact detail questions. Households were allocated randomly and equally to conditions (50% each). The controlling variable for this experiment is **l_ff_mobexp_w12** on record **l_hhsamp_ip**:

- 1 Control, usual contact detail questions
- 2 Mobile phone focused contact detail questions

The variables affected by this experiment are in the **record l_indresp_ip**:
l_rphmob_code, l_rhland_code

At Wave 15, an experiment tested amendments to questions in the Contact Details module, with the aim of increasing the reporting and updating of mobile phone numbers.

Households were randomly allocated to one of two treatment groups: one group received the 'mobile first' version of the questions that was asked in IP12/IP13; the other group were presented with amended wording aiming to justify/increase compliance with the request for mobile numbers. This affected the first questions in the 'prioritise mobile number' version of the Contact Details module used in IP12 (for half the sample) and IP13 (for everyone).

The wording asking for mobile contact details was experimentally allocated equally among households.

The controlling variables are on record **o_hhsamp_ip**:

ff_mobnow15 (1/2 each, allocation stratified by **sampleorig** **ff_gridmodew15**
ff_incentw15 **ff_consentcasiw15**)

1 = Mobile first version

2 = Increase compliance version

Modules affected:

contactdetails_ip15

9. Questionnaire design experiments: specific topics

9.1 Measures of consumption and expenditure

At Wave 1, an experiment was conducted to help develop efficient data collection instruments concerning household consumption. There is methodological concern about whether respondents consider all possible categories of expenditure to arrive at reports of household consumption. The way respondents formulate answers could also be affected by categorical cues in the question text. This three-way split-ballot (different questions) experiment in the household questionnaire addressed these concerns.

There was random allocation of households within PSUs to one of three treatments:

- | | |
|----------------|---|
| Group 1 | Question about overall expenditure without cues |
| Group 2 | Question about overall expenditure with detailed categorical cues |
| Group 3 | Separate questions about amounts of expenditure broken out into reporting on each category rather than an overall figure. |

The variable in the data that controls allocation to treatments is **a_group3** on the record **a_hhsamp_ip**.

The substantive information from each experimental treatment can be found in the following variables on the record **a_hhresp_ip**:

- | | |
|----------------|---|
| Group 1 | a_xpall_g1 |
| Group 2 | a_xpall_g2 |
| Group 3 | a_xpfood1_g3, a_xpfood2_g3, a_xpfood3_g3,
a_xpfood4_g3, a_xpfdout_g3, a_xpaltob_g3, a_xpclftw_g3, a_xphealth_g3,
a_xptrans_g3, a_xptel_g3, a_xprec_g3 |

At Wave 6, a further experiment was incorporated to address concerns about the way respondents report information on expenditure and consumption. While the experiment was administered in the household questionnaire, the questions were about “benefit unit” consumption rather than household consumption. For the purposes of this research, a ‘benefit unit’ was defined as an adult, their spouse or partner, and any dependent children under the age of 18 living in the household. The experimental questions were only asked if the household respondent qualified as responsible for paying the bills, such as rent, mortgage, gas or electricity. The PNO of this person is listed in item **f_consper** on the record **f_hhresp_ip**.

Households were randomly allocated to two experimental treatments. In one treatment (Version A), respondents are asked to give total benefit unit expenditure by adding up a set of expenditure categories using a showcard to trigger recall of expenditure on each category. In the alternative treatment (Version B), respondents provided an amount of expenditure for each expenditure category, then reconciled the total amount spent for accuracy.

The variable in the data that controls allocation to treatments is **f_consexpw6** on the record **f_hhsamp_ip** taking a value of 1 if Version A and 2 if Version B.

The substantive information from each experimental treatment can be found in the following variables on the record **f_hhresp_ip**:

f_consper and **f_consintro** for all treatments

Version A	f_constotatshh; f_constotatsmm; f_constotatsss; f_constotal_a; f_constotatehh; f_constotatemm; f_constotateess; f_constrat1 to f_constrat5; f_consoth_code; f_opusltshh; f_opusltsmm; f_opusltsss; f_opusl1; f_opusl2; f_opusltehh; f_opusltemm; f_opusltess
Version B	f_expbrktshh; f_expbrktsmm; f_expbrktsss; f_expmort; f_expbills; f_exptrans; f_expfood; f_expclotthes; f_expchild; f_expdiy; f_exphealth; f_exphobby; f_exptreats; f_exphols; f_expgive; f_expoth; f_breaktot; f_breakad; f_expbrktehh; f_expbrktemm; f_expbrktess

9.2 Measuring satisfaction

The way you ask subjective evaluation questions (e.g., satisfaction) may have a big influence on the types of answers that you get. A series of experiments with question wording, format and placement have been run in Waves 1-3 and 5 of the IP. Experiments across Waves 1-3 concerned the measurement properties of various questions on satisfaction whereas experimentation in Wave 5 concerned the nature of the judgement. In addition, an experiment in the Wave 5 youth questionnaire examines satisfaction measurement amongst young people.

In Wave 1, the experiment compared 11 and 7 point scales for job satisfaction. There is debate about the number of response categories which is substantively meaningful for respondents and for analysis.

Within **PSUs**, households were randomly allocated to one of two conditions. All interviewed adults within household received the same experimental treatments. The conditions are:

Group 1	11 point scale, no showcard, only end-points labelled
Group 2	7 point scale, no showcard, only end-points labelled

The variable that controls allocation to treatments is **a_group2** on the record **a_hhsamp_ip**. The variables **a_jbsat_g1** and **a_jbsat_g2**, for groups 1 and 2 respectively, on the record **a_indresp_ip** contain the substantive information.

At Wave 2, the experiment expanded to other satisfaction items about the participant's satisfaction with their health, family income, leisure, job (if applicable) and their life overall.

As well as the mode and showcard experiments, households were independently assigned to treatment groups formed by varying question design, delivery and position within the interview.

All eligible adults within a household received the same experimental treatment. The Table below sets out the 10 treatment groups at Wave 2 for this experiment.

Table: Definition of experimental groups in Wave 2 experiment: measurement of satisfaction.

Group	Response Mode	Timing of Question
1	CASI Full-labels	Late in questionnaire
2	CASI Polar-labels	Late in questionnaire
3	F2F + showcard Full-labels	Late in questionnaire
4	Tel + F2F Full labels: branched	Late in questionnaire
5	F2F + showcard Polar labels	Late in questionnaire
6	Tel + F2F Polar-labels	Late in questionnaire
7	F2F + showcard Full-labels	Early in questionnaire
8	Tel + F2F Full labels: branched	Early in questionnaire
9	F2F + showcard Polar-labels	Early in questionnaire
10	Tel + F2F Polar labels	Early in questionnaire
CASI = Computer assisted self interview; Tel = telephone; F2F = Face-to-face		

The standard questions have an initial question, “How dissatisfied or satisfied are you with the following aspects of your situation: (a) your health; (b) the income of your household; (c) the amount of leisure time you have”? This was followed by the question, “Using the same scale, how dissatisfied or satisfied are you with your life overall?” Respondents who were employed or self-employed were in addition asked, “All things considered, which number best describes how dissatisfied or satisfied you are with your job overall?”

The responses were measured using a seven-point scale. There were 3 different delivery methods (showcard, oral, CASI). Groups 3, 5, 7 and 9 had the response categories on a showcard. Groups 4, 6, 8 and 10 had no showcard; the question was purely oral. Groups 1 and 2 were presented with the computer and asked to complete the question by themselves (CASI).

The response scale was presented in three different ways. For groups 1, 3 and 7 each of the points on the seven-point scale was labeled (Full labels: 1-stage): *7 Completely satisfied; 6 Mostly satisfied; 5 Somewhat satisfied; 4 Neither satisfied nor dissatisfied; 3 Somewhat dissatisfied; 2 Mostly dissatisfied; 1 Completely dissatisfied*. Participants in groups 4 and 8 were also able to answer using the fully-labelled scale, but the question was broken into two parts (branched), with the participant first being asked, “How dissatisfied or satisfied are you with your (life/job) overall? Would you say that you are... (1 Dissatisfied; 2 Neither dissatisfied nor satisfied; 3 Satisfied).” If the participant indicated that they were either dissatisfied or satisfied they were asked the follow-up question, “Are you somewhat, mostly or completely (dissatisfied/satisfied) with your (life/present job) overall? (1 Somewhat; 2 Mostly; 3 Completely).” The third treatment, for groups 2, 5, 6, 9 and 10, is the “Polar labels” option. In this treatment group, only the labels for the extreme points on the scale were conveyed; i.e., completely dissatisfied and completely satisfied.

The timing of the job satisfaction question in the questionnaire was fixed for all participants who had a job, following a section about employment or self-employment. The life

satisfaction questions were either asked early in the interview (about a quarter of the way through, Groups 7-10) or late in the interview (very near the end, Groups 1-6).

The controlling variable for the job satisfaction split-ballot design is **b_ff_jobsatw2** on the record **b_hhsamp_ip**. Values of this variable correspond to groups 1-6 in the table above.

The controlling variable for the split-ballot design of the remaining satisfaction items is **b_ff_lifesatw2** on the record **b_hhsamp_ip** with values corresponding to the entire set of experimental treatments outlined in the table above.

The substantive data for job satisfaction can be found in the items **b_jbsat_a** through **b_jbsat_f** on the record **b_indresp_ip**. The substantive data for the remaining satisfaction items can be found **b_lfsat1_a** through **b_lfsato_j** on the record **b_indresp_ip**.

At Wave 3 and Wave 6 this experiment was repeated with a rotation in experimental treatments. The life and job satisfaction split-ballot question wording experiment carried at Wave 2 was repeated at Waves 3 and 6. The experimental allocation at Waves 3 and 6 varied the Wave 2 allocation, however. As well as the showcard experiments, described above, households were independently assigned to experimental groups formed by varying question design, delivery and position within the interview. All eligible adults within a household received the same experimental treatment. Also, while the Wave 2 allocation was nested within the showcard experiment, the Waves 3 and 6 allocations were not. Therefore, a separate showcard was required for the satisfaction items for interviewers not otherwise issued with showcards at Wave 3, and there was no showcard experiment in Wave 6. There was no manipulation of placement for the life satisfaction items.

This manipulation was of households within **PSU**'s, therefore interviewers received different versions of these questions depending on which household they were interviewing.

Note that there was an error in the implementation of the Wave 3 satisfaction experiment, which meant that some respondents were asked the life-satisfaction questions twice, once at the beginning and once at the end of the questionnaire. For details see section Known Data Issues. For this reason, the exact experimental allocation from Wave 3 was replicated in Wave 6.

The questions at Waves 3 and 6 match Wave 2. The standard question-set involved an initial question, "How dissatisfied or satisfied are you with the following aspects of your situation: (a) your health; (b) the income of your household; (c) the amount of leisure time you have". This was then followed by the question, "Using the same scale, how dissatisfied or satisfied are you with your life overall?" Those participants who were employed or self-employed were asked, "All things considered, which number best describes how dissatisfied or satisfied you are with your job overall?" The responses were measured using a seven-point scale. Whereas in Wave 2, 10 different experimental treatment groups captured variation in the presentation of satisfaction items, only 6 different groups were used at Waves 3 and 6. The group allocation was permuted for Waves 3 and 6 to achieve higher sample sizes across groups when waves are pooled.

For the Life Satisfaction Items, the permutation of treatments is as follows:

Table: Treatment groups in Wave 2 and Waves 3 and 6 for experiment on measurement of life Satisfaction

Group	Treatment at Wave 2	Group	Treatment at Waves 3 and 6
1	Full labels, CASI, beginning	2	Polar labels, CASI
2	Polar labels, CASI, beginning	5	Polar labels, showcards
3	Full labels, showcards, Beginning	1	Full labels, CASI
4	Full labels, unfolding design, Beginning	6	Polar labels, no showcards
5	Polar labels, showcards, Beginning	3	Full labels, showcards
6	Polar labels, no showcards, Beginning	4	Full labels, no showcard, branched rating
7	Full labels, showcards, end	4	Full labels, no showcard, branched rating
8	Full labels, unfolding design, end	3	Full labels, showcards
9	Polar labels, showcards, end	6	Polar labels, no showcards
10	Polar labels, no showcards, end	5	Polar labels, showcards

At Wave 3, the controlling variable is **c_ff_lifesatw3** on record **c_hhsamp_ip** while at wave 6 the controlling variable is **f_ff_lifesatw3** on record **c_hhsamp_ip**. In both Wave 3 and Wave 6, this variable is coded:

- 1 fully labelled CASI, end of interview
- 2 polar point labelled CASI, end of interview
- 3 fully labelled with showcards, end of interview
- 4 fully labelled without showcards, unfolding design, end of interview
- 5 polar-point labelled with showcards, end of interview
- 6 polar-point labelled without showcards, end of interview

Groups 3 and 5 had the response categories presented visually on a showcard. Groups 4 and 6 had no visual cue, the question was purely oral. Groups 1 and 2 were presented with the computer and asked to complete the question by themselves (CASI).

As well as the different delivery methods at Waves 3 and 6 (showcard, oral, CASI), the response scale was presented in three different ways. For groups 1, 3 and 4 each of the points on the seven-point scale was labelled; *7 Completely satisfied; 6 Mostly satisfied; 5 Somewhat satisfied; 4 Neither satisfied nor dissatisfied; 3 Somewhat dissatisfied; 2 Mostly dissatisfied; 1 Completely dissatisfied*. This is the “Full labels: 1-stage” response scale. Participants in group 4 were also able to answer using the fully-labelled scale, but the question was broken into two parts, with the participant first being asked, “How dissatisfied or satisfied are you with your (life/job) overall? Would you say that you are... (1 Dissatisfied; 2 Neither dissatisfied not satisfied; 3 Satisfied).” If the participant indicated that they were either dissatisfied or satisfied they were asked the follow-up question, “Are you somewhat, mostly or completely (dissatisfied/satisfied) with your (life/present job) overall? (1 Somewhat; 2 Mostly; 3 Completely)”. The third treatment, for groups 2, 5, and 6 is the

“Polar labels” option. In this treatment group, only the labels for the extreme points on the scale were conveyed; i.e., completely dissatisfied and completely satisfied.

Wave 3 substantive data for job satisfaction can be found in the items **c_jbsat_a** through **c_jbsat_f** on the record **c_indresp_ip**. Comparable wave 6 variables are **f_jbsat_a** through **f_jbsat_f** on the record **f_indresp_ip**.

Wave 3 substantive data for the remaining satisfaction items can be found **c_lfsat1_a** through **c_lfsato_f** on the record **c_indresp_ip**. Comparable variables for wave 6 are **f_lfsat1_a** through **f_lfsato_f** on the record **f_indresp_ip**.

9.3 Reference groups in measuring satisfaction

At Wave 5, an experiment examined the comparison groups women have in mind when they answer questions about life satisfaction. Respondents were randomly allocated to one of four treatments groups. The first three experimental groups had respondents rate their life satisfaction comparing themselves to (1) others of the same gender, (2) others with the same level of education, or (3) others of their gender and education. The fourth experimental treatment group was a control where respondents provided a measure of satisfaction without reference to any comparison group. All satisfaction items were subject to this experiment: employment, health, leisure, income and overall life satisfaction, as well as job satisfaction.

The variable **e_ff_lifesatw5** on the record **e_hhsamp_ip** controls allocation to versions of this question. This variable is coded:

- | | |
|---|-----------------------------------|
| 1 | Same gender comparison |
| 2 | Same education comparison |
| 3 | Same gender, education comparison |
| 4 | Control group |

The substantive variables containing data from this split-ballot experiment are **e_scflsat1_sg** through **e_scflsato**, and **e_jbsat_sg** through **e_jbsat** on the record **e_indresp_ip**.

9.4 Measuring identity

In Wave 2, the Innovation Panel tested a set of questions on identity based on items carried in the Citizenship Survey in 2007-2008. In total, respondents were given 13 categories with which they could identify (or not). This experiment compared endorsement of the words “profession” versus “occupation” via a split-ballot design. Within **PSUs**, households were randomly allocated to experimental treatment and all adults within households received the same treatment:

Group 1 – Occupation

Group 2 – Profession

The variable **b_ff_identityw2** on the record **b_hhsamp_ip** controls allocation to versions of this question. The substantive variables containing data from this split-ballot experiment are **b_ethexp_aa** through **b_ethexp_bm** on the record **b_indresp_ip**.

9.5 Measures of wealth

Wave 3 of the Innovation Panel was used to examine four different question designs for collecting the amount of money held in savings and investments:

Table: Schematic design of Wave 3 measuring wealth experiment treatments

	Individual Reporting	Financial Reporting
Aggregate Amounts	<i>Group 1</i>	<i>Group 2</i>
Itemized Amounts	<i>Group 3</i>	<i>Group 4</i>

For all groups, an initial question asks respondents to report which specific types of savings or investments they hold. The two factors concern whether reports are collected as aggregate amounts or itemized (itemization) and whether reports are from each individual adult or about the entire household (individual vs. financial reporting). Under *financial* reporting, all individuals reported on savings or deposit accounts, National Savings accounts, Individual Saving Accounts (ISAs), and Premium Bonds, whereas the ‘financial reporter’ at the household level reported on National Savings certificates, Unit or Investment Trusts, all Shares apart from ISAs, National Savings bonds, and any other investment the household may have.

Households within PSUs were randomly assigned to one of the four experimental treatments. Allocations were independent of the showcard experiment. All individuals interviewed as part of the household were treated the same. The controlling variable is **C_FF_HHDAW3** on record **C_HHSAMP_IP** and is coded as follows:

Group 1 – Aggregation, Individual reporting

Group 2 – Aggregation, Financial reporting

Group 3 – Itemization, Individual reporting

Group 4 – Itemization, Financial reporting

The variables affected by the design are as follows:

Table: Variables in experiment measuring wealth

Group 1	Group 2	Group 3	Group 4
	<i>Household Questionnaire:</i> C_NVESTRT_A4 to C_NVESTRT_B96 on record C_HHRESP_IP C_NVESTKRT to C_NVESTSPRT on record C_HHRESP_IP		<i>Household Questionnaire:</i> C_NVESTRT_A4 to C_NVESTRT_B96 on record C_HHRESP_IP C_INVESTRTCODE to C_NVESTLSPRT on record C_HHOLDINVEST_IP
<i>Individual Questionnaire:</i> C_NVEST_A1 to C_NVEST_A96, C_NVEST_B1 to C_NVEST_B96, C_NVEST1 to C_NVEST96 on record C_INDRESP_IP C_SVACK to C_NVESTC4 on record C_INDRESP_IP C_NVESTSJ to C_NVESTSP on record C_INDRESP_IP	<i>Individual Questionnaire:</i> C_NVEST_A1 to C_NVEST_A96, C_NVEST_B1 to C_NVEST_B96, C_NVEST1 to C_NVEST96 on record C_INDRESP_IP C_SVACK to C_NVESTC4 on record C_INDRESP_IP	<i>Individual Questionnaire:</i> C_NVEST_A1 to C_NVEST_A96, C_NVEST_B1 to C_NVEST_B96, C_NVEST1 to C_NVEST96 on record C_INDRESP_IP C_INVESTCODE to C_NVESTLSP on record C_WEALTH_IP	<i>Individual Questionnaire:</i> C_NVEST_A1 to C_NVEST_A96, C_NVEST_B1 to C_NVEST_B96, C_NVEST1 to C_NVEST96 on record C_INDRESP_IP C_INVESTCODE to C_NVESTLSP on record C_WEALTH_IP

9.6 Context of questions about consent to data linkage with administrative records

Often the survey design decisions regarding how to obtain consent for administrative data linkage to social survey data are based on anecdotal accounts and common sense rather than empirical evidence. This set of experiments and additional data collection examined: (a) the reasons for consenting or not consenting; (b) whether survey context matters in asking consent; (c) the stage in the life of a panel in which data linkage should be performed; and (d) the method of re-asking consent when maintaining consent in a panel is at issue.

The experiment has two factors. First, consent to link to benefit and tax credit records held by the Department for Work and Pensions was asked either just after questions on the amounts received in these forms of unearned income, or at the end of the questionnaire (*context* dimension). Second, respondents were either reminded of their consent or non-consent given previously and asked whether this consent should still apply (*dependent*

interviewing) or asked independent of any previous consent given (*independent interviewing*). If no information about prior consent was available, then the consent question was asked independently.

Table: Schematic design of Wave 4 consent experiment

	Independent Interviewing	Dependent Interviewing
Ask consent in context	<i>Group 1</i>	<i>Group 2</i>
Consent asked at end of interview	<i>Group 3</i>	<i>Group 4</i>

The controlling variable for this experiment is **d_ff_conexpw4** on record **d_hhsamp_ip**:

Group 1 – Independent question within context

Group 2 – Dependent question within context

Group 3 – Independent question at end of interview

Group 4 – Dependent question at end of interview

The relevant variables affected by this experiment are **d_bncn_a1** to **d_bcoi_a**, **b_bncn_b1** to **d_bcoi_b**, **d_bcsig**, **d_bcrat1** to **d_bcrat9**, and **d_bcchnge** on record **d_indresp_ip**.

9.7 Respondent preferences about mode of data collection

In Wave 4 a split-ballot experiment was used to collect information about the preferred mode of administration for the survey. The ordering of questions about mode preference was randomly varied across two different experimental groups: (a) first, ratings of how likely the survey participant is to respond in specific modes, followed by questions about which mode is preferred and least preferred (generalized preference) vs. (b) a generalized mode preference questions followed by ratings of specific modes. The experiment was repeated at Waves 5 and 6 with the same allocation.

The Wave 4 controlling variable for this split-ballot design is **d_ff_group2** on record **d_hhsamp_ip** and is coded:

Group 1 Ratings of specific modes then generalized preference.

Group 2 Generalized preference then ratings of specific modes.

The relevant variables affected by this experiment are **d_mpint_a** to **d_mpnot_a**, and **d_mpint_b** to **d_mpweb_b** on record **d_indresp_ip**.

The controlling variable for this design in Wave 5 is **e_ff_group2** on record **e_hhsamp_ip**, with relevant variables affected being **e_mpint_a** to **e_mpnot_a**, and **e_mpint_b** to **e_mpweb_b** on record **e_indresp_ip**.

The controlling variable for this design in Wave 6 is **f_ff_group2** on record **f_hhsamp_ip**, with relevant variables affected being **f_mpint_a** to **f_mpnot_a**, and **f_mpint_b** to **f_mpweb_b** on record **f_indresp_ip**.

9.8 Feasibility of directly measuring household energy use

This experiment assesses the feasibility of investigating certain key environmental behaviours and outcomes directly. Specifically, IP respondents were asked to report on two occasions meter readings for gas and electricity, and odometer readings for the motor vehicle used most often by the household. The experiment examines whether providing respondents with advance notice of the required data encourages the provision of the data within the interview. Collecting the readings on two occasions would allow the first reading to be subtracted from the second, thereby directly measuring vehicle and energy use over an identifiable period of time. Errors in implementing the experiment at Wave 5 meant that data from the experiment were not released. Instead, the experiment was re-implemented Wave 6.

The experiment takes a 2×2 design:

Table: Design of household energy use experiment

	Gas, electric and odometer reading	Odometer reading alone
Notification of meter readings in advance letter	Group 1	Group 2
No notification	Group 3	Group 4

Randomisation was of households within PSUs into one of the four groups for the experiment.

Approximately 4 weeks after the standard fieldwork period, a follow-up postal questionnaire collected the matching information gathered in the main Wave 6 interview from households providing at least one meter reading (N=825). The household questionnaire respondent was notified during the interview of this subsequent data collection and questionnaires were sent to this named individual. To be able to compare readings from two occasions, respondents to the follow-up confirmed their address and the make and mode of the vehicle about which they reported in the main interview.

After two-weeks, any non-response was followed-up with a reminder and a telephone follow-up was used after a further 10 days of non-response to collect information through the alternative mode.

Controlling variable if **f_ff_metersw6** on record **f_hhsamp_ip**, and takes the values:

- 1 Gas, Electric, Odometer, advanced letter warning
- 2 Odometer only, advanced letter warning
- 3 Gas, Electric, Odometer, no advanced letter warning
- 4 Odometer only, no advanced letter warning

Substantive variables containing initial meter readings are **f_gasuse**, **f_gasmeter**, **f_gasest**, **f_elecuse**, **f_elecmeter**, **f_varmeter1**, **f_varmeter2**, **f_elecest**, **f_meterfol**, **f_odouse**, **f_odometer**, **f_odoest**, **f_odofof** on the record **f_hhresp_ip**.

Follow-up meter reading data are contained in variables **f_intlen1**, **f_modetype**, **f_remind**, **f_ff_metersw6**, **f_qutype**, **f_mrfaddchck**, **f_mrfuelhave1**, **f_mrfuelhave2**, **f_mrfuelhave3**, **f_mrfuelhave4**, **f_mrfuelhave96**, **f_mrgasuse**, **f_mrgasometer**, **f_mrgasest**, **f_mrelecuse**, **f_mrelecmeter**, **f_mrvarmeter1**, **f_mrvarmeter2**, **f_mrelecest**, **f_mrvehchck**, **f_mrodouse**, **f_mrodometer**, **f_mrodoest**, **f_readdated**, **f_readdatem**, **f_readdatey** on record **f_meterreading_ip**.

9.9 Context effects in fertility decisions

This Wave 4 experiment examines priming effects that impact on the respondent's thoughts about expected total fertility. Fertility intentions can be highly dependent on wider context, such as partnership, age, actual childbearing, economic position and social pressures. Since long-running panel studies often rotate questionnaire content, there is concern that changes in answers over time may reflect the changing context effects to certain measures. That is, observed change in longitudinal data may reflect the priming effects of preceding questions. The experiment was repeated at Wave 5 with the same allocation.

Households within PSUs were randomly allocated to receive expected fertility questions either before or after questions about friendship networks.

The Wave 4 controlling variable is **d_ff_fertw4** on record **d_hhsamp_ip** and is coded:

- | | |
|----------------|---|
| Group 1 | Version A (before friendship questions) |
| Group 2 | Version B (after friendship questions) |

The expected fertility questions are *"Do you think you will have any (more) children?"* And, if the answer is yes, then *"How many (more) children do you think you will have?"* . The affected variables are **d_lchmor_a** to **d_lchmorn_a**, and **d_lchmor_b** to **d_lchmorn_b** on record **d_indresp_ip**.

The Wave 5 controlling variable is **e_ff_fertw4** on record **e_hhsamp_ip**, with affected variables **e_lchmor_a** to **e_lchmorn_a**, and **e_lchmor_b** to **e_lchmorn_b** on record **e_indresp_ip**.

Note, this experiment applies only to continuing respondents being administered the CASI self-completion instrument at Waves 4 and 5.

9.10 Vignettes: measuring partner satisfaction with division of household labour

In Waves 5 and 6 respondents evaluated a set of vignettes related to the sharing of domestic work.

The hypothetical arrangements varied along five dimensions: (1) paid work; (2) earnings; (3) presence of children; (4) housework allocations; and (5) use of paid help. All adult respondents were asked to report their expected level of satisfaction with the set of hypothetical household division of labour arrangements using a seven-point scale, from

completely dissatisfied, 1, to completely satisfied, 7. Each respondent was asked to rate three randomly allocated vignettes varying over the five dimensions.

Household within PSU were randomly allocated to treatments and all individuals within the household received the same set of hypothetical vignettes to evaluate. The allocation to vignettes allows sufficient number of cases to do population level analyses. The exact allocation at Wave 5 was replicated at Wave 6.

The table below outlines the variables and categories to which respondents were allocated, where “w” indicates wave and takes values “E” and “F” for Waves 5 and 6 respectively. The variables are **w_ff_paid_work1** to **w_ff_paid_work3**, **w_ff_earnings1** to **w_ff_earnings3**, **w_ff_children1** to **w_ff_children3**, **w_ff_housework1** to **w_ff_housework3**, and **w_ff_paid_housework1** to **w_ff_paid_housework3** on record **w_hhsamp_ip**. Please see the questionnaire for overall wording of each vignette.

Table: Dimensions in division of household labour vignettes

Controlling variables	Dimension	Value & wording
w_FF_PAID_WORK1 to w_FF_PAID_WORK3	Paid work	Group 1: you and your partner both have full time jobs Group 2: you and your partner have part time jobs and both of you work 2 and a half days per week Group 3: you and your partner both have jobs, you work full time while your partner works 2 and a half days per week Group 4: you and your partner both have jobs, your partner works full time while you work 2 and a half days per week
w_FF_EARNINGS1 to w_FF_EARNINGS3	Earnings	Group 1: and your partner has an hourly pay which is twice as much as yours Group 2: and your hourly pay is twice as much as your partner Group 3: and you have approximately the same hourly pay
w_FF_CHILDREN1 to w_FF_CHILDREN3	Children	Group 1: no children Group 2: one child aged 6 months Group 3: one child aged 5 years Group 4: one child aged 15 years
w_FF_HOUSEWORK1 to w_FF_HOUSEWORK3	Housework	Group 1: your partner does all of the housework while you do none of it Group 2: your partner does three quarters of the housework while you do one quarter of it

		Group 3: you and your partner share the housework equally Group 4: your partner does one quarter of the housework while you do three quarters of it Group 5: you do all of the housework while your partner does none of it
w_FF_PAID_HOUSEWORK1 to w_FF_PAID_HOUSEWORK3	Paid housework	Group 1: but you employ somebody to help with the housework one morning per week Group 2: and you do not employ anybody to help with the housework

The satisfaction ratings for the three vignettes are in **w_vig1**, **w_vig2**, **w_vig3** and **w_hwsat** on record **w_indresp_ip**.

9.11 Subjective expectations about the returns to higher education and decisions to attend university

This Wave 5 study tested ways of measuring perceptions about the economic returns to higher education, the costs involved, and experimentally tests the effects of information on actual decisions to attend university.

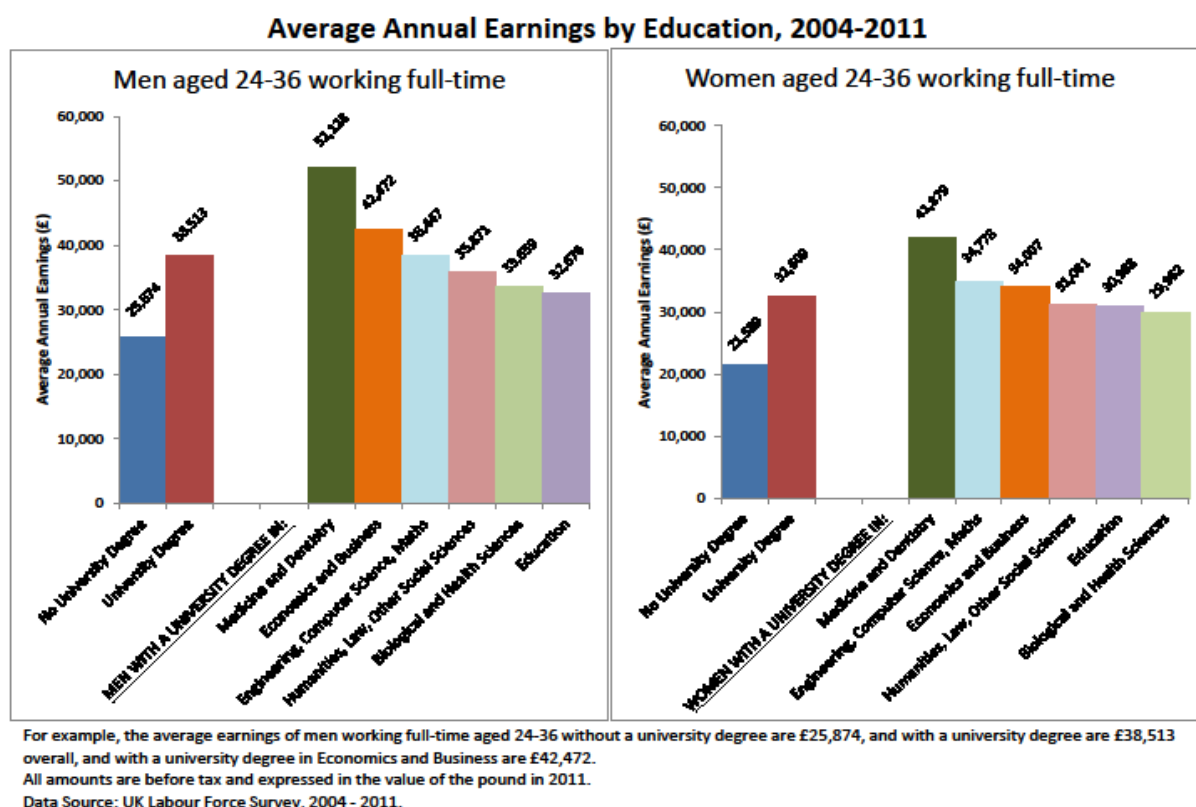
Respondents aged 16-21, and one of their parents, were asked questions about their expectations regarding the following: the likelihood of achieving A-levels, of applying to university for different subjects, of being accepted at university, expected costs, expectations of finding a job and of earnings conditional on having a university degree in a particular subject or conditional on having only a high school degree. This content is available at items **e_oddsqual** to **e_earninfo**, and **e_kidstat** to **e_pearinfo** on record **e_indresp_ip**.

In addition, half of the respondents were provided with information about the economic returns to higher education consisting of the distribution of wages among university graduates in various degree areas by gender. The other half of the sample did not receive any information. The controlling variable for this information treatment is **e_ff_wageinfow5** on record **e_hhsamp_ip**:

Group 1	Wage information
Group 2	No wage information

The information treatment provided to respondents is shown in Figure 3, Average annual earnings by education, 2004-2011

Figure 3. Average annual earnings by education, 2004-2011



In Wave 8 this experiment was repeated. All responding adults aged 16-21 not in higher education were asked about the perceived costs and benefits of obtaining a higher education degree. Parents of children age 10-21 not in higher education were asked about their expectations for their child, vis., the costs and benefits of that child obtaining higher education. For parents, one child was selected about whom they were asked to report: the eldest co-resident child aged 16-21 not in higher education and if no such child then the eldest aged 10-15, and if no such child then they were not asked.

In Wave 9, respondents who answered the Wave 8 questions were asked the same questions again, but without an information treatment. Between waves 8 and 9, all respondents who received the information during the wave 8 interview were posted a copy of the same information.

A random half of all adults responding to these questions received information on earnings for men and women across a range of occupations. Randomisation was at the household level, meaning households within **PSUs** were randomly allocated to treatment.

Controlling variable: ff_wageinfow8

0 = no wage information

1 = wage information

In **wave 9**, there is no random allocation for this experiment. Fed forward indicator flags identify those who responded at Wave 8, and for parents, the child asked about. In the survey the personal information relating to this youth pidp was pre-filled in the parental expectations module.

Fed forward indicators :

ff_yahechoiceip8

0 = Not asked young adult education expectations module
1 = Asked young adult education expectations module

ff_paredexpectip8

0 = Not asked parental education expectations module
1 = Asked parental education expectations module

ff_paredaboutip8

takes on the pidp of the child the parent was asked about in the wave 8 parental expectations module. If ***ff_paredexpectip8 = 1***

The contents is available in record **h_indresp**: **heage30, oddsqual, heapplied, oddsapply, oddsschol, xptuition, xpborrow, unifin1, unisub1, oddsemp1, oddsemp3, xpearn1, xpearn1dk, xpearn4, xpearn4dk, xpearn3, xpearn3dk, xpearn5, xpearn5dk, xpearng1, xpearng1dk, xpearng3, xpearng3dk, xploanpy, earnemail, kidstat, katuni, pedchpno, pheage30, poddsqual, pheapplied, poddsapply, poddsschol, pxptuition, pxpborrow, punifin1, punisub1, poddsemp1, poddsemp3, pxpearn1, pxpearn1dk, pxpearn4, pxpearn4dk, pxpearn3, pxpearn3dk, pxpearn5, pxpearn5dk, pxpearng1, pxpearng1dk, pxpearng3, pxpearng3dk, pxploanpy, pearnemail**

The contents is available in record **i_indresp**: **hemakeappl, heapplstat, heage30, oddsqual, oddsapply, oddsschol, xptuition, xpborrow, unifin1, unifin1_new, unisub1, oddsemp1, oddsemp3, xpearn1, xpearn1dk, Xpearn4, Xpearn4dk, xpearn3, xpearn3dk, xpearn5, xpearn5dk, xpearng1, xpearng1dk, xpearng3, xpearng3dk, xploanpy, kidstat, katuni, pheage30, poddsqual, pheapplied, poddsapply, poddsschol, pxptuition, pxpborrow, pacborrow, punifin1, pacunifin, punisub1, poddsemp1, poddsemp3, pxpearn1, pxpearn1dk, pxpearn4, pxpearn4dk, pxpearn3, pxpearn3dk, pxpearn5, pxpearn5dk.**

9.12 Measuring change in self-assessed disability

This experimentation uses dependent interviewing to investigate the measurement of change in self-assessed measures of long-standing illness or disability. The current UKHLS method of obtaining information about disability is to use a yes/no question about long-standing conditions which is followed-up by a check-list of 11 areas of everyday life where people may have difficulties (i.e., “Activities of Daily Living”, or ADL indicators). The **Wave 6** sample was split into two experimental treatments and a control group. The first experimental treatment compared the response to the initial question about long-standing conditions to the previous wave’s response and followed-up any responses which were different from previous with a “Why?” question. The ADL check-list was administered at a

later point in the questionnaire for this set of respondents. The second experimental condition asked only the ADL check-list without filtering on whether the respondent reported that they had a long-standing condition. Finally, the control group repeated the original UKHLS approach.

Randomisation was of households within PSUs into the experimental treatments and control as follows: one-half of the sample into the first experimental treatment, one-quarter of the sample into the second experimental treatment, and the remaining quarter of the sample into the control group. All new entrants received the control version by default.

This experiment is controlled by variable **ff_disabilityw6** on record **hhsamp_ip**. It takes the values:

- Group 1** – edit check/explain inconsistencies, move DISDIF to annual events
- Group 2** – only DISDIF, are not asked HEALTH
- Group 3** – original design

The experiment was repeated in Wave 7 with the same allocations to treatments. The IP7 refreshment sample was allocated to Group 3.

Affected variables on the **indresp** records for Wave 6 and Wave 7 are: **health, hthrdia11, hthrdia12, hthrdia13, hthrdia14, hthrdia15, hthrdia16, hthrdia21, hthrdia22, hthrdia23, hthrdia24, hthrdia25, hthrdia26, disdif1, disdif2, disdif3, disdif4, disdif5, disdif6, disdif7, disdif8, disdif9, disdif10, disdif11, disdif12, disdif96, disdifb1, disdifb2, disdifb3, disdifb4, disdifb5, disdifb6, disdifb7, disdifb8, disdifb9, disdifb10, disdifb11, disdifb12, disdifb96, disdifa1, disdifa2, disdifa3, disdifa4, disdifa5, disdifa6, disdifa7, disdifa8, disdifa9, disdifa10, disdifa11, disdifa12, disdifa96**

9.13 Associated Study: measuring time and risk preferences

These data are from an Associated Study, which was approved and implemented as part of Wave 6 and Wave 7 data collection.

Researchers using these data should acknowledge their source in any publications arising from analysis of the data. The citation for the data includes the following information:

"This paper also makes use of risk and time preferences data collected as part of the Future Research Leader project Linking Experimental and Survey Data: Behavioural Experiments in Health and Wellbeing (ES/K001965/1, PI: MM Galizzi), funded by the Economic and Social Research Council"

The study aims to combine survey data from the Innovation Panel with data on risk and time preferences. Risk preference is defined as the attitude for taking a gamble and is operationalized by systematically asking respondents to choose between lotteries yielding different pay-out probabilities and different pay out amounts. Through a series of 18

questions, pay-out amounts and pay-out probabilities were varied thus allowing for risk preference to be quantified.

Time preference is defined as the degree to which time closer to the present is valued more highly than time more distant in the future. It is operationalized by systematically asking respondents to choose between receipt of money closer or further away in time with interest. Through a series of 72 questions, the time differential and interest rate varied thus allowing for time preference to be quantified.

A target sample of around 580 respondents was selected such that only one respondent participated per household. Household within PSUs were randomly selected and then selection of respondents within households was made with a Kish grid of enumerated adults. One-tenth of selected participants were given a payment upon completion of the questions. Among those selected to receive a payment, the amount was based on one of the 91 questions that they answered. If one of the lottery questions was selected, the preferred lottery was played and pay-out was made accordingly. The study was replicated at Wave 7 among the same set of respondents as at Wave 6 but with re-randomisation of pay-outs.

At Wave 6, the selection of respondents and items for pay-out are controlled by the variables F_FF_TIMERISKW6, F_FF_TRSEL2, F_FF_TRSEL3, F_FF_TRSEL4, F_FF_TRSEL5, F_FF_TRWINW6, F_FF_TRQW6, and F_FF_TRDW6, on record F_HHSAMP_IP. The variable F_FF_TIMERISKW6 indicates selected households. The variables F_FF_TRSEL2 through F_FF_TRSEL5 were used to create the Kish grid. The variable F_FF_TRWINW6 was randomised within those respondents selected to receive the lottery questions and determine whether they received a pay-out. The variable F_FF_TRQW6 indicates from which question the pay-out was determined. Finally, if a 'lottery' question was selected, the variable F_FF_TRDW6 indicates the randomised outcome relevant for determining the lottery pay-out.

Wave 6 substantive variables on record **indresp** are **f_trflag**, **f_trnotes**, **f_trpre**, **f_trprei**, **f_choice01** through **f_choice71**, **f_riskpre**, **f_riskpreb**, **f_trenda**, **f_choice73** through **f_choice91**, **f_trend**, **f_selpay**, **f_dieroll**, **f_trnowin**, **f_trwin**, **f_drawball**, **f_payout0172**, **f_payout7390**, **f_payout91**, **f_trdie**, **f_trpayout**, **f_runtr**, **f_tmprf3**, **f_tmprf12**, **f_tmprf1**, **f_trriska**, **f_trhlrisk**, **f_trflrisk**, **f_trimpat**, **f_trimpul**, **f_trwemwba**, **f_trsmoker**, **f_trncigs**, **f_trevralc**, **f_tregalco**, **f_trdklm**, **f_tr5alcdr**, **f_trdrnkyr**, **f_trffdwk**, **f_trjfd**, **f_trfrutppd**, **f_trhlwtr**, **f_trtrydiet**, **f_trexttype1** through **f_trexttype97**, and **f_trpsprt** which will appear on record **f_indresp_ip**.

At Wave 7 the controlling variables are

On record **indsamp:**

ff_trflag – This is the fed-forward value of the IP6 individual level selection flag calculated in GRIDVARIABLES_IP6. At IP6, households were selected via **ff_timeriskw6** but a single individual within each household was selected post enumeration and the value of this individual selection is the variable “TRFLAG”. Since the time individuals are being asked the time-risk preference again in IP7, we are retaining the IP6 individual selection and feeding forward the IP6 value for individuals rather than the household selection mechanism. **trflag**

is Coded 1 = Participant, blank for continuing respondents not selected at IP6, all IP7 new entrants, and the IP7 refreshment sample.

On record HHSAMP:

ff_trwinw7 – Fresh randomisation for IP7. Takes the values 1-10 without any value labels, missing for all *ff_trflag = blank*. Takes values 1-10

ff_trqw7 – Fresh randomisation for IP7. Takes the values 1-91 without any value labels, missing for all *ff_trwinw7 > 1*. Takes values 1-91.

ff_trdw7 – Fresh randomisation for IP7. Takes the values of 1-10 without any value labels, missing for all *ff_trqw7 < 73*.

Wave 7 substantive variables on record INDRESP are:

“Time and Risk Preference” module: **trpre, choice01-choice91, riskpre, iskpreb, trenda, trend**

“CASI Time Preference Control” module: **runtr, tmprf3, tmprf12, tmprf1, trriska, trhlrisk, trflrisk, trimpat, trimpul, trevalc, tregalco, trdklm, tr5alcdr, trfrutppd, dospertf1, dospertf2, dospertf3, dospertf4, dospertf5, dospertf6, dosperth1, dosperth2, dosperth3, dosperth4, dosperth5, dosperth6**

9.14 Assessing how people think about environmental taxes

This experiment examines different wordings of questions on willingness to pay environmental taxes. Each respondent receives a single question concerning willingness to pay an environmental tax. There are 10 question variants. Households within PSUs are allocated to treatment. This control variable affects variables **envtax_a** to **envtax_j** on record **indresp**.

Controlling variable on record HHSAMP: **ff_envtaxqw7**

- Group 1 – Question Version A
- Group 2 – Question Version B
- Group 3 – Question Version C
- Group 4 – Question Version D
- Group 5 – Question Version E
- Group 6 – Question Version F
- Group 7 – Question Version G
- Group 8 – Question Version H
- Group 9 – Question Version I
- Group 10 – Question Version J

9.15 Validity of interviewer ratings of respondent health

This experiment aims to explore what factors contribute to interviewers' assessments of respondents' health. The experiment varies when in a face-to-face interview the interviewer is asked to rate the respondent's health: at the beginning of the interview before any substantive questions are asked or at the end of interview. For telephone interviews, the interviewer's rating of the respondent's health occurred at the end of the interview to allow for a comparison to this circumstance in which interviewers are not able to observe physical cues. Any respondents interviewed via Web were excluded from the experiment.

Random allocation to treatment occurred at the individual level and was computed via system randomization within the questionnaire itself. The controlling variable for this experiment is **h_inthlthw8** on record **h_indreps_ip**:

- | | |
|---|--|
| 1 | Assessment at the start of the interview |
| 2 | Assessment at the end of the interview |

The interviewer assessment variables are in record **h_indresp**:

inthlthe, inthlthl

9.16 Social desirability bias in attitudes towards immigration

Intolerance to ethnic minority immigrants is often masked, which can lead to systematically under-reported opposition and over-reported tolerance. This experiment is a longitudinal application of a list experiment.

Households within **PSUs** were randomly allocated to either a control group or one of three experimental treatments.

Together with the item count lists outlined under the experiment concerning the measurement of sexual orientation (see 9.17 Measuring sexual identity using direct and indirect questioning), the ordering of item counts was randomised across respondents. The statements which respondents counted also had a randomised order across respondents.

At **Wave 9**, $\frac{1}{2}$ of the participants from a given group were independently allocated to one of the other three groups, control or treatment, other than the one assigned at Wave 8.

Controlling variables:

ff_ictimw8

- | | |
|---|-----------------------------|
| 1 | Control |
| 2 | Muslim countries item count |
| 3 | East European item count |
| 4 | Caribbean item count |

ff_ictimw9

- 1 Control
- 2 Muslim countries item count
- 3 East European item count
- 4 Caribbean item count

The variables used for this experiment are in record **h_indresp** and **i_indresp**: **ictimc1**, **ictimc1o**, **ictimt1**, **ictimt1o**, **ictimt2**, **ictimt2o**, **ictimt3**, **ictimt3o**, **ictimc2**, **ictimc3**, **ictimc4**.

9.17 Measuring sexual identity using direct and indirect questioning

The identification of the best strategies to measure sexual orientation is needed to inform policy makers of experience of stigma and harassment suffered by the lesbian, gay and bisexual population. However, the measurement of sexual orientation faces methodological difficulties, since sexuality is among the most sensitive topics in surveys. This experiment proposes to evaluate the self-administered UKHLS question on sexual identity against the Integrated Household Surveys (IHS) interviewer administered version. The experiment uses a two-list item count sensitive questioning (ICT) technique to obtain something akin to validation data on sexual orientation in order to evaluate the IHS and UKHLS direct questioning approaches. By repeating the two-list ICT longitudinally and rotating allocation of the sensitive item to lists, respondent's sexual identity can be directly ascertained, permitting a validated micro-level analysis.

Together with the item count lists outlined under the experiment concerning the measurement of attitudes to immigrants (see 9.16 Social desirability bias in attitudes towards immigration), the ordering of item counts was randomised across respondents. The statements which respondents counted also had a randomised order across respondents. Randomisation occurred during the interview with associated variables indicating the randomised orderings, their variable and value labels, specified in the questionnaire.

The experiment was a 2×2 design where half of respondents received either the ONS direct sexual identity question or the UKHLS direct sexual identity question. The crossed-treatment assigned respondents to one of a pair of item-counts concerning sexual attraction, sexual behaviour and sexual identity respectively.

The 2×2 design:

		ff_ictsexw8	
		Group A	Group B
ff_sexidw8	1 = UKHLS protocol	1 = Lists: A, B+S1, C, D+S2, E, F+S3 Plus the UKHLS direct	2 = Lists: A+S1, B, C+S2, D, E+S3, F Plus the UKHLS direct
	2 = IHS protocol	1 = Lists: A, B+S1, C, D+S2, E, F+S3	2 = Lists: A+S1, B, C+S2, D, E+S3, F

The longitudinal element:

		IP8	IP9
ff_ictsexw8	1 = Group A	List A	List A + S1
		List B + S1	List B
		List C	List C + S2
		List D + S2	List D
		List E	List E + S3
		List F + S3	List F
		Plus Direct Question A or B	
	2 = Group B	List A + S1	List A
		List B	List B + S1
		List C + S2	List C
		List D	List D + S2
		List E + S3	List E
		List F	List F + S3
		Plus Direct Question A or B	

The variables used for this experiment are in record **h_indresp** and **i_indresp**: **ictsexa**, **ictsexbs**, **ictsexao**, **ictsexbso**, **ictsexabso**, **ictsexxc**, **ictsexds**, **ictsexco**, **ictsexdso**, **ictsexcdso**, **ictsexe**, **ictsexfs**, **ictsexeo**, **ictsexfso**, **ictsexefso**, **ictsexas**, **ictsexb**, **ictsexaso**, **ictsexbo**, **ictsexasbo**, **ictsexcs**, **ictsexd**, **ictsexcso**, **ictsexdo**, **ictsexcsdo**, **ictsexes**, **ictsexf**, **ictsexeso**, **ictsexfo**, **ictsexesfo**, **sexuor**, **sidqn**.

9.18 What do the general population regard as “successful ageing”?

Populations are ageing and there is substantial interest among researchers, policy makers and the general population as to what constitutes ‘successful ageing’. At **Wave 9** and **Wave 17**, respondents were presented with a series of vignettes, each describing an older person.

At **Wave 9**, this experiment used three vignettes which presented different scenarios about the circumstances of a 75-year-old, across six life dimensions. Gender of the person described in the vignette is also varied. The six life dimension are: Chronic disease; Disability; Physical functioning; Cognitive functioning; Interpersonal engagement; and Productive engagement. Each dimension has a favourable and unfavourable outcome which were independently chosen, and allocated randomly at the respondent level.

At **Wave 17**, respondents were again presented with vignettes. The individual in the vignette was described with a favourable or unfavourable outcome in each of six dimensions. Respondents were asked after each vignette “How successfully is [name] ageing?” and gave their response on a scale of 0 (not successfully) to 10 (very successfully). This allowed respondents to give different weights to different dimensions according to how they judge their relative importance. A 2^6 factorial design was used to randomly allocate a subset of three vignettes to each respondent, with one of six name sets (with variations of typically-female and typically-male names) also randomly allocated.

Allocations

Allocations at Wave 9

Random allocation to treatment occurred at the individual level within the questionnaire itself. The controlling variables were computed via system randomisation at the start of the individual interview. Each vignette required a unique set of six controlling variables, one for each life dimension, while the names were allocated such that only one control variable was needed. The controlling variables, to be assigned in the questionnaire and in the record **i_indresp**:

ff_vigname: *sequence of names used in the three vignettes*

- 1 George, Margaret, Harry
- 2 George, Margaret, Anne
- 3 George, Harry, Margaret
- 4 Margaret, George, Harry
- 5 Margaret, George, Anne
- 6 Margaret, Anne, George

ff_vdisease1 (for vignette1), **ff_vdisease2** (for vignette2), **ff_vdisease3** (for vignette3)

- 1 No long-term illness
- 2 Diabetes

ff_vdisability1 (vignette1), **ff_vdisability2** (vignette2), **ff_vdisability3** (vignette3),

- 1 No difficulties climbing stairs
- 2 Difficulties climbing stairs

ff_vcognit1 (vignette1), **ff_vcognit2** (vignette2), **ff_vcognit3** (vignette3),

- 1 No problems remembering
- 2 Problems remembering

ff_vphys1 (vignette1), **ff_vphys2** (vignette2), **ff_vphys3** (vignette3)

- 1 Opens food packaging easily
- 2 Struggles to open food packaging

ff_vinterp1 (vignette1), **ff_vinterp2** (vignette2), **ff_vinterp3** (vignette3),

- 1 Sees [his/her] family and friends regularly
- 2 Rarely sees [his/her] family and

ff_vvolun1 (vignette1), **ff_vvolun2** (vignette2), **ff_vvolun3** (vignette3)

- 1 Often volunteers
- 2 Doesn't volunteer

Allocations at Wave 17

The variables controlling allocation to treatments, randomised at the individual level within the questionnaire script:

vignameip17 (1/6 allocated to each condition)

- 1 George, Margaret, Harry
- 2 George, Margaret, Anne
- 3 George, Harry, Margaret
- 4 Margaret, George, Harry
- 5 Margaret, George, Anne
- 6 Margaret, Anne, George

vdisease1ip17 (1/2 allocated to each condition)

- 1 No long-term illness
- 2 Diabetes

vdisease2ip17 (1/2 allocated to each condition)

- 1 No long-term illness
- 2 Diabetes

vdisease3ip17 (1/2 allocated to each condition)

- 1 No long-term illness
- 2 Diabetes

vdisability1ip17 (1/2 allocated to each condition)

- 1 No difficulties climbing stairs
- 2 Difficulties climbing stairs

vdisability2ip17 (1/2 allocated to each condition)

- 1 No difficulties climbing stairs
- 2 Difficulties climbing stairs

vdisability3ip17 (1/2 allocated to each condition)

- 1 No difficulties climbing stairs
- 2 Difficulties climbing stairs

vcognit1ip17 (1/2 allocated to each condition)

1 No problems remembering

2 Problems remembering

vcognit2ip17 (1/2 allocated to each condition)

1 No problems remembering

2 Problems remembering

vcognit3ip17 (1/2 allocated to each condition)

1 No problems remembering

2 Problems remembering

vphys1ip17 (1/2 allocated to each condition)

1 opens food packaging easily

2 struggles to open food packaging

vphys2ip17 (1/2 allocated to each condition)

1 opens food packaging easily

2 struggles to open food packaging

vphys3ip17 (1/2 allocated to each condition)

1 opens food packaging easily

2 struggles to open food packaging

vinterp1ip17 (1/2 allocated to each condition)

1 Regularly

2 Rarely

vinterp2ip17 (1/2 allocated to each condition)

1 Regularly

2 Rarely

vinterp3ip17 (1/2 allocated to each condition)

1 Regularly

2 Rarely

vvolun1ip17 (1/2 allocated to each condition)

1 Often volunteers

2 Doesn't volunteer

vvolun2ip17 (1/2 allocated to each condition)

1 Often volunteers

2 Doesn't volunteer

vvolun3ip17 (1/2 allocated to each condition)

1 Often volunteers

2 Doesn't volunteer

Modules affected

At wave 9, the module was [scavignettes_ip9](#). The variables used for this experiment are in record **i_indresp**:

vign1, vign2, vign3

At Wave 17, the module was [scasuccessageing_ip17](#).

9.19 Benefits unit finances module

The first experiment presented respondents with an income summary screen, which displayed the person's derived monthly net income based on their responses to previous answers in the survey.

The second experiment was part of the Benefit Unit Finances module. This was a separate module that used derived information from the adult interviews about benefit unit net income (using net income and benefit receipt variables). In addition the module collected information about benefit unit expenditure in the last month, and changes in assets and debts. At the end of the module respondents were shown a reconciliation of whether their income minus spending matched their changes in assets and debts in the last months. This included an experiment with how respondents were asked about changes in their assets and debts. Two versions were used: the first version framed questions in terms of gross flows of money, the second asks about net flows of money.

The two experiments were crossed. Controlling variables are in the file **i_hhsamp_ip**:

ff_incomesummary

- 1 Income summary screen not used
- 2 Income summary screen used

ff_bufmodule

- 1 Version A: gross flows
- 2 Version B: net flows

The variables used for this experiment are in record **i_indresp**:

finsum, finsumchk, paynmnth_b, j2paymnth_b, jsmnth_b, frmnthtot_b, indinctot_b, further variables are in record i_bufind, i_bufinance, and i_benefitsum: rec1draw, rec1save, rec1buexp, rec1totin, rec1totout, rec1balance, rec1bal, buinctot_r1, rec1draw_r1, rec1buexp_r1, rec1save_r1, rec1totin_r1, rec1totout_r1, rec1balance_r1, rec2buexp, rec2savint, rec2bal1, rec2balamt1, csavings1, rec2bal2, rec2balamt2, csavings2, rec2bal3, rec2balamt3, csavings3, rec2bal4, rec2balamt4, csavings4, rec2detint, rec2detbal1, rec2detamt1, cdebts1, rec2detbal2, rec2detamt2, cdebts2, rec2detbal3, rec2detamt3, cdebts3, rec2detbal4, rec2detamt4, cdebts4, csavings, cdebts, rec2incbal, rec2acbal, rec2balance, rec2bal, buinctot_r2, rec2buexp_r2, csavings_r2, cdebts_r2, rec2incbal_r2, rec2acbal_r2, rec2balance_r2.

9.20 Presentation of response options in satisfaction questions

Recent findings suggest a significant drop in health satisfaction across waves of the Understanding Society survey. It is unclear if it is an actual change or artefact of the survey design. The presentation of response options also changed across waves, in terms of orientation of the response options, but research does not suggest such a change should be expected. The present experiment uses three different presentations of question and response options on the standard set of satisfaction question used in the self-completion component on the IP. One version had all satisfaction items presented on one page, with response options across the top and the items on the left in a grid format. The second version had each question presented on a separate page, with response options vertically aligned. The final version also had each question presented on a separate page, but response options were presented horizontally. The goal of the experiment is to identify if and how the presentation of response options in self-completion formats can impact estimates of satisfaction.

The allocation occurred at the household-level, with 1/3 of households randomly assigned to each of the version. The controlling variable is:

ff_satisw9 in file **i_hhsamp_ip**

- 1 Grid format
- 2 One question per screen, response options vertically aligned.
- 3 One question per screen, response options horizontally aligned.

The variables used for this experiment are in record **i_indresp**:

sclfsat1, sclfsat2, sclfsat7, sclfsato

9.21 Improving consent to link to the electoral register

Two experimental conditions exist for this experiment. First is the medium of obtaining consent; the second involves the wording to the consent request. Respondents were first assigned to either an “opt-in” or “opt-out” consent condition. In the opt-in condition, respondents were asked directly in the IP10 questionnaire about whether they consent to link their electoral registry data to their survey responses. The opt-out condition informed respondents of linkage during the inter-wave mailing, giving them a chance to send in a Freepost response asking to be excluded from the linkage.

For both the opt-in and opt-out conditions, there were two wording versions for the requests. The first used the BES 2015 wording, and the second adapted it to be more explanatory. Random allocation to treatments occurred at the household level with equal allocations of 50/50 to each condition, such that the end result is there is a 25/25/25/25 assigned to each of the possible combinations of the two conditions. The controlling variables for this experiment are

ff_eleclinkinw10 in record **j_hhsamp_ip**:

- 1 Electoral register linkage opt-in experimental group
- 2 Electoral register linkage opt-out experimental group

ff_elecwordw10 in record **j_hhsamp_ip**:

- 1 Electoral register linkage wording experiment Version A
- 2 Electoral register linkage wording experiment Version B

Variables used for this experiment are in record **j_indresp_ip**:

eleclink_a, eleclink_b, elecadd, elecnewctry, elecoptout10

At Wave 11, respondents in the opt-out condition were asked about their memory of the interwave mailing between IP10 and IP11. These follow-up measures were asked only of those in the opt-out condition (50% of the IP10 sample). The measures asked about memory about the request in the interwave mailing, their belief whether they consented or not given the information, the usefulness of linkage, and acceptability of this method to ask for linkage. Finally these respondents were asked directly for their consent to link to their data in the electoral register, to ensure their desires are accurately expressed. Random allocation occurred at IP10 and is fed forward to IP11 (variable **ff_eleclinkinw10** in record **k_hhsamp_ip**).

Variables used for this experiment are in record **k_indresp_ip**:

eleclinkmem, eleclinkcon, eleclinkuse, eleclinkacc, eleclink11

9.22 Financial management within couples

This experiment compares two ways of asking couples about their financial management and perception of money ownership. The first condition asks only about how joint expenses are shared, with a response option included about shared money. The second condition consists of eight questions surrounding management and perceptions of money in the relationship, all measured on a five-point, fully labelled scale ranging from strongly agree to strongly disagree. Random allocation to treatment occurred at the household level with an equal 50/50 split. The controlling variable for this experiment is:

ff_jointfinw10 in record **j_hhsamp_ip**:

- 1 Joint Finance Experiment Group 1 (Single Question)
- 2 Joint Finance Experiment Group 2 (Eight Questions)

The variables used for this experiment are in record **j_indresp_ip**:

shareexp, moneyrelat, moneybelong, moneygives, moneynow, moneyexpense, moneyborrow, moneyseparate, moneyequal

9.23 Non-resident parents and reasons for separation

Two question modules which have been used in the past in the Innovation Panel were adapted to capture information on non-resident parents. Both of these modules, Fertility History and Family Networks, were asked of all respondents using a within-respondent design (note: Fertility History is asked every wave of new entrants; Family Networks was last asked in the third wave of the IP, but has been asked in the main survey every other wave). Half of respondents were asked Fertility History first, with the other half asked Family Networks first; later in the survey respondents were asked the other module. Both modules were adapted to expand questions on children not residing in the household, including separation of biological, step-, and adopted children.

After whichever module is asked first, respondents were asked a series of questions about each non-resident child identified (e.g. name, DOB), and the respondent's reasons for separation from that child's other parent. After the second module, a summary screen of the children identified in the first asked module was presented, and any discrepancies identified based on the second module asked were reconciled, by either removing children incorrectly identified, updating the information previously provided, or adding new children not identified in the first module asked. For any updated children, the questions about information about the children and reasons for separation from the child's other parent were asked. There is one experimental control variable, determining which of the two adapted modules a respondent was asked first. Households were randomly assigned to one of the two following conditions with an equal 50/50 split. The variable controlling the treatment allocations is:

ff_nonresw10 on record **j_hhsamp_ip**:

- 1 Ask adapted Family Networks (Version A) module first
- 2 Ask adapted Fertility History (Version A) module first

The variables used for this experiment are the following.

On record **j_nonressum_ip**:

nonressumid, nrccsumdobm, nrccsumdoby, nrccsumsex, reltonrccsum, nrccsumnumber, fursumadd

On record **j_nonresid_ip**:

nrccidnumber, nrccdobm, nrccdoby, nrccsex, reltonrc

On record **j_nonresch2_ip**:

nrclives, nrcbirthc, staykidwnrc, ftexwnrc, childpno, reasonsep4, reasonsep8, reasonsep12, reasonsep16, reasonsep20, reasonsep24, reasonsep28, nonresseq, livedwithnrc, seekidnrc, farkidnrc, reltopwc, reasonsep1, reasonsep5, reasonsep9, reasonsep13, reasonsep17, reasonsep21, reasonsep25, reasonsepmain, reasonsepoth_code, nonresch2id, mthslwc, wekidnrc, relkidnrc, pwcseparatem, reasonsep2, reasonsep6, reasonsep10, reasonsep14, reasonsep18, reasonsep22, reasonsep26, sepdecision, sepdecisiono_code, pwcparnam, yrslwc, staykidnrc, kidsptnrc, pwcseparatey, reasonsep3, reasonsep7, reasonsep11, reasonsep15, reasonsep19, reasonsep23, reasonsep27

On record j_indresp_ip:

lvreladp11, lvreladp12, lvreladp14, lvreladp15, lvreladp16, lvreladp17, lvreladp18, lvreladp19, lvreladp110, lvreladp111, lvreladp112, lvreladp113, lvreladp196, mamostcon1, pamostcon1, maage1, paage1, nrels111, nrels112, nrels113, nrels114, nrels115, nrels116, nrels117, nrels118, parmar1, malone1, palone1, biou181, biou18num1, adopu181, adopu18num1, stepu181, stepu18num1, lchcliv1, lchcar1, lch2uk1, masee1, macon1, mafar1, pasee1, pacon1, pafar1, paaaid11, paaaid12, paaaid13, paaaid14, paaaid15, paaaid16, paaaid17, paaaid18, paaaid196, paaaid197, paidu11, paidu12, paidu13, paidu14, paidu15, paidu16, paidu17, paidu18, paidu196, paidu197, chsee1, chcon1, chfar1, chaid11, chaid12, chaid13, chaid14, chaid15, chaid16, chaid17, chaid18, chaid196, chaid197, caidu11, caidu12, caidu13, caidu14, caidu15, caidu16, caidu17, caidu18, caidu196, caidu197, lvreladp21, lvreladp22, lvreladp24, lvreladp25, lvreladp26, lvreladp27, lvreladp28, lvreladp29, lvreladp210, lvreladp211, lvreladp212, lvreladp213, lvreladp296, mamostcon2, pamostcon2, maage2, paage2, nrels211, nrels212, nrels213, nrels24, nrels25, nrels26, nrels27, nrels28, parmar2, malone2, palone2, biou182, biou18num2, adopu182, adopu18num2, stepu182, stepu18num2, lchcliv2, lchcar2, lch2uk2, masee2, macon2, mafar2, pasee2, pacon2, pafar2, paaaid21, paaaid22, paaaid23, paaaid24, paaaid25, paaaid26, paaaid27, paaaid28, paaaid296, paaaid297, paidu21, paidu22, paidu23, paidu24, paidu25, paidu26, paidu27, paidu28, paidu296, paidu297, chsee2, chcon2, chfar2, chaid21, chaid22, chaid23, chaid24, chaid25, chaid26, chaid27, chaid28, chaid296, chaid297, caidu21, caidu22, caidu23, caidu24, caidu25, caidu26, caidu27, caidu28, caidu296, caidu297, ladopted1, ladopted1, ladoptu181, nonresadoptc1, numadoptnrc1, lstep1, lstep1, lstepu181, nonresstepc1, numstepnrc1, lprntadp1, lprnt1, lbiou181, nonresbioc1, numbionrc1, totalnrca, totalnrca, ladopted2, ladopted2, ladoptu182, nonresadoptc2, numadoptnrc2, lstep2, lstep2, lstepu182, nonresstepc2, numstepnrc2, lprntadp2, lprnt2, lbiou182, nonresbioc2, numbionrc2

9.24 Variations of the EQ-5D questions

This module aims to explore the differences in measurement of the widely used EQ-5D instrument using two developed variants. The EQ-5D is a 5-item instrument, designed by a consortium called EuroQol, to measure health-related quality of life. Despite its obvious shortcomings, it is the principal measure recommended by NICE for measuring quality-adjusted life-years (QALYs). QALY measures are a key input into cost-effectiveness evaluations of drug therapies, medical technologies and other healthcare procedures commissioned by NICE for the English NHS. Largely because of the worldwide influence that NICE has, EQ-5D is also an important policy tool in at least nine other countries.

A new version of EQ-5D, designed to improve question wording and increase sensitivity, has been produced and is increasingly being adopted for clinical trials. The original version, with a 3-point response scale, is known as EQ-5D-3L, and the newer version with a 5-point scale is called EQ-5D-5L. The IP carried both versions, with the experiment varying whether the respondent was asked both the 3L and 5L versions or only the 5L version, and the location of these within the questionnaire. 1/3 of respondents were asked the 3L version in an early CASI module and 5L in the standard CASI module, 1/3 were asked the 5L in the early CASI module and 3L in the standard CASI module and 1/3 were only asked the 5L version in the standard version.

Random allocation to conditions occurred at the respondent level with an equal (1/3 each) chance. The controlling variable for this experiment is as follows.

eq5dcond on record **k_indresp_ip**:

- 1 ASK EQ-5D-3L Late and ASK EQ-5D-5L Early
- 2 ASK EQ-5D-3L Early and ASK EQ-5D-5L Late
- 3 ASK EQ-5D-5L Late ONLY

The variables used for this experiment are the following.

On record **k_indresp_ip**:

eq3lmobear, eq3lcareear, eq3lactear, eq3lpainear, eq3lanxear, eq5lmobear, eq5lcareear, eq5lactear, eq5lpainear, eq5lanxear, eq3lmoblat, eq3lcarelat, eq3lactlat, eq3lpainlat, eq3lanxlat, eq5lmoblat, eq5lcarelat, eq5lactlat, eq5lpainlat, eq5lanxlat

9.25 HMRC data linkage consent experiment

This experiment compares the impact of placement and wording of consent to data linkage to HMRC records. This consent question was originally carried in the mainstage of Understanding Society. This experiment varies the wording to this question to explore the standard version used in the mainstage survey and a new, “easier” to understand wording to ascertain the impact of this has on consent rates. These versions were varied with equal allocation across respondents in both web and face-to-face versions of the survey. For those responding face-to-face, the placement of this consent request also varied, being asked either early or late in the questionnaire. Respondents were provided with additional

information in the form of a leaflet with information and one with a flowchart on the linkage process and usage of data.

Random allocation to treatment occurred at the respondent level, and was stratified by response modes. Respondents were allocated to either standard or easy versions within modes to ensure a near 50/50 split within each. Within the face-to-face mode only, respondents were also allocated using a 50/50 split to being asked the consent question either early or late, which was stratified by question difficulty. After the question was allocated as easy or standard (within face-to-face), then the location was allocated for this version, to ensure another close to 50/50 within mode and question difficulty. The controlling variables for this experiment are:

condiffcawi on record **k_indresp_ip**:

- 1 Standard HMRC consent question for CAWI
- 2 Easy HMRC consent question for CAWI

condiffcapi on record **k_indresp_ip**:

- 1 Standard HMRC consent question for CAPI
- 2 Easy HMRC consent question for CAPI

conlocstand on record **k_indresp_ip**:

- 1 Ask HMRC standard consent early in CAPI
- 2 Ask HMRC standard consent late in CAPI

conloceasy on record **k_indresp_ip**:

- 1 Ask HMRC easy consent early in CAPI
- 2 Ask HMRC easy consent late in CAPI

The variables used for this experiment are the following.

On record **k_indresp_ip**:

consentq1, consentq2, consentq3, consentq4, intcread1, intflread1, intcread2, intflread2

On record **k_keystroke_paradata**:

k_keystrokes1, k_keystrokes2

At Wave 15, a follow-up experiment was conducted to examine reasons why consent rates are so much lower in web than face-to-face surveys. Respondents were asked a single consent question (tax record linkage) with some follow-up questions about how the respondent made the consent decision. Some of these questions were previously

implemented in IP11. This experiment exploited the mixed-mode design of IP15 (Web and CAPI), as well as the self-completion element (CASI) within the CAPI interview. The follow-up questions were all asked at the beginning of the CASI module, immediately after the consent question.

Face-to-face respondents were allocated to one of three groups, varying when the consent question was asked:

- 1) CAPI: the consent question was the last CAPI question before the start of the CASI module,
- 2) CASI: the consent question was the first in the CASI module, and
- 3) partial CASI: the consent question is the first in the CASI module, but the question is printed on a showcard and read by the interviewer and not displayed on screen, the respondent enters the answer in CASI.

Web respondents were allocated to one of two groups, varying how the consent question was asked:

- 1) the consent question included a photo of, and message from, the study director,
- 2) the control group were shown only the text of the consent question.

The variables controlling allocation to treatments, made equally across conditions at the household level, are on record **o_hhsamp_ip**:

ff_consentcasiw15 (1/3 each, allocation stratified by sampleorig ff_gridmodew15 ff_incentw15)

- 1 = CAPI
- 2 = CASI
- 3 = Partial CASI

ff_consentwebw15 (1/2 each, allocation stratified by ff_consentcasiw15 sampleorig ff_gridmodew15)

- 1 = Consent question only
- 2 = Consent question includes photo of, and message from, study director

The variables used for this experiment are the following, in file **o_indresp_ip**:

hmrcconsentq2, coundstd2b, coundstd2e, coundstd2f, coundstd2g, scac2, scac, scrx_code, scux_code, hmrcconsentq1, hmrcconsentq3, hmrcconsentq4, cdcsn1, cdcsn2, cdcsn3, cdcsn4, cdcsn5, cfactors1, cfactors2, cfactors3, cfactors4, cfactors5, cfactors6, cfactors7, cfactors8, cfactors9, cfactors10, scrf1, scrf2, scrf3, scrf4, scrf5, scrf6, scrf97, scun1, scun2, scun3, scun97

9.26 Does competition over public services decrease support for residency rights of immigrants?

This study contains two experimental sets of questions that explore respondents' views towards residency rights of immigrants. The two experimental sets are a conjoint experiment exploring the impact of various hypothetical immigrant characteristics on preferences and a vignette question that varies the immigrant's location to either Glasgow or the respondent's town of residence. Only respondents residing in England were asked these experimental questions. Both experimental sets were asked in the standard CASI module toward the end of the survey.

For the first, the fully-randomised conjoint experiments consist of respondents ranking or rating two hypothetical choices, in this case applicant profiles, to examine which applicants would be granted residency rights based on varying specific characteristics. These characteristics include health condition, number of children in state schools, housing situation, country of origin, employment history and gender. Three comparison questions were made, asking for which of two applicants are preferred to be allowed to remain, with each set then followed by rating on a 1 to 10 scale on the belief that each applicant should be allowed to remain.

The second uses a vignette where respondents were provided with a hypothetical applicant who uses public services, but varies on the location of residence of the applicant based on existing knowledge about where respondents live. The controlling variables for these experiments are as follows.

The controlling variable for the vignette experiment is:

cjitown on record **k_indresp_ip** (Applicant town of residence)

- 1 TOWN OF RESIDENCE from ff_post_code
- 2 Glasgow

As there is a large number of controlling variables for the conjoint experiment, they are listed in an appendix. See: **Appendix A: Controlling variables for conjoint preference experiment (9.26).**

The variables containing respondents' ratings and choices are on record **k_indresp_ip**: **cjimmig1, cji1app1, cji1app2, cjimmig2, cji2app1, cji2app2, cjimmig3, cji3app1, cji3app2, cjivig**

9.27 Reporting of height and weight

In Wave 12 data on height and weight of respondents were collected, in order to calculate body mass index, a measure of adiposity. The aim of this experiment was to determine if the interview mode affect the accuracy of individuals' responses on body weight and height.

Households in face-to-face modes (nurse and interviewer modes only, see the Section Mixed mode experiments: web and face-to-face, were randomly assigned to one of two

groups for the collection of self-reported height and body weight: self-completion mode group and interviewer-administered group.

Nurse and interviewers assessed height and weight using the [fieldwork protocols](#) previously established in Understanding Society. Briefly height without shoes was measured using a portable stadiometer. Measurements taken with the respondents stretching to the maximum height and the head in the Frankfort plane. Weight was measured using a portable electronic scale. Respondents were asked to remove their shoes and any bulky clothing.

All web respondents were asked height and weight in the same way as those in the self-completion mode. Other than the mode, the questions were the same.

Allocation to this experiment was done at the household level; households were allocated randomly and equally to conditions (50% each). The variable controlling the treatment allocations is **I_ff_height on record I_hhsamp_ip**:

- 1 Self-reported height and weight in CAPI
- 2 Self-reported height and weight in CASI

The variables affected by this experiment are on record **I_indresp_ip**:

I_hlht, I_hlhtf, I_hlhti, I_hlhtc, I_hlwt, I_hlwts, I_hlwtp, I_hlwtk, I_hlwte, I_hlwtl

9.28 Biomarker and sample collection

Wave 12 collected a number of biomarkers throughout the survey to examine biomarker collection and determine the impact of mode of completion. Which biomarkers were collected and how was determined by mode of completion. (See the Section Mixed mode experiments: web and face-to-face on the allocation to nurse, interviewer or web survey in Wave 12.) The exception is self-collected blood pressure described above – all respondents regardless of web were asked to complete these measures in the same way.

The following outlines the biomarker data and samples collected by mode of data collection.

Face-to-Face (nurse, interviewer)

1. Blood pressure. In the interview, both nurses and interviewers asked respondents for consent to take blood pressure readings. Blood pressure cuffs were provided to nurses and interviewers. For those consenting, three measures of blood pressure were taken. Feedback on the status of the blood pressure was given to respondents.
2. Height and weight. In the interview, both nurses and interviewers asked respondents to consent to take actual height and weight measures using a stadiometer and scale.

Nurse

In addition to those reported above, nurses also asked respondents to give

1. Whole blood sample. Drawn intravenously using several vials.
2. Dried blood spots. While the nurse helped lance the finger and place the blood spots, respondents were asked to send back the dried blood spot sheet after the interview.
3. Hair sample. The nurse cut and sent back a hair sample to test for hormones.

Interviewer

In addition to blood pressure, height and weight assessments, interviewers asked respondents to give

1. Dried Blood Spots
2. Hair Sample. Interviewers left a kit for respondents to take dried blood spot and a hair sample, both of which were to be sent back by the respondent after completion. The interviewer introduced the kits and provided explanations of the tests and the reasons for the requests, as well as pointing out instructions for completion.

Web

In addition to blood pressure, respondents were asked to collect

1. Dried Blood Spots
2. Hair Sample. For the dried blood sample and hair sample collection, web respondents were asked in interview if they were willing to receive these kits, with a short explanation of the tests and reasons for the request. For those accepting the request, the kits were shipped to their household, with instructions on how to complete and return the in-home sample.

Dried blood samples and Lab Closures

The closure of labs during the COVID-19 pandemic resulted in the dried blood samples being analysed at different times. The measurements that were made after the lab closure are impacted by that closure. To identify these the indicator variable **late_lab_batches** in **I_indresp_ip** shows whether the samples were processed before or after the lab closures. A value of 1 indicates they were analysed after the lab closures.

Analytes from hair samples:

- Cortisol (pg/mg)
- Cortisone (pg/mg)
- Testosterone (pg/mg)
- Progesterone (pg/mg)
- Dehydroandrosterone (pg/mg)

The following variables have a value of -8 or 1 (inapplicable or Yes). However respondents in the hair file do have lab results if consent was received.

- **I_haircut** take hair sample
- **I_hrconsc** explain hair consent
- **I_hairwill** hair sample consent

The variable **I_haircut** shows these inapplicable responses as a result of the hair being collected by the respondents themselves instead of a nurse.

Analytes from Blood samples:

These variables are derived from full blood samples

- **I_fb_trig** full blood triglycerides (unfasted) mmol/l
- **I_fb_chol** full blood cholesterol (total) mmol/l
- **I_fb_hdl** full blood high-density lipoprotein cholesterol mmol/l
- **I_fb_hba1c** full blood glycated hemoglobin (ifcc standardised) mmol/mol hb
- **I_fb_hba1c_corrected** full blood glycated hemoglobin (ifcc standardised) mmol/mol hb - corrected

These variables are derived from dried blood spot

- **I_dbs_trig** dried blood triglycerides (unfasted) mmol/l
- **I_dbs_chol** dried blood cholesterol (total) mmol/l
- **I_dbs_hdl** dried blood high-density lipoprotein cholesterol mmol/l
- **I_dbs_hba1c** dried blood glycated hemoglobin (ifcc standardised) mmol/mol hb

The variables used for this experiment are the following.

On record **I_indresp_ip**:

I_chbnd, I_bmcon, I_clota, I_clotb, I_fit, I_mastc, I_renald, I_dbswill, I_refbsc1, I_refbsc2, I_refbsc3, I_refbsc4, I_refbsc5, I_refbsc6, I_refbsc7, I_refbsc97, I_dbsconsc, I_dbssamptak, I_dbshand, I_dbsfingers1, I_dbsfingers2, I_dbsfingers3, I_dbsfingers4, I_dbsfingers5, I_dbstime, I_dbsvpalco, I_prick, I_attempts, I_dbsdifc1, I_dbsdifc2, I_dbsdifc3, I_dbsdifc4, I_dbsdifc5, I_dbsdifc6, I_dbsdifc97, I_nodbsm1, I_nodbsm2, I_nodbsm3, I_nodbsm4, I_nodbsm97, I_fullblwill, I_reffullbl1, I_reffullbl2, I_reffullbl3, I_reffullbl4, I_reffullbl5, I_reffullbl6, I_reffullbl7, I_reffullbl97, I_fbsconsc, I_sampf1, I_sampf2, I_sampf3, I_sampf4, I_sampf5, I_fbsamptak, I_faint1, I_samparm, I_samdifc1, I_samdifc2, I_samdifc3, I_samdifc4, I_samdifc5, I_samdifc6, I_samdifc97, I_nofbsm1, I_nofbsm2, I_nofbsm3, I_nofbsm97, I_vpsys, I_vpalco, I_vppress1, I_vppress2, I_vppress3, I_vpplaster, I_vpprob1, I_vpprob2, I_vpprob3, I_vpprob97, I_vpprob96, I_debslfbp, I_debbp, I_debbm, I_debhair, I_debdb, I_debfb, I_webkit, I_webkitkid, I_kit1, I_drcons, I_bmnores_code, I_bpnores_code, I_dbsnores_code, I_fbnores_code, I_nodbsmoth_code, I_noslfbpoth_code, I_nxtendoth_code, I_othbdif_code, I_othdifbp_code, I_othfullbl_code, I_othnbp_code, I_othrefdb, I_othreffbs_code, I_scrx_code, I_scux_code, I_selfbonores_code, I_slfbplocoth_code, I_slfbpprres_code, I_unreloth_code, I_unrelwgt_code, I_vpprobothr_code, I_bpcons, I_con30sb1, I_con30sb2, I_con30sb3, I_con30sb4, I_con30sb96, I_omronno, I_cufsize, I_airtemp, I_systolic1, I_diastolic1, I_pulse1, I_systolic2, I_diastolic2, I_pulse2, I_systolic3, I_diastolic3, I_pulse3, I_ynobp,

l_resbpd1, l_resbpd2, l_nattbpd1, l_nattbpd2, l_nattbpd3, l_nattbpd4, l_nattbpd5,
l_nattbpd97, l_difbpc1, l_difbpc2, l_difbpc3, l_difbpc4, l_difbpc5, l_difbpc6, l_difbpc97,
l_bpfeedc, l_bprisk, l_resphgt, l_stdhgt1, l_relhgt, l_unrelhgt1, l_unrelhgt2, l_unrelhgt3,
l_unrelhgt4, l_unrelhgt5, l_unrelhgt97, l_hgtref1, l_hgtref2, l_hgtref3, l_hgtref4, l_hgtref5,
l_hgtref6, l_hgtref97, l_nohgtob1, l_nohgtob2, l_nohgtob3, l_nohgtob4, l_nohgtob5,
l_nohgtob97, l_rwgtob, l_respwgt, l_rwgt1, l_floorc, l_relwgt, l_reswtref1, l_reswtref2,
l_reswtref3, l_reswtref4, l_reswtref5, l_reswtref6, l_reswtref97, l_nowtob1, l_nowtob2,
l_nowtob3, l_nowtob4, l_nowtob5, l_nowtob6, l_nowtob7, l_nowtob8, l_nowtob97,
l_fb_trig, l_fb_chol, l_fb_hdl, l_fb_hba1c, l_fb_hba1c_corrected, l_dbs_trig, l_dbs_chol,
l_dbs_hdl, l_dbs_hba1c, l_dbsvalidconsent, l_fbvalidconsent, l_fb_sample_dv,
l_dbs_sample_dv

On record l_hair_ip:

l_hairlen, l_headshk, l_hairwill, l_hrconsc, l_haircut, l_haircol, l_hairdye, l_dyetime,
l_dyetype, l_orighair, l_hairtrea, l_hairtrea2, l_hairwet, l_hairstyle, l_stylefreq, l_hairnosa,
l_haircloth_code, l_hairnores_code, l_hairtrea3_code, l_orighairoth_code, l_stylewhat,
l_hairmass, l_hair_cortisol, l_hair_cortisone, l_hair_testosterone, l_hair_progesterone,
l_dhea, l_hairsamplestatus_coded, l_hairsamplecomment_coded, l_hairvalidconsent,
l_childpno, l_hairkid, l_hairkidcode, l_haircutkid, l_nohaircutkid, l_kitchild

On record l_experience_ip:

l_version, l_timespent, l_studyintrst, l_futstdylik, l_slfbpease, l_kitarr7d, l_kithowarr,
l_kitcond, l_instrcteasy, l_taksmpeasy, l_samptoenlv, l_clearsndbck, l_notprvdsamp,
l_notconsent, l_bsease, l_bsretease, l_bsfultr, l_bsofffdbk, l_hairease, l_hairretease,
l_hairfultr, l_hairnurse, l_bsnurse, l_kitself, l_moresample, l_discussfdbk, l_signature,
l_contactnum

Questions asked in the main interview are in **l_indresp**. For each measure there are set of variables about why measures weren't taken. They generally fall into these three groups, participant reasons, mechanical reasons and problems affecting the sample:

Blood pressure

- **l_bpcons** blood pressure consent
- **l_ynobp** reason no bp reading
- **l_nattbpd1** problems with pc
- **l_nattbpd2** respondent upset/anxious/nervous
- **l_nattbpd3** error reading
- **l_nattbpd4** problems with cuff fitting/painful
- **l_nattbpd5** problems with equipment (not error reading)
- **l_nattbpd97** other reason (please specify)
- **l_difbpc1** no problems taking blood pressure
- **l_difbpc2** reading taken on left arm because right arm not suitable
- **l_difbpc3** respondent was upset/anxious/nervous
- **l_difbpc4** problems with cuff fitting/painful

- **I_difbpc5** problems with equipment (not error reading)
- **I_difbpc6** error reading
- **I_difbpc97** other problems or departures from protocol (please specify)

Height

- **I_resphgt** standing height intro
- **I_hgtref1** cannot see point / height already known / doctor has measurement
- **I_hgtref2** too busy / taken too long already / no time
- **I_hgtref3** respondent too ill / frail / tired
- **I_hgtref4** considered intrusive information
- **I_hgtref5** respondent too anxious / nervous / shy / embarrassed
- **I_hgtref6** refused (no other reason given)
- **I_hgtref97** other, please specify
- **I_nohgtob1** respondent is unsteady on feet
- **I_nohgtob2** respondent cannot stand upright / too stooped
- **I_nohgtob3** respondent unable to remove shoes
- **I_nohgtob4** ill or in pain
- **I_nohgtob5** stadiometer faulty or not available
- **I_nohgtob97** other, please specify

Weight

- **I_rwgtob** respondent weight observation
- **I_respwgt** respondent weight intro
- **I_reswtref1** cannot see point/weight already known/doctor has measurement
- **I_reswtref2** too busy/taken long enough already/no time
- **I_reswtref3** respondent too ill/frail/tired
- **I_reswtref4** considered intrusive information
- **I_reswtref5** respondent too anxious/nervous/shy/embarrassed
- **I_reswtref6** refused (no other reason given)
- **I_reswtref97** other, please specify
- **I_nowtob1** respondent is unsteady on feet
- **I_nowtob2** respondent cannot stand upright
- **I_nowtob3** respondent is chairbound
- **I_nowtob4** confined to bed
- **I_nowtob5** respondent unable to remove shoes
- **I_nowtob6** respondent weighs more than 198 kg
- **I_nowtob7** ill or in pain
- **I_nowtob8** scales not working
- **I_nowtob97** other, please specify

Hair

- **I_hairnosa** no hair sample obtained
- **I_nohaircutkid** reason no child hair sample

Dried blood spot (DBS)

- **I_clota** clotting disorder
- **I_clotb** anticoagulants
- **I_fit** have a fit
- **I_mastc** mastectomy
- **I_renald** renal dialysis
- **I_dbswill** participant willing to have a dried blood spot sample taken by nurse
- **I_refbsc1** dislike/fear of giving a blood sample
- **I_refbsc2** respondent felt it would be unsafe
- **I_refbsc3** respondent recently had blood test
- **I_refbsc4** refused because of current illness
- **I_refbsc5** worried about h.i.v. or aids
- **I_refbsc6** not receiving feedback on results
- **I_refbsc7** refusal - no other reason given
- **I_refbsc97** other please specify
- **I_dbsdifc1** no problems
- **I_dbsdifc2** unable to obtain enough blood
- **I_dbsdifc3** difficulty stopping bleeding
- **I_dbsdifc4** respondent became light-headed/faint/nauseous
- **I_dbsdifc5** respondent fainted
- **I_dbsdifc6** equipment problems
- **I_dbsdifc97** other (specify at next question)
- **I_nodbsm1** no suitable finger or thumb
- **I_nodbsm2** respondent was too anxious/nervous
- **I_nodbsm3** respondent felt light-headed/faint/nauseous
- **I_nodbsm4** respondent fainted
- **I_nodbsm97** other (specify at next question)

Full blood

- **I_fullblwill** participant willing to have a full blood sample taken by nurse
- **I_reffullbl1** previous difficulties with venepuncture
- **I_reffullbl2** dislike/fear of needles
- **I_reffullbl3** respondent recently had blood test/health check
- **I_reffullbl4** refused because of current illness

- **I_reffullbl5** worried about h.i.v. or aids
- **I_reffullbl6** not receiving feedback on results
- **I_reffullbl7** no information about what blood will be tested for
- **I_reffullbl97** other (specify at next question)
- **I_fbsconsc** signed full blood consent obtained by nurse
- **I_faint1** respondent feeling faint
- **I_samdifc1** no problem
- **I_samdifc2** incomplete sample
- **I_samdifc3** collapsing/poor veins
- **I_samdifc4** second attempt necessary
- **I_samdifc5** some blood obtained but respondent felt faint/fainted
- **I_samdifc6** unable to use tourniquet
- **I_samdifc97** other (specify at next question)
- **I_nofbsm1** no suitable or no palpable vein/collapsed veins
- **I_nofbsm2** respondent was too anxious/nervous
- **I_nofbsm3** respondent felt faint/fainted
- **I_nofbsm97** other (specify at next question)

9.29 LinkedIn consent

This experiment tested different ways of asking respondents for consent to link data from their LinkedIn profiles to their Understanding Society survey data. LinkedIn is a platform that serves mainly as a professional networking site, including allowing individuals to search for jobs.

This experiment included two aspects. The first experimental condition tested the impact of placing this consent question early or late in the questionnaire. The second condition tested the impact of including a motivational statement prior to the consent request. Half of respondents received a motivational statement similar to those used to reduce item missingness on other survey questions (Al Baghal and Lynn 2015) and now regularly used in Understanding Society. The statement was placed before the consent question: “The data this would provide is key to this study. This data will enhance the understanding of your survey responses.” The control group received the consent question without this statement.

For the ongoing Wave 14 sample (identified by **hhorig<19**), the allocations to treatments were generated at the individual level by the questionnaire script. The experiments are controlled by the following variables on the record **n_indresp_ip**:

n_lkincon (1/2 each):

- 1 LinkedIn consent is early
- 2 LinkedIn consent is late

n_linkstat (1/2 each):

- 1 LinkedIn consent statement included
- 2 LinkedIn consent statement not included

For the wave 14 refreshment sample the allocations were however made at the household level before the start of fieldwork. This was done to test the differences in experiment variances based on whether the allocations occur at the household-level (so everyone in the same household receives the same treatment) or at the individual-level (so treatment is allocated in during the survey to each individual independent of household) impact outcomes, including precision, differently. These allocations are recorded in the following variables on record **n_hhsamp_ip**:

n_ff_lkincon (1/2 each, allocation stratified by **sampleorig**, **ff_incentw14**, **ff_gridmodew11**, **l_hhmodes**):

- 1 LinkedIn consent is early
- 2 LinkedIn consent is late

n_ff_lkinstat (1/2 each, allocation stratified by **ff_lkincon**, **sampleorig**, **ff_incentw14**, **ff_gridmodew11**, **l_hhmodes**):

- 1 Additional consent statement included
- 2 Additional consent statement excluded

The variables affected by these experiments are on record **n_indresp_ip**:

lkinhavear, **lkinlinkear**, **lkinurlear**, **lkinurleardk**, **lkinforenear**, **lkinsurnear**, **lkinempear**, **lkintitear**, **lkinlocear**, **lkinscheear**, **lkinhavlat**, **lkinlinklat**, **lkinurllat**, **lkinurllatdk**, **lkinforenlat**, **lkinsurnlat**, **lkinemplat**, **lkintitlat**, **lkinloclat**, **lkinschlat**

9.30 Vignettes: intentions to prepare for automation

This experiment tested how what is communicated about the impact of automation affects the extent to which individuals are concerned about automation, are interested in re/upskilling to prepare for automation and intend to undertake actions accordingly. To address this question, a 2×2 experiment was conducted. Both the effect of the prospected severity of technological developments (severe threat vs. minor threat) and the prospected timeframe of such developments (short-term vs. long-term) on individuals' re/upskilling attitudes and behaviours were tested. The experiment was introduced in a vignette about the effect and timeframe of automation, where the text of the vignette was varied experimentally. Following presentation of the vignette, a series of follow-up questions about re/upskilling were asked. Additional questions about the respondent's current job skills and security were included before the vignette for controls. Only employed respondents were included in this experiment. The two dimensions (effect and timeframe) and a control text were combined to be controlled by one variable. The allocation was made at the household level, with 1/5 of households allocated to each group. The treatment allocation variables is **n_ff_techworkw14** on record **n_hhsamp_ip** (allocation stratified by **ff_incentw14**, **sampleorig**, **ff_gridmodew13**, **l_hhmodes**):

- 1 Major severity; long-term
- 2 Major severity; short-term
- 3 Minor severity; long-term
- 4 Minor severity; short-term
- 5 Control group

The variables affected by this experiment are on record **n_indresp_ip**:

datawork, algorithmwork, robotwork, losejob, losejoblong, techreplace, techdoes, techopp, techimport, techoutdate, techskills, techtrain, techtime, techchange

9.31 Comparison of ReQoL-10 to other scales

This experiment compared the 10-item Recovering Quality of Life questionnaire (ReQoL-10) to other health quality of life scales: the instrument (EQ-5D-3L) currently preferred by the National Institute for Health and Clinical Excellence (NICE) and the 12-item General Health Questionnaire (GHQ). To test for ordering effects, half of respondents were presented with ReQoL-10 and EQ-5D-3L earlier in the survey and the GHQ later on, with the other half receiving the reverse, with the GHQ coming earlier in the survey and ReQoL-10 and EQ-5D-3L later. For the continuing wave 14 sample (identified by **hhorig<19**), the treatment allocations were made within the questionnaire script, with half the sample allocated to each group. The allocations are recorded in the variable **n_req10cond** on record **n_indresp_ip**:

- 1 Ask ReQoL and EQ5D early and GHQ late
- 2 Ask ReQoL and EQ5D late and GHQ early

For the wave 14 refreshment sample (identified **n_hhorig=19**) the allocations were made prior to fieldwork at the household level, again with half the sample allocated to each group. This was done to test the differences in experiment variances based on whether the allocations occur at the household-level (so everyone in the same household receives the same treatment) or at the individual-level (so treatment is allocated in during the survey to each individual independent of household) impact outcomes, including precision, differently. The allocations are recorded in the variable **n_ff_req10cond** on record **n_hhsamp_ip** (allocation stratified by **sampleorig, ff_incentw14, ff_gridmodew13, l_gor_dv**):

- 1 ReQoL and EQ5D early and GHQ late
- 2 ReQoL and EQ5D late and GHQ early

The variables affected by this experiment are on the record **n_indresp_ip**: **reqol1ear, reqol2ear, reqol3ear, reqol4ear, reqol5ear, reqol6ear, reqol7ear, reqol8ear, reqol9ear, reqol10ear, reqphysear, eq3lmobear, eq3lcareear, eq3lactear, eq3lpainear, eq3lanxear, scghqaear, scghqbear, scghqcear, scghqdear, scghqeear, scghqfear, scghqgear, scghqhear, scghqiear, scghqjear, scghqkear, scghqlear, reqol1lat, reqol2lat, reqol3lat, reqol4lat, reqol5lat, reqol6lat, reqol7lat, reqol8lat, reqol9lat, reqol10lat, reqphyslat, eq3lmoblat,**

eq3lcarelat, eq3lactlat, eq3lpainlat, eq3lanxlat, scghqalat, scghqblat, scghqclat, scghqdlat, scghqelat, scghqflat, scghqglat, scghqhlat, scghqilat, scghqjlat, scghqklat, scghqlat

9.32 Proxy nomination

This experiment tested different ways of asking respondents to nominate a proxy who could complete a short questionnaire on their behalf in future waves, if they themselves were unable to complete it or moved into an institution. The aim was to seek for the best way to identify who the alternative person to ask is and ensure we have permission to ask about the respondent in future waves.

The experiment tested two different versions of this initial consent question. Respondents who consented were then asked a number of follow-up questions collecting information about the person nominated. These follow-up questions were based on the Stable Contacts module. The experiment also tested the effects of rotating the order of the Stable Contacts and Proxy Nomination modules, where one comes first and the other follows immediately after. The two parts of the experiment are controlled by two variables. For the wave 14 continuing sample (identified by **hhorig**<19) these were generated within the questionnaire script. These controlling variables are on the record **n_indresp_ip**:

n_prxver (1/2 each):

- 1 Proxy nomination question is version 1
- 2 Proxy nomination question is version 2

n_prxear (1/2 each):

- 1 Proxy nomination is early
- 2 Proxy nomination is late

For the Wave 14 refreshment sample (identified by **n_hhorig**=19), the allocations were made at the household level prior to fieldwork. This was done to test the differences in experiment variances based on whether the allocations occur at the household-level (so everyone in the same household receives the same treatment) or at the individual-level (so treatment is allocated in during the survey to each individual independent of household) impact outcomes, including precision, differently. The allocations are recorded in the following variables on record **n_hhsamp_ip**:

n_ff_prxver (1/2 each, allocation stratified by **sampleorig**, **ff_incentw14**, **ff_gridmodew11**):

- 1 Proxy nomination question is version 1
- 2 Proxy nomination question is version 2

n_ff_prxear (1/2 each allocation stratified by **ff_prxver**, **sampleorig**, **ff_incentw14**, **ff_gridmodew11**):

- 1 Proxy nomination before Stable Contacts
- 2 Proxy nomination after Stable Contacts

The variables affected by this experiment are on record **n_indresp_ip**: **prxconv1ear, prxconv2ear, prxnomear, prxrelear, prxconv1lat, prxconv2lat, prxnomlat, prxrellat**. The variables containing the contact details are not released with the data (Questions **prxnameear, prxadd1ear, prxadd2ear, prxtownear, prxcntyyear, prxpcodeear, prxtel1ear, prxtel2ear, prxemlear, prxnamelat, prxadd1lat, prxadd2lat, prxtownlat, prxcntylat, prxpcodelat, prxtel1lat, prxtel2lat, prxemllat**).

9.33 Twitter consent

This module repeats the consent asked (non-experimentally) at IP10 (see 10.5 Consent to link Twitter data) and aims to look at the feasibilities and practicalities of linking social media (in particular Twitter) and survey data in a longitudinal context, and how they can be combined to improve the quality of both.

The IP10 consent questions were updated using cognitive interviewing techniques.

In addition to wording updates, an experiment on additional consent information placement was included. Eight help links with additional information were devised for this question. Half of respondents were presented with these links between the text and immediately before the consent question, all on one page. The other half were first asked to indicate whether they wanted more information in the response options before deciding, and if so the eight help links were presented on a separated page with the consent question.

Respondents were asked about whether they use Twitter and, if so, their willingness to link their Twitter account to their survey responses. The main goal of the research was to obtain consent to linkage. Given consent had previously been sought at IP10, this module allows for testing in change of consent for those asked previously, as well as asking consent for the first time of IP11 and IP14 refreshment samples.

The controlling variable, equally allocated at the household-level, is on record **o_hhsamp_ip**:

ff_twithelpw15 (1/2 each, allocation stratified by sampleorig ff_incentw15
ff_gridmodew15 l_hhmodes)

- 1 = Help links on same page
- 2 = Help links on separate page

The questions used for this experiment are in file **o_indresp_ip**:

twithave, twitlinka, twitlinkb, twitlinkinfo

9.34 Marginal propensity to consume

A large and influential literature aims to measure consumers' spending, saving and debt responses to income changes using questions of this kind. These studies aim to estimate how much consumers would spend out of each £1 change in income. Average estimates of

this parameter vary hugely across studies in ways that are unlikely to be due to differences in individuals' economic environments. They may instead reflect differences in the scenarios put to consumers, or in the way questions are worded.

This experiment was designed to explain the large differences in findings across studies through varying methods of eliciting individuals' spending responses to hypothetical income windfalls. Innovation Panel respondents were assigned to different conditions asking about how they would respond if they were to receive a one-time payment of either £500 or £2500. The condition of amount was crossed with a second condition, varying the duration that the windfall payment would be spent over, either 3 or 12 months. For both conditions, respondents were equally allocated to conditions. Respondents were asked if they would spend more, less or the same based on the amount and duration, and if more or less, the amount more/less.

The controlling variables, which were crossed and equally allocated at the household-level, are on record **o_hhsamp_ip**.

ff_mpctreatment (1/2 each, allocation stratified by ff_mpcduration
ff_mpcamount sampleorig ff_gridmodew15)

1 = Two-part question on marginal propensity to consume
2 = Direct question on marginal propensity to consume

ff_mpcduration (1/2 each, allocation stratified by ff_incentw15 sampleorig
ff_gridmodew15)

1 = 3 months
2 = 12 months

ff_mpcamount (1/2 each, allocation stratified by ff_mpcduration
ff_mpcamount sampleorig ff_gridmodew15)

1 = £500
2 = £2500

The variables used for this experiment are in the file **o_indresp_ip**:

t1mpc1, t1mpc2inc, t1mpc2red, t1mpc3same, t2mpc1500, t2mpc12500, mpcsustain, save, saved, t2mpc21, t2mpc22, t2mpc23, t2mpc24

9.35 Informal care

Measurement of informal care is problematic for a number of reasons. Caregiving is a complex concept and ontologically vague. Linguistically speaking, 'care' can refer to emotions and/or practices and, in common parlance, 'caregiving' tends to embody a nebulous set of activities. Informal care is often defined as unpaid, non-contractual or that which happens outside of formal employment, which may seem straightforward, but this can also pose measurement challenges where boundaries of paid/unpaid, job/non-job-

related are blurred. In addition to the lack of a clear definition of what constitutes informal care, public perceptions and personal identities shape whether individuals recognise themselves as unpaid carers and/or recipients.

Three different sets of questions were used to identify informal caregivers to identify the best set to use in Understanding Society going forward.

One set of questions was the current Caring module and represents no additional questionnaire time, as it is core content, used as the control set of questions.

Another set was a revised version of the Caring module, with amendments to wording to better reference all types of caring.

The third set was a combination of existing questions in the Disability module and the ADL/IADL questions in the Social Care module. This condition is designed to ask comparable questions of co-residents: recipients and providers. For receipt, individuals would get the ADL/IADL questions if they have any health problems or if aged 65 or over. For provision, individuals would be asked a tailored question on health problems to see if anyone in their household has difficulties. If so, questions follow about what support they give, if any, in terms of ADLS/IADLS. That way the questions on both sides were broadly the same.

The controlling variable, determining which of the three sets of questions respondents received, equally and randomly allocated at the household-level, is on record **o_hhsamp_ip**:

ff_infcarew15 (1/3 each, allocation stratified by sampleorig ff_incentw15
ff_gridmodew15 ff_mpcduration)

- 1 = Amended Caring questions
- 2 = Combining comparable questions
- 3 = Control

The variables affected by this experiment are in the **o_indresp_ip** module:

aidhh, aidxhh, naidxhh, aidhu1, aidhu2, aidhrs, aideft, casch, aida, naidxhha, aidhu1a, aidhu2a, aidhrsa, aidefta, cascha, naidxhhb, aidhu1b, aidhu2b, aidhrsb, aideftb, caschb, aidhua1, aidhua2, aidhua3, aidhua4, aidhua5, aidhua6, aidhua7, aidhua8, aidhua9, aidhua10, aidhua11, aidhua12, aidhua13, aidhua14, aidhua15, aidhua16, aidhuaa1, aidhuaa2, aidhuaa3, aidhuaa4, aidhuaa5, aidhuaa6, aidhuaa7, aidhuaa8, aidhuaa9, aidhuaa10, aidhuaa11, aidhuaa12, aidhuaa13, aidhuaa14, aidhuaa15, aidhuaa16, impdishh1, impdishh2, impdishh3, impdishh4, impdishh5, impdishh6, impdishh7, impdishh8, impdishh9, impdishh10, impdishh11, impdishh96, impdishh97, impdiswho1, impdiswho2, impdiswho3, impdiswho4, mpdiswho5, impdiswho6, impdiswho7, impdiswho8, impdiswho9, impdiswho10, impdiswho11, impdiswho12, impdiswho13, impdiswho14, impdiswho15, impdiswho16, caretaskhh1, caretaskhh1_1, caretaskhh1_2, caretaskhh1_3, caretaskhh1_4, caretaskhh1_5, caretaskhh2, caretaskhh2_1, caretaskhh2_2, caretaskhh2_3, caretaskhh2_4, caretaskhh2_5, caretaskhh3, caretaskhh3_1, caretaskhh3_2, caretaskhh3_3, caretaskhh3_4, caretaskhh3_5, caretaskhh4, caretaskhh4_1, caretaskhh4_2, caretaskhh4_3, caretaskhh4_4, caretaskhh4_5, caretaskhh5, caretaskhh5_1, caretaskhh5_2, caretaskhh5_3,

caretaskhh5_4, caretaskhh5_5, caretaskhh6, caretaskhh6_1, caretaskhh6_2, caretaskhh6_3, caretaskhh6_4, caretaskhh6_5, caretaskhh7, caretaskhh7_1, caretaskhh7_2, caretaskhh7_3, caretaskhh7_4, caretaskhh7_5, caretaskhh8, caretaskhh8_1, caretaskhh8_2, caretaskhh8_3, caretaskhh8_4, caretaskhh8_5, caretaskhh9, caretaskhh9_1, caretaskhh9_2, caretaskhh9_3, caretaskhh9_4, caretaskhh9_5, caretaskhh10, caretaskhh10_1, caretaskhh10_2, caretaskhh10_3, caretaskhh10_4, caretaskhh10_5, caretaskhh11, caretaskhh11_1, caretaskhh11_2, caretaskhh11_3, caretaskhh11_4, caretaskhh11_5, caretaskhh12, caretaskhh12_1, caretaskhh12_2, caretaskhh12_3, caretaskhh12_4, caretaskhh12_5, caretaskhh13, caretaskhh13_1, caretaskhh13_2, caretaskhh13_3, caretaskhh13_4, caretaskhh13_5, caretaskhh14, caretaskhh14_1, caretaskhh14_2, caretaskhh14_3, caretaskhh14_4, caretaskhh14_5, caretasknonhh1, caretasknonhh2, caretasknonhh3, caretasknonhh4, caretasknonhh96, caretasknonhh97

9.36 Alcohol consumption

A review of Understanding Society questionnaires identified our usage of the Audit-C alcohol consumption questions differed somewhat from the questions the NHS uses for its measures. They differ in how the surveys ask about alcohol in either drinks or units, with the latter being a specific amount of alcohol. Understanding Society surveys have asked those who have had a drink in the past 12 months about how often they have a drink with alcohol, how many drinks they have on day they have alcohol, and how often they have 6+ (for women) or 8+ units (for men) of alcohol on days they are drinking, a drinks-drinks-units pattern.

The NHS asks similar questions but asks for them in terms of drinks-units-units.

Given people may not think in terms of units, drinks may be more intuitive and conform to how respondents formulate responses. One purpose of this experiment is to determine whether asking about drinks or units alters responses significantly. To address this question as clearly as possible, the NHS version was compared to a version asking these same questions in a drinks-drinks-drinks pattern.

For respondents completing the survey with a face-to-face interviewer, this module was asked in self-completion module, to minimise social desirability bias.

The controlling variable, determining which of the two sets of questions respondents received, equally and randomly allocated at the household-level, is on record **o_hhsamp_ip**:

ff_alcoholw15 (1/2each, allocation stratified by sampleorig ff_gridmodew15 ff_incentw15 ff_infcarew15)

1 = NHS version

2 = Drinks-only version

The variables affected by this experiment are in the file **o_indresp_ip**:

auditc1, auditc2, auditc3, auditc4a, auditc5a, auditc4b, auditc5b

9.37 Body Volume Index app and body measurements

IP15 was used to test different methods of obtaining body measurements. First, all respondents were sent a tape measure with instructions to measure their waist and hip circumference, as well as to self-report their height and weight. Second, respondents interviewed face-to-face were asked to repeat the waist and hip measurement during the interview, with instructions from the interviewer. Third, all respondents were invited to install a mobile app on their device and use it to upload two photos of themselves. The app used the images as well as profile information entered by the respondent (age, sex, height, weight, level of activity) to calculate different body measures, including body fat percentage, visceral fat content, and waist/hip ratio.

In the initial invitation letter sent before the survey, all respondents were sent a metric-only tape measure (150cm long). The mailing included instructions and a request to do a self-measurement of their waist and hips, and a card for respondents to record their dimensions. During the interview, respondents were asked if they had completed these measures and, if so, to report what these were. If a high proportion of respondents can provide these measures reliably, this adds to the feasibility of conducting biomarker collection via the web.

The reliability of these self-reported measures was tested through the in-interview measures of waist and hips led by interviewers. The interviewers did not physically touch the respondent, but provided the tape measures, led the respondent through the measurement, and visually inspected and recorded the results. Not only did these test interviewer-collection of these measures, they also serve as “objective” measures to compare to the self-reported waist and hip measures collected elsewhere in the interview. These interviewer-led measures are also part of the validation of the BVI app data collected.

Interviewers collected height and weight measures using stadiometers and scales, which was previously done both at W2/3 and IP12. The IP12 study showed interviewers were able to collect these measures using these tools. The IP15 measurement differs slightly in regards to the scales that were used. Regular scales collecting only weight were used at IP12; at IP15, the scales also provided measures of body fat and water percent. The body fat percent is directly comparable to the BVI estimate, and the standard BMI measure (captured by height and weight) can be more broadly used in comparison to the BVI measures in regards to outcomes.

These requests for these measures were not conducted experimentally, but rather as a pilot testing feasibility and quality, and to act as validation indicators for other measures collected in the study.

These questions were asked in the questionnaire modules **bodymasscapi_ip15**, **waistandhipcapi_ip15**, **selfwaistandhip_ip15**, **scaheightandweight_ip15**.

All IP15 respondents were asked to download the app during the interview; however, it was expected that respondents would use the app after the interview or otherwise in the absence of an interviewer. A variant of the BVI app, branded and customised specifically for use in Understanding Society, was created and made available on both Android and Apple app stores.

The invitation to the app study was based on previous research done in the IP asking respondents to download and use an app. After the invitation, several follow-up questions were asked to identify if a person downloaded the app successfully or reasons why they were not participating. Those respondents for whom we had a valid email address who did not refuse to participate in the app study, were sent an email reminder after the interview, reminding them of their login credentials and where to find the app.

The experimental components of this consent request were along two dimensions: the incentive given for using the app; and how feedback was referenced. For incentives, respondents received £5 either unconditionally or conditionally based on participation. The feedback experiment had conditions of whether feedback was referenced in regards to total body fat, feedback referenced in regards to visceral body fat, or no feedback was referenced.

The controlling variables for incentives and feedback, crossed and equally allocated at the household-level, are on record `o_hhsamp_ip`:

ff_bvfeedback_w15 (1/3 each, allocation stratified by `sampleorig`
`ff_gridmodew15` `ff_incentw15`)

- 1 = Feedback on total body fat
- 2 = Feedback on visceral body fat
- 3 = Feedback not mentioned

ff_bvincentive_w15 (1/2 each, allocation stratified by `ff_bvfeedback_w15`
`sampleorig` `ff_gridmodew15`)

- 1 = Unconditional £5 incentive
- 2 = Conditional £5 incentive

The variables related to the invitation to the Body Volume Index app are in the file `o_indresp_ip`:

osdetection, bviappoutc1, bviappoutc2, bviinsthow, bvicklnk, bviopencam, bviscanqr, bviopenas, bvifindapp, bviinst, bviopn, bviacode bvilogin, bviappemlhas, bvinotwhy1, bvinotwhy2, bvinotwhy3, bvinotwhy4, bvinotwhy5, bvinotwhy6, bvinotwhy7, bvinotwhy8, bvinotwhy9, bvinotwhy97

The data collected with the app are in the file `o_bviapp_ip`. Appendix B: Specification for the Body Volume Index app contains the specification of the app contents as well as screen shots of the app. The table below documents the variables in the file `o_bviapp_ip`.

Contents of the data file `o_bviapp_ip`

Source	Variable	Description
Identifier	<code>pidp</code>	Cross-wave person identifier
Timestamp	<code>date</code>	Date app measurement completed
Profile questions	<code>age</code>	Age
	<code>sex</code>	Sex. Values: 1 Male, 2 Female
	<code>height</code>	Height (cm)

	weight	Weight (kg)
	active	Activity level. Values: 1 Sedentary, 2 Active, 3 Fit, 4 Very fit
Body measurements	bfat	Body fat (%)
	vfat	Visceral fat (litres)
	whratio	Waist hip ratio
	wcirc	Waist circumference (cm)
	bmi	Body Mass Index (BMI)
	rsnktflr	Right side of the neck to the floor (cm)
	lsnktflr	Left side of the neck to the floor (cm)
	bonktflr	Back of the neck to the floor (cm)
	rmpshldrtflr	Right midpoint of the shoulder to the floor (cm)
	lmpshldrtflr	Left midpoint of the shoulder to the floor (cm)
	vertcrotfl	Vertical crotch to the floor (cm)
	linlgtflr	Left length of inside leg from crotch to floor (cm)
	rinlgtflr	Right length of inside leg from crotch to floor (cm)
	rhknetflr	Right height of the knee to the floor (cm)
	lhknerflr	Left height of the knee to the floor (cm)
	nkskinwst	From the neck along the skin to waist (cm)
	rwsttflr	Right side of the waist to the floor (cm)
	lwsttflr	Left side of the waist to the floor (cm)
	rlegtfllr	Right Leg length from the outside to the floor (cm)
	llegtfllr	Left Leg length from the outside to the floor (cm)
	fnklegbknk	Front of the neck through the legs to the back of the neck (cm)
	lcirctrnk	Left circumference of trunk from the top of the shoulder (cm)
	rcirctrnk	Right circumference of trunk from the top of the shoulder (cm)
	lengshldr	Length between the shoulders (cm)
	circneck	Circumference of neck at the base (cm)
	lengneck	Length between two sides of the neck (cm)
	rbicpcirc	Right bicep circumference (cm)
	lbicpcirc	Left bicep circumference (cm)
	relbcirc	Right elbow circumference (cm)
	lelbcirc	Left elbow circumference (cm)
	rfrarm	Right forearm circumference (cm)
	lfrarm	Left forearm circumference (cm)
	chstcirc	Chest circumference (cm)
	wstcincrw	Waist circumference at narrowest point (cm)
	propwst	Proportion of the waist in the front part of the abdomen
	wstwho	Waist circumference to WHO guidance (cm)
	propwstwho	Proportion of the WHO waist in the front part
	circhipwho	Circumference of the hips at the widest point to WHO guidance (cm)
	fpopwhohip	Front proportion of the WHO hip measurement
	rcirctgh	Right circumference of the thigh (cm)
	lcirctgh	Left circumference of the thigh (cm)
	rcirccknee	Right circumference of the knee (cm)
	lcirccknee	Left circumference of the knee (cm)

	rcfirc	Right calf circumference at the widest point (cm)
	lclfcirc	Left calf circumference at the widest point (cm)
	lclfcircm	circumference of calf at the mid point between knee and floor (cm)
	rcfircm	circumference of calf at the mid point between knee and floor (cm)
	rlengft	Right length of foot from base of the heel to the toe (cm)
	llengft	Left length of foot from base of the heel to the toe (cm)
Debrief questions	scan	Who took the photos. Values: 1 I did, 2 Someone else
	backgrd1	Background in photos: Furniture
	backgrd2	Background in photo: Door(s)
	backgrd3	Background in photos: Window(s)
	backgrd4	Background in photos: Picture(s)
	backgrd5	Background in photos: Mirrors
	backgrd6	Background in photos: Something else
	backgrd7	Background in photos: Nothing, empty wall
	wallclr	Colour of background wall. Values: 1 White or off-white, 2 Red, 3 Orange, 4 Yellow, 5 Green, 6 Blue, 7 Purple, 8 Grey, 9 Brown, 10 Black, 11 Multipole colours
	wallsd	Shade of background wall. Asked if wallclr is not 1. Values: 1 Light, 2 Medium, 3 Dark
	clothes	Type of clothing worn. Values: 1 Underwear, 2 Tight clothing, 3 Baggy clothing, 4 No clothes, 5 Mixture of tight and baggy clothing
	light	Lighting when photos were taken. Values: 1 Light, 2 Medium, 3 Dark

9.38 National identity

National identities inform behaviours and shape social relations, with evidence for changes both in majority and minority identities in recent years with respect to self-reported identity choice and the meaning people ascribe to it. Combining different identity measures and positioning of the questions might help better understand the meaning respondents ascribe to listed identity categories. This experiment aimed to study whether affective priming and question order of a set of national identity questions are associated with a change of identity choice and meaning.

Respondents were asked a multiple-choice national identity question and an affective Britishness question (importance of being British). Those respondents who chose “Other” as a national identity were also asked a follow-up question to list their other national identity. Half of respondents were asked the national identity question(s) first; the other half were asked the affective Britishness question first.

Respondents were asked follow-up questions on the importance of identity and their political views on Brexit and Scottish independence. These questions were adopted from questions used in previous waves of Understanding Society.

The controlling variable, equally allocated at the household level, is on record **o_hhsamp_ip**:

ff_identityw15 (1/2 each, allocation stratified by sampleorig ff_gridmodew15
ff_incentw15 ff_infcarew15)

1 = Importance of being British first
2 = National identity first

The variables affected by this experiment are in the file **o_indresp_ip**:

schlwtl, scbritida, scnatidoa_code, scnatidob_code, scbritidb, scwhorurac, vote1, vote2, vote4, vote4_all, vote3, vote3_all, eumempast, scotuk2, scnatida1, scnatida2 , scnatida3, scnatida4, scnatida5, scnatida6, scnatida97, scnatidb1, scnatidb2, scnatidb3, scnatidb4, scnatidb5, scnatidb6, scnatidb97

9.39 Vignettes: measuring flexibility stigma - double whammy or femininity stigma

This experiment uses a vignette design to examine what the rise in homeworking practices mean with regards to gender inequality in the labour market. Employed respondents between ages 18-65 were shown three vignettes, all with the same basic narrative describing candidates for a job based on their background and work circumstances. These candidates varied on five dimensions: 1) gender, 2) parenthood status, 3) home/hybrid working practice, 4) prevalence of hybrid working in the company, and 5) company policies.

After each of the three vignettes, respondents were asked their views on the worker. Respondents were asked if they would recommend the candidate for the job; the perception the candidate is committed to work; how productive the candidate is; and how much of a team-player the candidate is. All these questions were measured on a 0 to 10 scale. There was also a single manipulation check question asked at the end.

The allocations for the vignette dimensions were done within the survey script in equal proportions. For each dimension there is one controlling variable for each of the three vignettes in file **p_indresp_ip**:

fsvigname1

1 = Charlotte Davies
2 = Paul Taylor

fsvigname2

1 = Anna Wright
2 = Adam Newton

fsvigname3

1 = Sarah Walker

2 = James Harrison

fsvigparent1/ fsvigparent2/ fsvigparent3

1 = Has no children

2 = Has two children

fsvighybrid1/ fsvighybrid2/ fsvighybrid3

1 = Works in the office 5 days a week

2 = Has worked from home 1-2 days a week on a regular basis in the last year

3 = Has worked from home 3-4 days a week on a regular basis in the last year

fsvigorguse1/ fsvigorguse2/ fsvigorguse3

1 = Less than 20% of all workers

2 = More than half of all mothers

3 = More than half of all parents

4 = More than half of all workers

fsvigorgpol1/ fsvigorgpol2/ fsvigorgpol3

1 = Are no company level policies that allow for a mix of working from home and in the workplace.

2 = Are company level policies that allow for a mix of working from home and in the workplace but only for mothers.

3 = Are company level policies that allow for a mix of working from home and in the workplace but only for parents.

4 = Are company level policies that allow for a mix of working from home and in the workplace for all workers.

The variables affected by this experiment are in file **p_indresp_ip**:

fsvrecomm1, fsvrecomm2, fsvrecomm3, fsvcommit1, fsvcommit2, fsvcommit3, fsvproduct1, fsvproduct2, fsvproduct3, fsvteam1, fsvteam2, fsvteam3, fsvrecal

For initial findings, see Chung and Wang (2024).

9.40 Cognitive reflection and politically motivated reasoning

This experiment uses a variation of a known cognitive reflection test (CRT) to test whether politically motivated reasoning is a manifestation of heuristic or deliberative thinking. The CRT is a four-question battery asking cognitive ability-like questions (e.g. 'If you're running a race and you pass the person in second place, what place are you in?' Answer: 'first'). The CRT questions are not experimentally manipulated.

The political questions ask about a petition to require a verified identification to open a social media account. The experiment leverages people's identification with 'Leave' or 'Remain' following the Brexit vote. There are three versions of the requirement of the verified identification using these. One version says the "petition was popular in areas that voted Remain in the Brexit referendum", and the second version substitutes 'Remain' with

‘Leave’. The third version is the control, which does not contain any such phrase to prime respondents.

The allocations were done at the household level, such that all members of a household were allocated the same treatment, with 1/3 of households allocated to each group. The allocations were stratified by sample origin, wave 16 incentive treatment allocation, wave 16 mode allocation, and government office region. The allocation variable is **ff_polreasoningw16** in file **p_hhsamp_ip**:

- 1 = Control
- 2 = Remain version
- 3 = Leave version

The variables affected by this experiment are in file **p_indresp_ip**:

scacrt2acor, scacrt2b, scacrt2ccor, scacrt2d, petition

For initial findings, see [Butler and Fowler \(2024\)](#).

9.41 Robustness of climate change worries measurement

This experiment compares two versions of a single question on worry about climate change. Since the effects of climate change aggravate over time, it may mostly impact future generations. For this reason (elderly) people’s worries about climate change partly capture their valuation of the welfare of future generations. However, the climate change worry question in the Understanding Society questionnaire specifically refers to personal worries: "The effects of climate change are too far in the future to really worry me."

This experiment directly compares this version with an alternative that changes the last phrases from “worry me” to “worry about”.

The allocation to these two versions of this question was done at the household level, with equal proportions allocated to each group. The allocations were stratified by sample origin, wave 16 incentive treatment allocation, wave 16 mode allocation, and government office region. The controlling variable is **ff_climatechangew16** in file **p_hhsamp_ip**:

- 1 = ‘Worry me’ version
- 2 = ‘Worry about’ version

The variables affected by this experiment are in file **p_indresp_ip**:

scenv_nowo, scenv_nowoexp

For initial findings, see [Voorin Holt, Soetevent and van den Berg \(2024a\)](#).

9.42 Understanding of the long-term future

People’s perception of the long-term future plays an important role in the support for policies that address intergenerational challenges such as climate change, nuclear waste

storage and technological risks. Yet, many people find it difficult to think about the future beyond their own lifespan. This experiment explores how good are people's grasp of the long-term rather than short-term future, whether this understanding related to own remaining life expectancy and if this understanding depends on whether its measurement is framed in a specific context such as climate change.

Respondents were asked one or both of two sets questions. One set of questions focused on interest compounding and the other on rising seas levels. Each set had a one-sentence introduction and three follow-up questions. Questions each asked about the compounding impacts at 5, 30, and 100 year time spans. All questions asked about compounding at 1% per year.

There are two dimensions to this experiment: the placement of the questions and which set of questions were asked. The questions were placed in the early and late CASI sections. Respondents were randomly allocated to be asked either the 1) financial numeracy questions, 2) environmental numeracy questions, or 3) both sets. They were also randomly allocated to answer these questions early or late in the questionnaire. Those asked both sets received one (random) set early, and the other late.

There are two allocations controlling this experiment, both were allocated at the household level and stratified by `ff_climatechangew16`, sample origin, wave 16 mode allocation, and wave 16 incentive treatment allocation. The treatment allocation variables are in file **p_hhsamp_ip**:

ff_numeracyw16 (1/2 each)

- 1 = Environmental questions early/financial questions late
- 2 = Financial questions early/environmental questions late

ff_numquesw16 (0.4/0.4/0.2 allocation)

- 1 = Ask financial questions only
- 2 = Ask environmental questions only
- 3 = Ask both financial and environmental questions

The variables affected by this experiment are in file **p_indresp_ip**:

envnumear5y, envnumear30y, envnumear100y, finnumear5y, finnumear30y, finnumear100y, envnumlat5y, envnumlat30y, envnumlat100y, finnumlat5y, finnumlat30y, finnumlat100y

For initial findings, see [Voorin Holt, Soetevent and van den Berg \(2024b\)](#).

9.43 Mental health questions comparisons

At IP16, a study was conducted to compare the mental health questions that had been deployed in Understanding Society over the years.

The main survey questions about diagnoses of health conditions remained close to static for the first nine waves of the main survey. At Wave 10 and then again at Wave 14 of the main

survey the question wording changes about whether a doctor had ever diagnosed mental health conditions. Reviewing the data, there is some evidence that the change in wording has resulted in changes in prevalence. This experiment compares the three versions of asking the questions, pre-Wave 10, with Wave 10 changes, and the Wave 14 changes (current version). The treatment allocation variable, allocated at the household level, with one-third of the sample allocated to each version of the questions, is in file **p_hhsamp_ip**:

ff_mhealthexpw16 (1/3 each, allocation stratified by sampleorig ff_numeracyw16 ff_climatechangew16 ff_gridmodew16 ff_incentw16)

1 = Wave 8 version

2 = Wave 10 version

3 = Wave 14 version

The variables affected by this experiment are in file **p_indresp_ip**:

ahconda*, hcondsa*, hcondaa*, hcondb*, arthtypb, cancertypb*, diabetestypb, mhealthtypb*, hcondcodeb*, hcondab*, hcondsb*, mhcondc*, hcondc?, hcondc??, arthtypc, cancertypc*, diabetestypc, brainnervtypc*, spleenoutc, hcondcodec*, hcondac*, hcondsc*

For initial findings, see [Kumari and Burton \(2024\)](#).

9.44 Asking for child red book pictures

Child development up to the age of 2 is a key determinant of child and adult health. However, directly and regularly measuring physical health is invasive and time consuming for families, and costly for the Study. We therefore wished to investigate whether it is possible to collect such data from administrative records parents should hold for the child in the 'red book'.

The experiment asked respondents with children under 16 in the household to upload a photo or provide red book information. Based on IP15 data on household composition, half were allocated to being asked to provide photos online ahead of the interview, this request was sent to them four weeks before the interview invitation. The other half were informed in their advanced letter they would be asked within the interview. Those in the pre-interview group that did not provide the photo or information were asked to do so in the interview along with the second group. Similarly, any new parents identified in the IP16 interview were asked to provide a photo or information within the interview. Parents completing the interview by telephone were asked to provide data as part of the interview.

Those with children under one years old were asked to upload a picture, and if not, report from the red book, the height and weight of children at 6-8 weeks or the closest date to that. Respondents with children aged one or older were asked for these for when the children were one year old or the closest time to that. Respondents providing these data were given a conditional £2 incentive.

Eligible households were randomised to pre-interview or during-interview requests in equal proportions. The controlling variable is in file **p_hhsamp_ip**:

ff_redbookw16

(1/2 each of those with children under 16 at IP15, allocations stratified by
ff_kids015w16 ff_gridmodew16 ff_incentw16 ff_perksw16 sampleorig)

1 = Pre-interview request

2 = In-interview request

This experiment affects the following variables in file **p_indresp_ip**:

rbletter, rbdone, rbnottall, rbwill96, rbentint96, rbmeasure, rbothmeas, rbnottused1, rbnottused2, rbnottused3, rbnottused4, rbnottused5, rbnottused97, rbnottusedoth_code, nfrbnum

It also affects all variables in the file **p_redbook_ip**. This is a child level file, although on a few occasions both parents entered data for the same child, which has been retained. The following are documented in the IP16 questionnaire:

rbwill, rbwpimp, rbwoimp, rbhimp, rbwkmet, rbwogmet, rbhmet, rbusedwebtel, rbusedin, rbdated, rbdatem, rbdatey, nfrbdoby, redbooktype

The **p_redbook_ip** file contains additional variables that are not documented in the IP16 questionnaire. For information on these, see [Benzeval and Payne \(2024\)](#):

rbtwins, rbwhich, dobmatch, rbentint, rblength, smokeexp, namematch, rbsomewho, weighting, rbbooktype, rbfollowup, rbheadcirc, weightinkg, birthweight, rbimingsource, rbnottselctd, weight_g_dv, height_cm_dv, rbvisitweeks, uploadedpage, weight_kg_dv, breastfeeding, datesagematch, birthweighting, breastfedweeks, rbdetsource_dv, birthweightinkg, nodateofcontactreason

For initial findings, see [Benzeval, Aguirre, Baghal, and Mitchell \(2024\)](#).

9.45 Spatial cognition mobile app game

We asked respondents to download and use a game app for their smart phones. The data derived from playing this game can be used to measure spatial cognition, which is important in several aspects of life and is shown to be related to general cognition. This app is a game and is intended to be fun and less work than other app studies that IP respondents have been asked to complete in the past.

Respondents were randomly assigned to an incentive condition, half being offered £10 or £30 to participate. The allocation was at the individual level. The treatment allocation variable is in file **p_indsamp_ip**:

ff_navapincentw16

(1/2 each, allocations stratified by ff_incentw16 ff_perksw16 sampleorig)

1 = £10 conditional incentive

2 = £30 conditional incentive

For initial findings related to the experiment, see [Burton, Jäckle and Couper \(2024\)](#).

The data collected by this app game are available in file `p_sqh_results_ip`. For documentation on this data file, see [Coutrot et al \(2024\)](#).

9.46 Consent decision process

This experiment aimed to test whether participants can be encouraged to think about a consent request more reflectively and whether this then affects the consent decision and related indicators of effort and understanding.

The request was for consent to link data held by the Department for Work and Pensions (DWP) to survey responses.

Participants who completed in CAPI were equally allocated to one of three groups; those who complete online were equally allocated to one of five groups:

- (a) [All modes] **Standard**: standard 'BENLINK' question, previously carried in IP, most recently at IP10.
- (b) [All modes] **Benefits/risks**: version of question focusing on the benefits to the respondent of consenting and reassurance about things they might perceive as risks to them.
- (c) [All modes] **Value for science**: version of question focusing on the benefits to research and policy of having people consent.

- (d) [Web only] **Reasons for/against consenting**: respondents will be asked to list reasons for and against consenting before being asked the consent question.
- (e) [Web only] **Objective understanding**: respondents will be asked knowledge-check questions about information they have been told about the data linkage, before being asked the consent question.

(CATI respondents were not asked these questions.)

Within-script randomisation was used to allocate participants at the individual level.

Allocations

The variable controlling allocation to treatments, randomised at the individual level within the questionnaire script, varies depending on whether the interview was conducted face-to-face or via the web.

Allocation variable for face-to-face respondents:

congrpftf

(1/3 allocated to each of conditions 1, 2 and 3 for face-to-face respondents)

- 1 Standard DWP consent
- 2 Benefit-risk DWP consent
- 3 Research and policy DWP consent

Allocation variable for web respondents:

congrpweb

(1/5 each allocated to each of the five conditions for web respondents)

- 1 Standard DWP consent
- 2 Benefit-risk DWP consent
- 3 Research and policy DWP consent
- 4 Reasons pro con DWP consent
- 5 Understanding DWP consent

An apparent inconsistency between these allocation variables and the mode of completion variables for 12 cases is documented in the section on known data issues, under Inconsistency between mode of allocation variable and mode of completion.

Follow-up questions

After the consent questions, follow-up questions were asked of respondents in all arms, seeking to measure subjective understanding of the task, confidence in the consent decision, the consent decision process, and subjective effort in answering the question. These were variants of follow-up questions previously asked in relation to similar consent requests (related to linking HMRC data rather than DWP data), at Waves IP11 and IP15.

In addition, questions seeking to measure respondents' objective understanding of the consent request were asked as follow-up questions after the consent question, except for respondents in arm 5, who had already have answered such questions prior to the consent question.

Modules affected

benefitconsents_ip17, scabenefitsconsentfollowup_ip17

9.47 Identification of informal caregiving

The Understanding Society questionnaire includes questions designed to identify people who provide informal care and seeking to measure how much time they spend providing care.

The existing questions adopt an approach of asking about informal care in general. Alternative possible approaches exist, and it is important to understand whether these

alternative approaches can identify carers who might not be identified using the existing approach, or more accurately measure how much time they spend caring.

This experiment aimed to study an alternative set of questions, comparing them against the existing questions.

The new questions tested can be characterised as ‘activity-based questions’. A series of questions were asked of each respondent, asking them how long they spend each week undertaking a certain class of caring activity.

To facilitate within-person analysis of the results, a question-order experiment was used. All respondents were asked both the existing and the new (activity-based) questions. Respondents were randomly allocated to be asked one set of the questions early in the questionnaire and the other later in the questionnaire. The two places where the blocks of questions were asked were spaced out within the questionnaire to lessen the impact of the same information being requested too close together.

Allocations

The variable controlling allocation to treatments, randomised at the household level:

ff_careexpw17

(1/2 allocated to each condition, allocation stratified by **sampleorig**, **ff_incentw17**, **ff_gridmodew17**, {recoded version of **p_hhmodes** grouping (1,4) (2) (3, 5, 7)})

1 Activity-based question asked late

2 Activity-based question asked early

Modules affected

caringearly_ip17, caringlate_ip17

9.48 Labour market expectations

This experiment was intended to investigate respondents’ reporting of their labour market expectations. Specifically, it investigated how these vary, depending on a random allocation to have options presented to them either anchored to their current wages or their assessment of the market wage given their skills and experience.

All included respondents were asked, non-experimentally, to estimate their probabilities of being offered a new job within the next year, for both offers from their current employer and a different employer.

They were also be asked, non-experimentally, to estimate the probabilities that over the next 12 months they will (a) remain in their job, (b) quit their job, and (c) be laid off from their job, which should add up to 100%.

Respondents were also be asked to provide their current earnings from their (main) job, plus the minimum and maximum they might expect to be offered if they received a job offer, and the minimum they would accept.

For those respondents allocated to be asked their expectations based on a 'market wage' anchor, they were also asked what they think a typical amount offered would be.

Respondents were asked to rate the probability that their earnings in a year's time will be in each bucket describing a range:

Less than 85% of anchor	85%-95% of anchor	95%-100% of anchor	100%-105% of anchor	105%-115% of anchor	Greater than 115% of anchor
-------------------------	-------------------	--------------------	---------------------	---------------------	-----------------------------

The bounds for each of these buckets were expressed as monetary sums, calculated by the script based on the anchor values provided previously by the respondent.

For respondents who had indicated a reasonably high probability of receiving at least one job offer, they were also be asked to provide probability estimates that their best job offer will be in each of the same buckets as they used for the previous question.

Inclusion criteria

These questions were asked of respondents who completed via the web or completing the self-complete section within the face-to-face interview (CASI). Telephone (CATI) respondents were excluded from this module.

The questions were only be asked of respondents who report their main economic activity is paid employment.

Allocations

The variable controlling allocation to treatments, randomised at the household level:

ff_lmeexpw17

(1/2 allocated to each condition, allocation stratified by **sampleorig**, **p_nemp_dv**, **ff_gridmodew17**, **ff_incentw17**)

1 Current wage anchor

2 Market wage anchor

Modules affected

scalabourmarketexpectations_ip17

9.49 Indoor residential environment: consent for in-home sensor

Wave 17 included an overarching project to investigate people's indoor residential environment and their energy use. This included collection of questionnaire responses about people's homes and associated factors related to their energy use (see 10.16 Indoor residential environment and energy use: background information), as well as data collection via placing sensors in consenting respondents' homes and collecting information from their smart meters, if they have them and consent to it.

See also:

- 9.50 Domestic energy use: consent for smart meter data linkage
- 10.16 Indoor residential environment and energy use: background information

As part of this, we ran experiments related to the questions seeking consent to place sensors in people's home.

Follow-up questions were asked after the consent questions, to investigate variation based on the versions of the consent question asked. For the sensor consent, the follow-up questions were debrief questions, varying by whether the respondents provided or declined consent, seeking to measure their reasons for the decision they made.

Allocations

The variable controlling allocation to treatments, randomised at the household level:

ff_esensinfow17

(1/2 allocated to each condition, allocation stratified by **sampleorig**, **ff_gridmodew17**, **ff_incentw17**)

1. Full information in question text
2. Key information in question text, rest in help text

ff_esensdurw17

(1/3 allocated to condition 1, 1/6 allocated to each of the other conditions, allocation stratified by **sampleorig**, **ff_gridmodew17**, **ff_incentw17**, **ff_esensinfow17**, **ff_smetinfow17**)

1. Up to 6 months [1/3]
2. Up to 12 months [1/6]
3. Up to a year [1/6]
4. Up to 24 months [1/6]
5. Up to 2 years [1/6]

ff_esensfbw17

(1/2 allocated to each condition, allocation stratified by **sampleorig**, **ff_esensinfow17**, **ff_esensdurw17**, **ff_smetinfow17**)

1. Feedback from sensor promised
2. Feedback from sensor not mentioned

Modules affected

The consent questions (and associated follow-ups) were carried within the household questionnaire.

9.50 Domestic energy use: consent for smart meter data linkage

Wave 17 included an overarching project to investigate people's indoor residential environment and their energy use. This included collection of questionnaire responses about people's homes and associated factors related to their energy use (see 10.16 Indoor residential environment and energy use: background information), as well as data collection via placing sensors in consenting respondents' homes and collecting information from their smart meters, if they have them and consent to it.

See also:

- 9.49 Indoor residential environment: consent for in-home sensor
- 10.16 Indoor residential environment and energy use: background information

As part of this, we ran experiments related to the questions seeking consent to collect smart meter data.

Follow-up questions were asked after the consent questions, to investigate variation based on the versions of the consent question asked. For the smart meter consent, the follow-ups were knowledge-check questions, intended to measure whether respondents had taken on board accurate knowledge when they were providing (or declining) consent.

Allocations

The variables controlling allocation to treatments, randomised at the household level:

ff_smetinfow17

(1/3 allocated to each condition, allocation stratified by **sampleorig**, **ff_gridmodew17**, **ff_incentw17**, **ff_esensinfow17**)

1. Full information in bullet point format
2. Full information in paragraph format
3. Key information in bullet point format, rest in help text

Modules affected

These consent questions (and associated follow-ups) were carried within the household questionnaire.

10. Non-experimental studies

Non-experimental studies have occasionally been carried in the Innovation Panel since wave 5.

Since Wave 16, the annual Innovation Panel competition has included a specific call for proposals for studies comprising new questions without significant experimental or survey methods components. Each following wave will see a set of new questions included once for that specific wave. Questions cover such things as people's perceptions, experiences and opinions on a particular topic and are limited to one minute of questionnaire time.

10.1 Questions about twins

In Wave 5, this was a non-experimental study that tested the feasibility of asking respondents whether they are a twin (**e_twin**), the person number of their co-resident twin (**e_twinpno**), whether they would disclose their twin's address (**e_getwin**) and the likelihood that twin would participate in a follow-up study (**e_twinresp**).

Variables are in the "Twin Module" in record **e_indresp_ip**.

10.2 Measuring finger length ratios as indicator of prenatal testosterone exposure

This study tested the feasibility of measuring the finger-length ratio of respondents' second and fourth digit (2D:4D). 2D:4D has been proposed as a stable marker for prenatal testosterone exposure which in turn has organizational effects on human development and predicts a range of traits and outcomes later in life. Direct measurement of prenatal testosterone is expensive and invasive and can only be done on the embryo. Therefore 2D:4D as an indirect measurement is of potentially high value for researchers interested in human development and the life course. Given stability of 2D:4D as people age, measurement can occur ex-post at any panel wave with yet predictive value for respondents' life course outcomes.

In the face-to-face interviews the lengths of the ring and index fingers of both hands were measured by interviewers using digital Vernier calipers; in the web survey respondents were asked to measure themselves. The procedure is described in detail in the fieldwork document [IP6 Project Instructions for Interviewers](#).

In IP7 the measurement was repeated for new sample members, including the IP7 refreshment sample, and added to the youth self-completion questionnaire.

10.3 Associated Study: time-use diary

An Understanding Society Associated Study application was received from the Centre for Time Use Research, Department of Sociology at the University of Oxford. This was approved and implemented as part of IP7 data collection. This project aimed to use a diary study as a

basis for calibrating time-use estimates using the stylized questions in the main Understanding Society panel. Data from the time diaries are in the record **g_timediary_ip**.

It was proposed that responding adults would be asked to complete two time-use diaries, one of which would cover a randomly selected weekday (Monday-Friday) and the second a randomly selected weekend-day (Saturday-Sunday). All adults within the same household were allocated the same two days of the week. However, it should be noted that differences in date of interview may mean that adults in the same household weren't completing their diaries in the same weeks. For example, in a household which was allocated a Thursday and Saturday, an adult interviewed on Monday would complete the diary that week whilst another adult interviewed on the Friday would complete their diary for the same Saturday but the Thursday diary would be completed on the following week.

Controlling variables on record **hhsamp**:

ff_diaryweekw7

- | | |
|---|-----------|
| 1 | Monday |
| 2 | Tuesday |
| 3 | Wednesday |
| 4 | Thursday |
| 5 | Friday |

ff_diarywew7

- | | |
|---|----------|
| 1 | Saturday |
| 2 | Sunday |

Format of the time diary

The time diary used was a "light-touch" diary, that is, the respondents would indicate their activities against a pre-coded list, rather than writing down what they were doing. The week-day and weekend diaries were identical in layout, except that the week-day diary (green cover) had five boxes on the cover labelled Monday to Friday, and the weekend diary (blue cover) had just two boxes for Saturday and Sunday. The front cover also had boxes in which the interviewer filled in their interviewer number, the Serial number of the household and the person number of the respondent. The interviewer also filled in the first name, date-of-birth and sex of the respondent on the cover, along with the date of interview, diary day and diary date. The next two pages had instructions on how to complete the diary, and the following two pages had an example of how to complete the diary.

The recording of activities started on page 6. The time of day was indicated along the top of the page, with columns for each 10-minute period. Pages 6-7 covered Early morning and Morning (4.00am – 11.50am), pages 8-9 covered Afternoon and Early evening (12.00 to 7.50pm) and pages 10-11 covered Evening and Night (8.00pm to 3.50pm).

The rows of the diary contained the activities. There were 29 pre-coded activities, along with a 30th which was "Other" with an instruction to write the activity in. In addition to the activities, there were also requests for additional information. Where a respondent indicated that they had travelled, they were also asked to indicate their mode of travel. For

each time-slot, the respondent was also asked to indicate how much they enjoyed that activity (from 1=not at all to 7=very much) and who was with them.

To indicate that they had participated in a particular activity, the respondent marked with a line the time they were doing that activity. The start and finish time for the activity was to be marked by an X with the line joining the two. If they were doing a secondary activity at the same time, they were to mark the start and end times of this with a dot (•) and join these with a line.

Fieldwork procedures

Since IP7 employed a mixed-mode design, the arrangements for passing the diaries to the respondents differed by mode of interview. Responding adults in households that completed all their interviews online were sent their diaries in the mail, with a covering letter and a pre-paid return envelope. For households which required an interviewer visit – those in the F2F sample, and those in the mixed-mode sample who were not completed online – the interviewer handed over the diaries to the respondent at the end of the interview. In households where the interviewer visited, and where someone had completed online, the interviewer also gave the adult who responded online their diary. The interviewer also gave the respondent(s) a pre-paid return envelope in which to return the diaries.

The CAPI script included instructions that the interviewer read out to the respondent, whilst they showed them the diary. Online, there was no mention of the time-diary, but the covering letter which accompanied the time diaries introduced the time diary, and gave instructions on how to complete it. There was an additional £5 incentive for accepting the time-diary; this was included in the letter for those completing online and was handed over by the interviewer for those who accepted the time diary in person.

At the start of fieldwork, respondents were sent an SMS text message and/or email reminder the day before the diary day and one on the diary day itself. Where we had the mobile telephone number and an email address for a respondent, this meant they got two reminders (SMS + email) on each of these days. However, we had a small number of complaints about these multiple reminders, and so the reminder the day before the diary day was dropped. All respondents who had accepted or been sent time diaries, but who hadn't returned them within two weeks of the completion dates were telephoned to remind them to complete and return their diaries.

Response to time diaries

Adults who participated online were more likely to return time diaries than those who participated face-to-face. Almost half of face-to-face respondents returned a diary, compared to two-thirds of online participants (49.9% compared to 66.1%). There was no difference between the original IP1 sample and the IP4 refreshment sample in proportion of respondents returning diaries (54.0% compared to 57.1%). Just under half of the IP7 refreshment sample respondents returned diaries (49.3%), less than both the other samples.

10.4 Spending Study 1

Sample members in respondent wave 9 households were invited to download an app and record their spending for one month. This study was fielded after the close of wave 9 fieldwork, in autumn 2016. The data from the receipt scanning study have been deposited with the UK Data Service (SN 8749, <http://doi.org/10.5255/UKDA-SN-8749-1>) and can be linked with the Innovation Panel data using the individual identifier “pidp”.

10.5 Consent to link Twitter data

This module aims to look at the feasibilities and practicalities of linking social media (in particular Twitter) and survey data in a longitudinal context, and how they can be combined to improve the quality of both. There was no experimental allocation. All respondents were asked about whether they use Twitter, and if so, their willingness to link their Twitter account to their survey responses.

Variables on record **j_indresp_ip**:

twithave, twitlink

A version of this module was used at Wave 15 with experimental components. See 9.33 Twitter consent.

10.6 An investigation of children’s consistency in reporting their parents’ occupations

This study adds two questions to the youth paper questionnaire (for 10-15 year olds) and to the young adult self-completion module used in the IP. These two questions ask youths about what their mother’s and father’s occupation is, which can then be converted to a SOC code as is done in the adult version of the survey when asking about occupation.

There was no random allocation to conditions in this study. All youth respondents (10-15 years old) that are given the paper youth questionnaire had these two questions added, all in the same location, while all young adults were asked in the self-completion module in the face-to-face surveys, or on the web. The following questions from the youth questionnaire are those added to the survey.

The variables used for this study are the following.

On the record **k_youth_ip** (youth questionnaire, 10-15 year olds):

ypmatsoc90, ypmatsoc90_cc, ypmatsoc00, ypmatsoc00_cc, ypmatsoc10, ypmatsoc10_cc, ypfatsoc90, ypfatsoc90_cc, ypfatsoc00, ypfatsoc00_cc, ypfatsoc10, ypfatsoc10_cc

On the record **k_indresp_ip** (young adults module, 16-21 year olds):

matocc90, matocc00, matocc10, fatocc90, fatocc00, fatocc10

10.7 Panel conditioning

Respondents that have taken the survey before were asked about how the survey affected the way they think, their behaviours, and their attitudes. All respondents were asked why they are participating in the survey. The module was asked in self-completion mode, and any telephone respondents were not asked this module. Due to time constraints, this module was included in the [fieldwork compression experiment](#) set of modules, both longer and shorter sets. That is, it was asked of all respondents allocated to groups 1-4 for the fieldwork compression experiment, but not of respondents allocated to the control group (group 5). Other than the routing from being in the fieldwork compression experiment and past survey participation, there are no other controlling variables. The questions are in the questionnaire module **scapanelconditioning_ip13**. the variables are in the record **m_indresp_ip**.

10.8 Living Apart Together (LATs) partner and survey

There are a growing number of people with partners that live apart in different abodes. The survey already asks respondents about any partner they have not living with them; this module is asked every wave in the self-completion section.

At Wave 13 respondents that have identified as having a partner living apart were asked for that partner's contact details (sample 1). These questions were added to the end of the existing module asking about non-resident relationships. The added questions were not experimental in nature – every respondent identifying a non-resident partner was asked the questions. The additional questions collected contact details of these partners in a manner similar to questions asked in the Stable Contact Details module. The request made it clear that the information collected would potentially be used to invite these partners to answer a survey, with no information about the respondent being given other than to say they gave the contact details. The new questions are in the questionnaire module **scalat_ip13**. Only a small percentage of respondents provided usable contact details for their partners.

At Wave 14 this module was improved and repeated but with an experimental component (samples 2 and 3) as described in the section describing the experiment (see: 7.23 Asking for Living Apart Together partner details). Again, only a small percentage of participants provided usable contact details.

Following Wave 14, all IP13 and IP14 participants who were in a living apart together (LAT) relationship, and those partners for whom they had provided contact details, were invited to take part in a 15-20 minute web survey between March and May 2022. The aims of this survey were four-fold to:

- trial approaching 'significant others' to participate in data collection;
- gather comprehensive information from LAT couples to enhance data collected from existing sample members' lives;
- create unique opportunity for research on such couples.
- test new content for LAT relationships to inform future waves.

A dedicated webpage⁴ was created for participants with tailored videos aimed at explaining the survey in more detail and the research and policy value of participating in the study. This contained links to participant information sheet.

Participants were offered a £5 gift voucher or option to make a charitable donation for completing the survey.

Of the 252 sample members in LAT relationships invited to take part, 8% reported they were no longer in a LAT relationship; 73% did not reply to the invitation, and 19% took part. Among the 36 partners of sample members for whom we had contact details, 5% replied to say they were no longer in a LAT relationship with the sample member, 86% did not respond and 8% took part.

See the Non-experimental studies section 9 Asking for Parents Living Apart other parent details, for information on similar content associated with a different group of respondents — parents living apart.

Table LAT survey fieldwork

Action	Date
Survey invites via letter and email (samples 1 and 2)	31 March 2022
Reminder 1 sent to all sample members not completing survey (samples 1 and 2)	23 April 2022
Reminder 2 sent to sample members (sample 1)	6 May 2022
Request for contact details email/letter sent (sample 3)	4 April 2022
Reminder request for details (sample 3)	23 April 2022
Survey invites sent to remaining non-coresident partners (sample 3)	May 2022

The participant letters and information sheet are available as part of the IP data documentation.

The follow-on LATS Survey was fielded in Qualtrics. The questionnaire was based on the literature and other surveys of couples living apart together (the Generations and Gender Survey and studies by NatCen and University of Birkbeck). We shared a version for consultation with key experts in the field and more widely via social media and the website. The survey included:

- Characteristics of sample members' partner living apart
- Age, gender, ethnicity, religion, national identity, country of birth, citizenship, health and education level directly asked of significant other (rather than partner reports)
- Information on living arrangements, housing status, employment status and benefit eligibility
- Legal marital status and brief marital and cohabitation histories.
- Family dynamics within the LAT couple – existence of mutual children, coparenting
- Reasons for living apart
- How do people feel about their LAT relationship

⁴ <https://www.understandingsociety.ac.uk/participants/lats>

- How do LAT relationships function and do they provide financial, practical and emotional support to each other
- Future intentions – e.g. intention to move, live together, marry.
- The impact (if any) of the COVID-19 pandemic on these relationships

We also sought feedback relating to:

- The nature and appropriateness of the questions;
- Willingness to participate in possible future studies;

The full questionnaire is available as part of the IP data documentation, along with the other IP questionnaires.

10.9 Asking for Parents Living Apart other parent details

At IP13 and IP14, respondents with partners that live apart (LAT) in different abodes were asked for additional contact details for these partners (see: 10.8 Living Apart Together (LATs) partner and survey). The additional questions collected contact details of these partners in a manner similar to questions asked in Stable Contact Details. The request made clear the information collected would potentially be used to invite these partners to answer a survey, with no information about the respondent being given other than to say they gave the contact details.

At IP15, this asking for others' contact details was done for parents living apart. When a respondent identifies as having children they do not live with, follow-up questions are asked about those children, including who they live with. Where non-resident children are living with their other parent, respondents are asked for contact details of that other parent. As respondents may have multiple non-resident children with multiple partners, the questions loop all non-resident children with different parents, so respondents might be asked for multiple contact details. As with asking for LATs' contact details, respondents were told information collected would potentially be used to invite these other parents to answer a survey, with no information about the respondent being given other than to say they gave the contact details.

This content was not experimental, in that there was no manipulation of the way these questions were asked, but was intended as a pilot exercise to test the feasibility of asking contact details for parents living apart in this manner.

The questions were in the questionnaire module **nonreschild**.

10.10 What constitutes an interesting size effect when measuring people's psychological state?

At IP16, respondents were asked to rate how their life satisfaction at the time of the interview compared to their life satisfaction at the previous interview.

In combination with respondents' self-reported life satisfaction scores at both the current and previous interviews, the information from the comparison question was intended to

contribute knowledge on what the smallest effect sizes of interest is for the life satisfaction measure.

For initial findings, see [Mansfield and Przybylski \(2024\)](#).

10.11 People's expectations of gender discrimination related to work

At IP16, respondents were asked whether they expected to face gender discrimination when seeking work.

For initial findings, see [Popli and Ratcliffe \(2024\)](#).

10.12 The extent and nature of the use of domestic workers

At IP16, respondents were asked about the use of paid domestic labour in their household, including the number of people doing paid work and the types of work (e.g., cleaning, laundry, ironing, gardening, house repairs).

For initial findings, see [Lieutaud \(2024\)](#).

10.13 Judging the passage of time

At IP16, respondents were asked questions about their subjective experience of the passage of time over a range of time periods.

For initial findings examining how these passage of time judgements vary with employment status, see [Raj and Delaney \(2024\)](#).

10.14 Adaptation to eco-climate emergency

At Wave 17, this study aimed to investigate associations between respondents' affect with respect to climate change and their degree of support/opposition for various proposed policies that are intended to respond to it.

Respondents were asked about the extent to which they feel different emotions when they think about climate change: worried, disappointed, interested, hopeful, fearful, and calm.

They were also asked how much they support or oppose a set of policies: ending the sale of petrol / diesel vehicles, ending expansion of fossil fuels, investment in R&D including green innovations, and increasing offshore wind capacity.

Modules affected

Scaclimateadaptation_ip17

10.15 Youth online survey

Where the household grid and at least one parent completes online, a youth self-completion questionnaire is usually posted to the household. In the design at wave 17, the invitation letter included a URL to the online youth survey and introducing it as a 'new, redesigned, more fun' survey on which the young people's feedback would be sought to improve it further.

The paper copy was posted with the invitation letter, as usual. At the second stage, if the youth self-completion had not been completed within one month, a reminder letter followed but it included a link to the online survey and the accompanying motivating message.

Uptake of the online survey vs. the paper booklet can be compared with the IP16 'default' youth questionnaire design, albeit these were not tested against each other experimentally.

The online Youth survey was scripted by the Understanding Society in-house team, rather than the fieldwork agency.

See also: 7.26 Youth online survey.

10.16 Indoor residential environment and energy use: background information

Wave 17 included an overarching project to investigate people's indoor residential environment and their energy use.

See also:

- 9.49 Indoor residential environment: consent for in-home sensor
- 9.50 Domestic energy use: consent for smart meter data linkage

As well as the experiments testing variants of consent questions (see sections above), additional content about homes and energy usage were carried non-experimentally within the household questionnaire.

Some of these questions were variants of questions previously carried within the survey, such as questions on the heating systems and consumer durables within people's homes, featuring new response options. Others were new questions, seeking to measure features of respondents' homes we have not previously measured.

New questions

- **heatcont** [Heating controls]
- **hwtaps** [Hot water taps]
- **hwshower** [Hot water shower]
- **warmroom** [Living room warm in winter]

- **fueleffort** [Effort to reduce energy use]
- **meterown** [Household energy meters]

Modified or adapted questions

- **heat** [Household heating]: Adapted version of **heatch**
- **cduseadp** [Consumer durables in accommodation]: Adapted version of **cduse** with revised response options.

11. Support and resources

Understanding Society has a wealth of information for data users of all levels.

It has a highly comprehensive online source of information regarding its variables, methodology, survey design and implementation. It is also an up-to-date source of [training courses](#), [data releases](#) and other relevant news regarding longitudinal research.

Useful documentation links

The Innovation Panels documentation provides links resources such as the questionnaires, technical and fieldwork documents and variable search. A [pathway for experienced users page](#) explores the data, the Code creator and highlights the different valued added datasets we produce.

Variable search

Find the variables you need for your research by searching by variable name or data file. It can be used to find other variables with related data throughout the dataset.

Creating syntax with the Code creator

To help researchers get started with their research the Code creator extracts data from the Main Study and produces a simple flat data file. Select the variables needed from the [Variable Search](#), 'save' the variables and 'build' the code. It provides you with ready-made Stata syntax to run on the downloaded data.

Topic page

The [Survey methods topic page](#) provides an overview of the methodological processes conducted on the Study.

11.1 User Forum

The Understanding Society study has a dedicated [User support forum](#) where after a short registration users can browse Frequently Asked Questions and raise new data issues. We request that researchers notify us about errors, inconsistencies, and other problems with the data identified during use of the data.

Some tips about reporting data issues: only use data from the latest release, if you are using data from past releases the problem may have been corrected. Provide a brief description of the issue and, if possible, include examples of syntax and tabulations. Specify the instrument and wave of the data with the problem, for example, IP 4.

If you have a question about the data, post your question at the online data [User forum](#). Users should read the "How to raise an issue" guidance before posting a question. If you have a suggestion for improving the data, such as creating new derived variables, suggestions for data harmonisation, adding variables linked from external datasets, you can

also post this via the User forum by selecting the category 'Suggestions for data improvements'.

The forum is monitored Monday-Friday and we aim to answer simple questions within 2 working days and more complex questions within 7.

Email

Users may also email User Support directly using our [email address](#). Our preferred mode of communication is via the forum as other users may then also benefit from the information provided.

Online helpdesk

If you'd like to speak to a member of the User Support team you can join an online helpdesk session. These are run via video conferencing software and are one-on-one sessions with a member of the User Support team.

If you would like to access the online helpdesk [please email us](#) and we will respond with joining information and arrange a convenient time for the conversation.

11.2 Training

We offer both [DIY training courses](#) (via Open Essex MoodleX) and [tutor-assisted workshops](#) which give a general overview of the Main Understanding Society Study and demonstrate how to prepare the data for analysis using multiple statistical software packages.

Workshops are also available on specific aspects of the Study such as weights, biomarker and genetics data.

To learn more about these training workshops and how to register visit the [training hub](#).

References – general

- Al Baghal, T. and Lynn, P. (2015). Using Motivational Statements in Web Instrument Design to Reduce Item Missing Rates in a Mixed-Mode Context. *Public Opinion Quarterly*, 79(2), 568-579.
- Benzeval, M., Aguirre, E., Baghal, T. A., & Mitchell, L. (2024). 9. Obtaining measurement data from the 'red book'. In Vine, J. et al., *Understanding Society Innovation Panel Wave 16: Results from methodological experiments and new data* (pp. 46–50). Institute for Social and Economic Research, University of Essex. <https://www.understandingsociety.ac.uk/wp-content/uploads/working-papers/2024-11.pdf#page=46>
- Benzeval, M. and Payne, J. (2024). 19. Child Development Measures from the “Red Book”. In Vine, J. et al. *Understanding Society Innovation Panel Wave 16: Results from Methodological Experiments and New Data*, Understanding Society Working Paper Series 2024–11. Colchester, Essex: Institute for Social and Economic Research, University of Essex. <https://www.understandingsociety.ac.uk/wp-content/uploads/working-papers/2024-11.pdf#page=148>
- Burton, J. (2024). 11. Use of a government logo. In Vine, J. et al., *Understanding Society Innovation Panel Wave 16: Results from methodological experiments and new data* (pp. 62–66). Institute for Social and Economic Research, University of Essex. <https://www.understandingsociety.ac.uk/wp-content/uploads/working-papers/2024-11.pdf#page=62>
- Burton, J., Jäckle, A., & Couper, M. (2024). 10. The effect of incentives on the willingness to participate in a gamified app study. In Vine, J. et al., *Understanding Society Innovation Panel Wave 16: Results from methodological experiments and new data* (pp. 51–61). Institute for Social and Economic Research, University of Essex. <https://www.understandingsociety.ac.uk/wp-content/uploads/working-papers/2024-11.pdf#page=51>
- Butler, C., & Fowler, C. (2024). 18. Cognitive reflection and politically motivated reasoning. In Vine, J. et al., *Understanding Society Innovation Panel Wave 16: Results from methodological experiments and new data* (pp. 133–147). Institute for Social and Economic Research, University of Essex. <https://www.understandingsociety.ac.uk/wp-content/uploads/working-papers/2024-11.pdf#page=133>
- Chung, H., & Wang, S. (2024). 16. Flexibility stigma in the UK post-pandemic: Intersectionality of stigma across gender and parental lines. In Vine, J. et al., *Understanding Society Innovation Panel Wave 16: Results from methodological experiments and new data* (pp. 105–123). Institute for Social and Economic Research, University of Essex. <https://www.understandingsociety.ac.uk/wp-content/uploads/working-papers/2024-11.pdf#page=105>
- Coutrot, A., Al Baghal, T. and Spiers, H. (2024). 20. Sea Hero Quest: Spatial Navigation Data Linked to the Innovation Panel Wave 16. In Vine, J. et al. *Understanding Society Innovation Panel Wave 16: Results from Methodological Experiments and New Data*, Understanding Society Working Paper Series 2024–11. Colchester, Essex: Institute for Social and Economic Research, University of Essex. <https://www.understandingsociety.ac.uk/wp-content/uploads/working-papers/2024-11.pdf#page=156>

Jäckle, A., Burton, J., Couper, M. P., Perelli-Harris, B. (2023). Participation of household panel members in daily burst measurement using a mobile app: effects of position of the invitation, bonus incentives, and number of daily questions. *Understanding Society Working Paper 2023-01*. Colchester: University of Essex.

<https://www.understandingsociety.ac.uk/research/publications/working-paper/understanding-society/2023-01/>

Kaminska, O. and Foulsham, T. (2016). Eye-tracking social desirability bias. *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique*, 130(1), 73-89.

Kumari, M., & Burton, J. (2024). 6. The prevalence of depression in Understanding Society. In Vine, J. et al., *Understanding Society Innovation Panel Wave 16: Results from methodological experiments and new data* (pp. 25–28). Institute for Social and Economic Research, University of Essex. <https://www.understandingsociety.ac.uk/wp-content/uploads/working-papers/2024-11.pdf#page=25>

Lieutaud, M. (2024). 14. Domestic workers and platforms. In Vine, J. et al., *Understanding Society Innovation Panel Wave 16: Results from methodological experiments and new data* (pp. 89–95). Institute for Social and Economic Research, University of Essex. <https://www.understandingsociety.ac.uk/wp-content/uploads/working-papers/2024-11.pdf#page=89>

Lynn, P. (2009). Sample design for Understanding Society. *Understanding Society Working Paper 2009-01*. Colchester: University of Essex. <https://www.understandingsociety.ac.uk/sites/default/files/downloads/working-papers/2009-01.pdf>

Lynn, P. and Jäckle, A. (2019). Mounting multiple experiments on longitudinal social surveys: design and implementation considerations, in P.J. Lavrakas, M.W. Traugott, C. Kennedy, A.L. Holbrook, E.de Leeuw, and B.T. West (eds.) *Experimental Methods in Survey Research: Techniques that Combine Random Sampling with Random Assignment*. Hoboken, N.J.: Wiley. Ch. 15:293-308. <https://doi.org/10.1002/9781119083771>.

Lynn, P. and Kaminska, O. (2010). Weighting strategy for Understanding Society. *Understanding Society Working Paper 2010-05*. Colchester: University of Essex. <https://www.understandingsociety.ac.uk/sites/default/files/downloads/working-papers/2010-05.pdf>

Mansfield, K. L., & Przybylski, A. K. (2024). 12. Establishing a databank of smallest effect sizes of interest (SESOI) for prominent psychological constructs. In Vine, J. et al., *Understanding Society Innovation Panel Wave 16: Results from methodological experiments and new data* (pp. 67–75). Institute for Social and Economic Research, University of Essex. <https://www.understandingsociety.ac.uk/wp-content/uploads/working-papers/2024-11.pdf#page=67>

Parutis, V., & Burton, J. (2024a). 7. Using e-vouchers for unconditional incentives. In Vine, J. et al., *Understanding Society Innovation Panel Wave 16: Results from methodological experiments and new data* (pp. 29–35). Institute for Social and Economic Research, University of Essex. <https://www.understandingsociety.ac.uk/wp-content/uploads/working-papers/2024-11.pdf#page=29>

Parutis, V., & Burton, J. (2024b). 8. Youth survey experiments. In Vine, J. et al., *Understanding Society Innovation Panel Wave 16: Results from methodological experiments and new data* (pp. 36–45). Institute for Social and Economic Research, University of Essex. <https://www.understandingsociety.ac.uk/wp-content/uploads/working-papers/2024-11.pdf#page=36>

Popli, G., & Ratcliffe, A. (2024). 13. Anticipated gender discrimination in the labour market. In Vine, J. et al., *Understanding Society Innovation Panel Wave 16: Results from methodological experiments and new data* (pp. 76–88). Institute for Social and Economic Research, University of Essex. <https://www.understandingsociety.ac.uk/wp-content/uploads/working-papers/2024-11.pdf#page=76>

Raj, S., & Delaney, L. (2024). 17. The relationship between employment status, passage of time judgements and wellbeing. In Vine, J. et al., *Understanding Society Innovation Panel Wave 16: Results from methodological experiments and new data* (pp. 124–132). Institute for Social and Economic Research, University of Essex. <https://www.understandingsociety.ac.uk/wp-content/uploads/working-papers/2024-11.pdf#page=124>

Schuman, H. and Presser, S. (1981). *Questions and Answers in Attitude Surveys: Experiments on Question Form, Wording, and Context*. Thousand Oaks, CA.: Sage.

University of Essex. Institute for Social and Economic Research (2021) Understanding Society: Spending Study 1, 2016-2017. [data collection]. UK Data Service. SN: 8749, DOI: <http://doi.org/10.5255/UKDA-SN-8749-1>.

University of Essex. Institute for Social and Economic Research (2022a) Understanding Society: Spending Study 2, 2018-2019. [data collection]. UK Data Service. SN: 8909, DOI: <http://doi.org/10.5255/UKDA-SN-8909-1>.

University of Essex. Institute for Social and Economic Research (2023) Understanding Society: Innovation Panel Wellbeing App Study, 2020. [data collection]. UK Data Service. SN: 9065, DOI: <http://doi.org/10.5255/UKDA-SN-9065-1>.

University of Essex, Institute for Social and Economic Research. (2022b). *Understanding Society: Innovation Panel Life Events Study, 2020*. [data collection]. UK Data Service. SN: 8990, DOI: <http://doi.org/10.5255/UKDA-SN-8990-1>.

Voorintholt, L., Soetevent, A. R., & van den Berg, G. J. (2024a). 5. Measuring worries about climate change: The effect of a subtle wording change. In Vine, J. et al., *Understanding Society Innovation Panel Wave 16: Results from methodological experiments and new data* (pp. 19–24). Institute for Social and Economic Research, University of Essex. <https://www.understandingsociety.ac.uk/wp-content/uploads/working-papers/2024-11.pdf#page=19>

Voorintholt, L., Soetevent, A. R., & van den Berg, G. J. (2024b). 15. Numeracy and the long-term future. In Vine, J. et al., *Understanding Society Innovation Panel Wave 16: Results from methodological experiments and new data* (pp. 96–104). Institute for Social and Economic Research, University of Essex. <https://www.understandingsociety.ac.uk/wp-content/uploads/working-papers/2024-11.pdf#page=96>

Williams, J., Ward, L., Burton, J., Carpenter, H., Cole, K., Hayward, B., Nicolaas, G., Parutis, V., Seymour, B., Thornton, K., Wood, M., and Woodward, L. (2022) *Understanding Society Wave 14 Boost Trial: Experiments with methods of recruiting a probability online boost sample*, *Understanding Society Working Paper 2022-07*. Colchester: University of Essex.
<https://www.understandingsociety.ac.uk/sites/default/files/downloads/working-papers/2022-07.pdf>.

References – summary results from IP experiments and new data

Burton, J. (2008). Understanding Society: Some preliminary results from the wave 1 Innovation Panel. *Understanding Society Working Paper 2008-03*. Colchester: University of Essex.

<https://www.understandingsociety.ac.uk/sites/default/files/downloads/working-papers/2008-03.pdf>

Burton, J, Laurie, H, & Uhrig, SCN. (2010). Understanding Society Innovation Panel wave 2: methodological experiments. *Understanding Society Working Paper 2010-04*. Colchester: University of Essex.

<https://www.understandingsociety.ac.uk/sites/default/files/downloads/working-papers/2010-04.pdf>

Burton, J, Budd, S, Gilbert, E, Jäckle, A, McFall, SL, & Uhrig, SCN. (2011). Understanding Society Innovation Panel wave 3: results from methodological experiments. *Understanding Society Working Paper 2011-05*. Colchester: University of Essex.

<https://www.understandingsociety.ac.uk/sites/default/files/downloads/working-papers/2011-05.pdf>

Budd, S, Gilbert, E, Burton, J, Jäckle, A, Kaminska, O, Uhrig, SCN, et al. (2012). Understanding Society Innovation Panel wave 4: results from methodological experiments. *Understanding Society Working Paper 2012-06*. Colchester: University of Essex.

<https://www.understandingsociety.ac.uk/sites/default/files/downloads/working-papers/2012-06.pdf>

Burton, J. (ed.) (2013). Understanding Society Innovation Panel wave 5: results from methodological experiments. *Understanding Society Working Paper 2013-06*. Colchester: University of Essex.

<https://www.understandingsociety.ac.uk/sites/default/files/downloads/working-papers/2013-06.pdf>

Al Baghal, T. (ed.) (2014). Understanding Society Innovation Panel wave 6: results from methodological experiments. *Understanding Society Working Paper 2014-04*. Colchester: University of Essex.

<https://www.understandingsociety.ac.uk/sites/default/files/downloads/working-papers/2014-04.pdf>

Al Baghal, T. (ed.) (2015). Understanding Society Innovation Panel wave 7: results from methodological experiments. *Understanding Society Working Paper 2015-03*. Colchester: University of Essex.

<https://www.understandingsociety.ac.uk/sites/default/files/downloads/working-papers/2015-03.pdf>

Al Baghal, T. (ed.) (2016). Understanding Society Innovation Panel Wave 8: results from methodological experiments. *Understanding Society Working Paper 2016-02*. Colchester: University of Essex.

<https://www.understandingsociety.ac.uk/sites/default/files/downloads/working-papers/2016-02.pdf>

Gaia, A. (ed.) (2017). Understanding Society Innovation Panel Wave 9: results from methodological experiments. *Understanding Society Working Paper 2017-07*. Colchester: University of Essex.

<https://www.understandingsociety.ac.uk/sites/default/files/downloads/working-papers/2017-07.pdf>

Al Baghal, T. (ed.) (2018). Understanding Society Innovation Panel Wave10: results from methodological experiments. *Understanding Society Working Paper 2018-06*. Colchester: University of Essex.

<https://www.understandingsociety.ac.uk/sites/default/files/downloads/working-papers/2018-06.pdf>

Al Baghal, T. (ed.) (2019). Understanding Society Innovation Panel Wave 11: results from methodological experiments. *Understanding Society Working Paper 2019-03*. Colchester: University of Essex.

<https://www.understandingsociety.ac.uk/sites/default/files/downloads/working-papers/2021-09.pdf>

Al Baghal, T. (ed.) (2021) Understanding Society Innovation Panel Wave 13: results from methodological experiments, *Understanding Society Working Paper 2021-09*, Colchester: University of Essex.

<https://www.understandingsociety.ac.uk/sites/default/files/downloads/working-papers/2021-09.pdf>

Jäckle, A. (ed.) (2022) Understanding Society Innovation Panel Wave 14: results from methodological experiments, *Understanding Society Working Paper 2022-06*, Colchester: University of Essex.

<https://www.understandingsociety.ac.uk/sites/default/files/downloads/working-papers/2022-06.pdf>

Vine, J. et al. (2023) Understanding Society Innovation Panel Wave 15: results from methodological experiments, *Understanding Society Working Paper 2023-10*, Colchester: University of Essex.

<https://www.understandingsociety.ac.uk/sites/default/files/downloads/working-papers/2023-10.pdf>

Vine, J. et al. (2024) Understanding Society Innovation Panel Wave 16: Results from methodological experiments and new data, *Understanding Society Working Paper 2024-11* Colchester: University of Essex.

<https://www.understandingsociety.ac.uk/sites/default/files/downloads/working-papers/2024-11.pdf>

Appendix A: Controlling variables for conjoint preference experiment (9.26)

This appendix contains details of the controlling variables related to the conjoint preference experiment described at: 9.26 Does competition over public services decrease support for residency rights of immigrants?

On record K_INDRESP_IP:

cjisex1a – Sex of applicant 1, pairing 1

cjisex1a =1 {Male}

cjisex1a =2 {Female}

cjimar1a- Marital status of applicant 1, pairing 1

cjimar1a =1 {Single}

cjimar1a =2 {Divorced}

cjimar1a =3 {Married}

cjichild1a- Number of children of applicant 1, pairing 1

cjichild1a =1 {0}

cjichild1a =2 {1}

cjichild1a =3 {2}

cjichild1a =4 {4}

cjichild1a =5 {6}

cjihome1a – Home ownership of applicant 1, pairing 1

cjihome1a = 1 {Owns own home}

cjihome1a = 2 {Lives in private rental}

cjihome1a = 3 {Lives in social housing}

cjihealth1a -Health status of applicant 1, pairing 1

cjihealth1a =1 {Pre-existing medical conditions}

cjihealth1a =2 {No medical conditions}

cjiwork1a- Employment status of applicant 1, pairing 1

cjiwork1a = 1 {Stay at home parent}

cjiwork1a = 2 {Unemployed, receiving benefits}

cjiwork1a = 3 {Unemployed, receiving no benefits}

cjiwork1a = 4 {Employed}

cjires1a – Years of UK residency of applicant 1, pairing 1

cjires1a =1 {2}

cjires1a =2 {5}

cjires1a =3 {10}

cjires1a =4 {15}

cjinat1a - County of origin of applicant 1, pairing 1

cjinat1a = 1 {India}
cjinat1a = 2 {Somalia}
cjinat1a = 3 {Jamaica}
cjinat1a = 4 {Poland}
cjinat1a = 5 {Germany}
cjinat1a = 6 {Spain}
cjinat1a = 7 {New Zealand}
cjinat1a = 8 {Mexico}

cjisex1b – Sex of applicant 2, pairing 1

cjisex1b =1 {Male}
cjisex1b =2 {Female}

cjimar1b- Marital status of applicant 2, pairing 1

cjimar1b =1 {Single}
cjimar1b =2 {Divorced}
cjimar1b =3 {Married}

cjichild1b- Number of children of applicant 2, pairing 1

cjichild1b =1 {0}
cjichild1b =2 {1}
cjichild1b =3 {2}
cjichild1b =4 {4}
cjichild1b =5 {6}

cjihome1b – Home ownership of applicant 2, pairing 1

cjihome1b = 1 {Owns own home}
cjihome1b = 2 {Lives in private rental}
cjihome1b = 3 {Lives in social housing}

cjihealth1b -Health status of applicant 2, pairing 1

cjihealth1b =1 {Pre-existing medical conditions}
cjihealth1b =2 {No medical conditions}

cjiwork1b- Employment status of applicant 2, pairing 1

cjiwork1b = 1 {Stay at home parent}
cjiwork1b = 2 {Unemployed, receiving benefits}
cjiwork1b = 3 {Unemployed, receiving no benefits}
cjiwork1b = 4 {Employed}

cjires1b – Years of UK residency of applicant 2, pairing 1

cjires1b =1 {2}
cjires1b =2 {5}
cjires1b =3 {10}
cjires1b =4 {15}

cjinat1b - County of origin of applicant 2, pairing 1

cjinat1b = 1 {India}
cjinat1b = 2 {Somalia}
cjinat1b = 3 {Jamaica}
cjinat1b = 4 {Poland}
cjinat1b = 5 {Germany}
cjinat1b = 6 {Spain}
cjinat1b = 7 {New Zealand}
cjinat1b = 8 {Mexico}

cjisex2a – Sex of applicant 1, pairing 2

cjisex2a =1 {Male}
cjisex2a =2 {Female}

cjimar2a- Marital status of applicant 1, pairing 2

cjimar2a =1 {Single}
cjimar2a =2 {Divorced}
cjimar2a =3 {Married}

cjichild2a- Number of children of applicant 1, pairing 2

cjichild2a =1 {0}
cjichild2a =2 {1}
cjichild2a =3 {2}
cjichild2a =4 {4}
cjichild2a =5 {6}

cjihome2a – Home ownership of applicant 1, pairing 2

cjihome2a = 1 {Owns own home}
cjihome2a = 2 {Lives in private rental}
cjihome2a = 3 {Lives in social housing}

cjihealth2a -Health status of applicant 1, pairing 2

cjihealth2a =1 {Pre-existing medical conditions}
cjihealth2a =2 {No medical conditions}

cjiwork2a- Employment status of applicant 1, pairing 2

cjiwork2a = 1 {Stay at home parent}
cjiwork2a = 2 {Unemployed, receiving benefits}
cjiwork2a = 3 {Unemployed, receiving no benefits}
cjiwork2a = 4 {Employed}

cjires2a – Years of UK residency of applicant 1, pairing 2

cjires2a =1 {2}
cjires2a =2 {5}
cjires2a =3 {10}

cjires2a =4 {15}

cjinat2a - County of origin of applicant 1, pairing 2

cjinat2a = 1 {India}

cjinat2a = 2 {Somalia}

cjinat2a = 3 {Jamaica}

cjinat2a = 4 {Poland}

cjinat2a = 5 {Germany}

cjinat2a = 6 {Spain}

cjinat2a = 7 {New Zealand}

cjinat2a = 8 {Mexico}

cjisex2b– Sex of applicant 2, pairing 2

cjisex2b =1 {Male}

cjisex2b =2 {Female}

cjimar2b - Marital status of applicant 2, pairing 2

cjimar2b =1 {Single}

cjimar2b =2 {Divorced}

cjimar2b =3 {Married}

cjichild2b - Number of children of applicant 2, pairing 2

cjichild2b =1 {0}

cjichild2b =2 {1}

cjichild2b =3 {2}

cjichild2b =4 {4}

cjichild2b =5 {6}

cjihome2b – Home ownership of applicant 2, pairing 2

cjihome2b = 1 {Owns own home}

cjihome2b = 2 {Lives in private rental}

cjihome2b = 3 {Lives in social housing}

cjihealth2b -Health status of applicant 2, pairing 2

cjihealth2b =1 {Pre-existing medical conditions}

cjihealth2b =2 {No medical conditions}

cjiwork2b- Employment status of applicant 2, pairing 2

cjiwork2b = 1 {Stay at home parent}

cjiwork2b = 2 {Unemployed, receiving benefits}

cjiwork2b = 3 {Unemployed, receiving no benefits}

cjiwork2b = 4 {Employed}

cjires2b – Years of UK residency of applicant 2, pairing 2

cjires2b =1 {2}

cjires2b =2 {5}

cjires2b =3 {10}

cjires2b =4 {15}

cjinat2b- County of origin of applicant 2, pairing 2

cjinat2b = 1 {India}

cjinat2b = 2 {Somalia}

cjinat2b = 3 {Jamaica}

cjinat2b = 4 {Poland}

cjinat2b = 5 {Germany}

cjinat2b = 6 {Spain}

cjinat2b = 7 {New Zealand}

cjinat2b = 8 {Mexico}

cjisex3a – Sex of applicant 1, pairing 3

cjisex3a =1 {Male}

cjisex3a =2 {Female}

cjimar3a- Marital status of applicant 1, pairing 3

cjimar3a =1 {Single}

cjimar3a =2 {Divorced}

cjimar3a =3 {Married}

cjichild3a- Number of children of applicant 1, pairing 3

cjichild3a =1 {0}

cjichild3a =2 {1}

cjichild3a =3 {2}

cjichild3a =4 {4}

cjichild3a =5 {6}

cjihome3a – Home ownership of applicant 1, pairing 3

cjihome3a = 1 {Owns own home}

cjihome3a = 2 {Lives in private rental}

cjihome3a = 3 {Lives in social housing}

cjihealth3a -Health status of applicant 1, pairing 3

cjihealth3a =1 {Pre-existing medical conditions}

cjihealth3a =2 {No medical conditions}

cjiwork3a- Employment status of applicant 1, pairing 3

cjiwork3a = 1 {Stay at home parent}

cjiwork3a = 2 {Unemployed, receiving benefits}

cjiwork3a = 3 {Unemployed, receiving no benefits}

cjiwork3a = 4 {Employed}

cjires3a – Years of UK residency of applicant 1, pairing 3

cjires3a =1 {2}

cjires3a =2 {5}
cjires3a =3 {10}
cjires3a =4 {15}

cjinat3a - County of origin of applicant 1, pairing 3

cjinat3a = 1 {India}
cjinat3a = 2 {Somalia}
cjinat3a = 3 {Jamaica}
cjinat3a = 4 {Poland}
cjinat3a = 5 {Germany}
cjinat3a = 6 {Spain}
cjinat3a = 7 {New Zealand}
cjinat3a = 8 {Mexico}

cjisex3b– Sex of applicant 2, pairing 3

cjisex3b =1 {Male}
cjisex3b =2 {Female}

cjimar3b - Marital status of applicant 2, pairing 3

cjimar3b =1 {Single}
cjimar3b =2 {Divorced}
cjimar3b =3 {Married}

cjichild3b- Number of children of applicant 2, pairing 3

cjichild3b =1 {0}
cjichild3b =2 {1}
cjichild3b =3 {2}
cjichild3b =4 {4}
cjichild3b =5 {6}

cjihome3b– Home ownership of applicant 2, pairing 3

cjihome3b = 1 {Owns own home}
cjihome3b = 2 {Lives in private rental}
cjihome3b = 3 {Lives in social housing}

cjihealth3b -Health status of applicant 2, pairing 3

cjihealth3b =1 {Pre-existing medical conditions}
cjihealth3b =2 {No medical conditions}

cjiwork3b- Employment status of applicant 2, pairing 3

cjiwork3b = 1 {Stay at home parent}
cjiwork3b = 2 {Unemployed, receiving benefits}
cjiwork3b = 3 {Unemployed, receiving no benefits}
cjiwork3b = 4 {Employed}

cjires3b – Years of UK residency of applicant 2, pairing 3

cjires3b =1 {2}
cjires3b =2 {5}
cjires3b =3 {10}
cjires3b =4 {15}

cjinat3b- County of origin of applicant 2, pairing 3

cjinat3b = 1{India}
cjinat3b = 2 {Somalia}
cjinat3b = 3 {Jamaica}
cjinat3b = 4 {Poland}
cjinat3b = 5 {Germany}
cjinat3b = 6 {Spain}
cjinat3b = 7 {New Zealand}
cjinat3b = 8 {Mexico}

Appendix B: Specification for the Body Volume Index app

See also: 9.37 Body Volume Index app and body measurements

App name: Body Volume

App icon: use the Understanding Society logo

APP STRUCTURE

- Log in screen
- Introductory pages automatically shown the first time the participant opens the app
- Landing page

LOGIN SCREEN

<Understanding Society logo>

BodyVolume

Enter Access Code

Please sign in using your unique app access code below

[Textbox]

- Continue

Pop-up error message if re-using access code:

Login Failed. The ID you have entered has already been used to complete a scan, please use a new unique identifier to log into the app.

LANDING PAGE

- Scan Body Volume
 - Video on how to take the scan
 - Profile questions
 - Body scan
 - Debrief questions
- Results
- FAQs
- Settings
 - Introductory pages (same as those shown when app first launched)
 - Tutorial Videos
 - Help
 - Privacy Policy
- Log Out

INTRODUCTORY PAGES

Screen 1

Thank you for downloading the Understanding Society BodyVolume app. We will ask you to take two photos of yourself: one from the front and one from the side. Please wear tight fitting clothes for the photos. This will only take you a few minutes to complete.

[Swipe to next screen]

Screen 2

You will see a brief video tutorial on how to take the photos. You will then be asked to report your age, gender, height, weight and activity levels, before taking the photos. At the end you will be asked a few debrief questions about how you took your photos.

[Swipe to next screen]

Screen 3

You can view your results in the Results section of the app.

For more information about this study, the data we collect, and how the body volume measures are derived, see the FAQ section. If another person in your household has an Understanding Society access code for the BodyVolume app, you can log out of the app and the other person can log in with their access code.

- Continue [go to landing page]

SCAN BODY VOLUME

Scan Type Selection

- I will scan myself → tutorial video
- Someone else will scan me → tutorial video

Your Profile

Text: Please answer these questions about yourself. The answers are needed to compute your body measures from the body scans.

Age (scroll)

Sex (Male, Female)

Height (Metric/Imperial, scroll)

Weight (Metric/Imperial/Imperial US, scroll)

Activity level (Sedentary, Active, Fit, Very Fit).

Help text: Activity Levels Explained

Sedentary: No/Low Exercise

Active: Moderate Daily Exercise

Fit: Regular Gym Use

Very Fit: Athlete or Daily Gym Use

- Continue

Pop-up: “Understanding Society” Would Like to Access the Camera

Understanding Society needs access to your camera in order to perform a body scan”

- Don’t Allow
- OK

Body scan

Starts automatically after completing the profile questions

Debrief questions

Specification notes: Allow respondents to continue without answering questions.

debriefintro (Introduction to debrief questions)

Text: Thank you for completing your body scan, to end we would like to ask you a few questions about the scan

- Continue

scan (Who took the photos)

Text: Who took the photos?

- 1 I did
- 2 Someone else

backgrd (Background in photos)

Scripting notes: Response option 7 is mutually exclusive.

Text: What was visible in the background of your photos?

Please select all that apply.

- 1 Furniture
- 2 Door(s)
- 3 Window(s)
- 4 Picture(s)
- 5 Mirrors
- 6 Something else
- 7 Nothing, empty wall

wallclr (Colour of background wall)

Text: What was the colour of the background?

- 1 White or off-white
- 2 Red
- 3 Orange
- 4 Yellow
- 5 Green
- 6 Blue
- 7 Purple
- 8 Grey
- 9 Brown
- 10 Black
- 11 Multiple colours

wallsd (Shade of background wall)

IF wallclr = 2, ..., 11 // Ask if background wall was colour other than white

Text: What was the shade of the background?

- 1 Light
- 2 Medium
- 3 Dark

clothes (Type of clothing worn)

Text: What type of clothes did you wear in the photo?

- 1 Underwear
- 2 Tight clothing
- 3 Baggy clothing
- 4 No clothes
- 5 Mixture of tight and baggy clothing

light (Lighting when photos were taken)

Text: How was the lighting when you took the photos?

- 1 Light
- 2 Medium
- 3 Dark

thankyou (Thank you for completing the measurement)

Text: Thank you for completing your body scan and the debrief questions

- Upload Results
- Update Answers to questions about the scan

Uploading your results, please wait until complete...

Success

Your results have successfully been uploaded, if you choose you can view your results on the next screen

- OK [back to landing page]

RESULTS SECTION

Keep as is in the myBVI app: total body fat, visceral fat, waist to hip – with 'help' pages explaining what these are

Error message if 'Results' clicked before taking body scan:

"Please Scan First

Please complete a body scan before attempting to view the results

- OK”

Information: Tap on the dials to see more information about your results

- OK

FAQs SECTION

How do I use this app?

After you select “Scan Body Volume” on the main menu, you will get a chance to watch a short video that will explain how you can scan yourself, or how someone else can scan you.

For more information please navigate to Settings then help.

On that website you can also see more detailed information on how to use the app.

How do I get help?

If you have technical problems with the app, then please contact Select Research at:

[email]

For questions about the study in general, then please contact the Understanding Society Participant Liaison team on:

[email]

or call them during office hours on

0800 xxx xxx

What does the BodyVolume app measure?

The app uses the outline of your body shape, plus the information about you from the ‘Your Profile’ questions (age, gender at birth, height, weight, and activity level), to calculate your waist-hip ratio, your waist circumference, total body fat, visceral fat (around your inner organs), lengths of body parts (e.g. right length of inside leg from crotch to floor), and the Body Volume Index (BVI).

The BVI is a measure of health risks based on your body shape and volume.

BVI technology is the world’s first system to measure body volume and produce the unique BVI health risk indicator number.

BVI uses patented 3D technology – the first time an affordable and easily accessible method has been available to professionals to measure body volume. Mayo Clinic developed the algorithm that provides the BVI result.

Has the Body Volume Index (BVI) received certification and review?

Yes – BVI is certified as a Class 1 Medical Device by the UK Medicines and Healthcare products Regulatory Agency (MHRA) and has achieved CE Marking throughout Europe.

In the US, the Food and Drug Administration (FDA) has reviewed BVI and considered it to be a ‘general wellness’ product allowing for widespread use amongst companies and with patients.

Does the BodyVolume app retain any of my data?

None of the images processed during the scan process are stored or sent anywhere.

The images are converted to outlines of your body shape, which are uploaded to a server in the UK, along with the information you entered on ‘Your Profile’ (your age, gender at birth, height, weight, and activity level).

The body outlines and profile information are used to calculate your body measures.

The body outlines are then deleted. The only information that is kept are your profile information, your results (the body measurements), and your responses to the questions about taking the scan.

How are my results calculated?

<link to YouTube video: https://www.youtube.com/watch?v=Vjxku_7ERiY&t=3s>

SETTINGS SECTION

Introductory Pages

Use the introductory pages shown the first time respondent launches the app

Tutorial Videos

Choose Video

Which tutorial video would you like to watch

Scan Myself Video

Scan Someone Else Video

Cancel

Help

<link to <https://www.understandingsociety.ac.uk/bodyvolume>>

Privacy policy

<link to <https://www.understandingsociety.ac.uk/bodyvolume/privacy>>

LOGOUT

Log Out?

This will log you out of the app, if you have completed your scan you will not be able to log in with your ID again and will non longer be able to view your results, are you sure you want to continue?

- Log Out [back to Enter Access Code page]
- Stay [back to Landing Page]

Screenshots from the BodyVolume app



