DOI: 10.1111/1475-5890.12357

Fiscal Studies
The Journal of Applied Public Economics

SYMPOSIUM PAPER

Sample composition and representativeness on Understanding Society

Peter Lynn | **Pablo Cabrera-Álvarez** | **Paul Clarke**

Institute for Social and Economic Research, University of Essex

Correspondence
Peter Lynn, Institute for Social and Economic Research, University of Essex, Wivenhoe Park, Colchester, Essex, CO4 3SQ, UK.
Email: plynn@essex.ac.uk

Submitted: November 2023

Abstract

In this paper, we provide an overview of the sample design of Understanding Society and the consequent nature of design weights as well as a description of procedures that are implemented in order to maximise participation by sample members and procedures that are implemented to produce non-response adjustments to the design weights. We then present some indicators of sample representativeness at the initial wave and of the impact that subsequent sample attrition has on this before concluding with some reflections on the nature of representativeness and estimation methods in the context of a highly complex sample design and complex patterns of missing data arising from non-response.

KEYWORDS

pseudo-likelihood estimation, sample design, survey-based estimation, survey non-response, weighting

JEL CLASSIFICATION

C81, C83

1 | INTRODUCTION

Understanding Society, the UK Household Longitudinal Study, is a research resource intended to be useful to a very wide range of users, most of whom are not specialist statistical analysts. Consequently, to maximise the potential value of the study, it is essential that the data can be analysed using relatively simple (model-assisted) design-based methods. With this in mind, our approach has been to use probability-based sampling for all components of the study, with known (or at least well-estimated) design weights, and to make efforts to maximise response rates and minimise non-response bias in order that the participating sample should well represent the initially selected sample. However, non-response is inevitable and does not occur at random, so non-response bias will occur. We attempt to minimise its impact on estimation by modelling the non-response process and providing users with non-response-adjusted design weights suitable for a broad spectrum of analysis purposes. Both the sample design and the techniques used to adjust design weights are rather complicated, but our intention has been to limit the extent to which users need to understand the details of design and weighting in order to utilise the data. A basic understanding is, however, necessary in order for analyses to be specified appropriately.

In this paper, we provide an overview of the sample design of Understanding Society and the consequent nature of design weights (Section 2) and of the procedures that we implement in order to maximise participation by sample members and to produce non-response adjustments to the design weights (Section 3). We then present some indicators of sample representativeness at the initial wave (Section 4) and of the impact that subsequent sample attrition has on this (Section 5) before concluding (Section 6) with some reflections on the nature of representativeness and estimation methods in the context of a highly complex sample design and complex pattern of missing data arising from non-response.

2 | SAMPLE DESIGN AND DESIGN WEIGHTING

The central component of Understanding Society is a large, equal-probability,[1] sample of the general population of the UK.[2] This is referred to as the General Population Sample (GPS). However, to afford a range of valuable analysis options that would not be possible with this sample alone, the study also combines a number of other samples of individuals that boost the sample size of important subgroups and/or add a long-term longitudinal component. These samples represent different, overlapping, subsets of the population, so this results in individuals having had multiple opportunities to be selected. For example, one sample represents persons resident in Scotland while another represents ethnic minorities, so a member of an ethnic minority living in Scotland would have had three chances of selection. Furthermore, the samples were selected at different times, so population mobility will also affect the extent to which individuals had multiple chances of selection. For example, two samples of residents of Northern Ireland have been selected, one in 2000 and one in 2009. Most persons in the first sample would have had a second chance of being selected, in 2009, but any who moved from Northern Ireland between 2000 and 2009 would not. To generalise, the overall inclusion probability of an individual is (approximately) the sum of the probability of that individual being selected into each of the component samples of the study.

An additional complicating factor is that while most of the component samples are equal-probability samples, some are not. Thus, it is not sufficient merely to know in which samples a person could have been selected; it is necessary to know their selection probability in each sample, including the ones in which they were not actually sampled. The component samples are listed in Table 1, in chronological order of selection. For each individual sample member, their selection probability for the sample within which they were selected is known by design and those for future samples can be determined from the survey data (as the data collected for each individual includes ethnic group, country of birth and location of residence at each wave). But for earlier samples, extra survey questions are required: all members of new samples at their initial wave are asked where they lived in the years when previous samples were selected. The relevant level of geography varies by sample subgroup and by year of sample. For example, sample 5 (the GPS) involved a higher sampling fraction in Northern Ireland (NI) than elsewhere in Great Britain, so members of sample 7 onwards are asked whether, in 2009, they lived in GB, NI, or elsewhere. Sample 6 – the Ethnic Minority Boost (EMB) – employed different selection probabilities in different geographical areas, depending on the estimated composition and density of the ethnic minority population in the area. This was done in order to strike an appropriate

[1] With one exception: Northern Ireland Panel Survey is over-sampled.

[2] Lynn, 2009.

TABLE 1 Component samples of Understanding Society

Sample (j)	Population covered	Year selected	Wave entered Understanding Society	Size
1. BHPS: Original	England, Scotland, Wales	1991	2	13,840
2. BHPS: Welsh Boost	Wales	1999	2	3,577
3. BHPS: Scottish Boost	Scotland	1999	2	3,395
4. Northern Ireland Panel Survey (NIPS)	Northern Ireland	2001	2	5,188
5. General Population Sample (GPS)	England, Scotland, Wales, Northern Ireland	2009	1	63,948
6. Ethnic Minority Boost (EMB)	Ethnic minorities in England, Scotland, Wales	2009	1	13,361
7. Immigrant and Ethnic Minority Boost (IEMB)	Ethnic minorities and those born outside UK: England, Scotland, Wales	2014	6	8,409
8. General Population Boost (GPS2)	England, Scotland, Wales, Northern Ireland	2022	14	~15,000

Note: 'Size' indicates the number of persons enumerated at the first wave of data collection and therefore becoming study sample members.

balance between sample efficiency (distribution of selection probabilities) and data collection costs.[3] Consequently, ethnic minority members of samples 7 and 8 were asked where in Great Britain they lived in 2009.

It should be noted that an individual's design weight is dependent on the analysis base. An analysis base is defined by the combination of waves and survey instruments that contribute data to the analysis. Instruments applied annually include the household enumeration grid, household questionnaire, individual questionnaire, 'extra 5 minutes', proxy questionnaire, self-completion component and youth questionnaire. Occasional instruments include nurse assessment and blood samples. For example, analysis of data from waves 3 and 4 can only be based on samples 1 to 6 (Table 1) so, for such analysis, the inclusion probability of each member of each of those samples is the sum of six sample-specific selection probabilities (some of which will be zero). However, for analysis of wave 14 data, members of all eight samples can be included, so each member of samples 1 to 6 will now have a larger selection probability, being the sum of eight sample-specific probabilities.[4] In general, Understanding Society design weights for analysis base k are estimated as follows. For individual i selected into sample j with probability π_{ij}, where I_{jk} is a 0/1 indicator of whether sample j contributes to analysis k:

$$w_{ijk} = \frac{I_{jk}}{\pi_{ijk}}, \text{ where } \pi_{ijk} = \pi_{ij} + \sum_{y<j} I_{yk}\hat{\pi}_{iy} + \sum_{y>j} I_{yk}\pi_{iy}.$$

This expression highlights that there is some uncertainty around the probabilities for all samples selected prior to the sample to which individual i belongs (i.e. samples for which $y < j$). For these samples, we have only an estimate, $\hat{\pi}_{iy}$, of the individual's selection probability, based on limited information collected through the survey instruments. The uncertainty relates primarily to temporality and geographical location. For example, the GPS is randomly divided into 24 monthly subsamples

[3] Berthoud et al., 2009.

[4] It should be noted that for any analysis that requires variables from wave 1 (for example, longitudinal analysis of change from wave 1 to wave $n > 1$) only samples 5 and 6 can be used, whereas for any analysis that only uses variables from wave 2 or later, samples 1 to 4 can also be included. Thus, all members of the GPS and EMB – the majority of the Understanding Society sample – will have different probabilities of selection into these two types of analysis.

(January 2009 to December 2010), so *de facto* eligibility for inclusion depends on being resident in the relevant sample month, specifically on the day the interviewer made contact at the address, whereas the residency information collected for members of later samples refers simply to '2009'. Similarly, sampling fractions for the EMB varied at the level of postcode sectors, but ethnic minority members of later samples are only asked in which local authority area they resided in 2009, so the sample-specific selection probability is approximated as the local authority mean sampling fraction.

Attempting to represent the resident population not just at one point in time, as cross-sectional surveys do, but at all points in time over a long period, is particularly challenging and raises some coverage issues.[5] In principle, the study represents all persons who meet the eligibility criteria for at least one of the component samples – and all those descended from them. But this leaves a few inconsistencies over time. Each component sample is initially restricted to the household population, excluding persons resident in institutions. Over time, in principle, the study follows sample members if they move into an institution (though there are often practical barriers that make this impossible). Thus, for example, coverage of residents of institutional accommodation for the elderly is initially absent but then increases over time. People who have spent some time temporarily in an institution (e.g. prison, armed forces accommodation) will tend to be under-represented, as one or more of their component selection probabilities may have been zero. Recent immigrants to the UK should be well-represented initially in each sample, but coverage will then gradually reduce (as subsequent immigrants are not added to the sample) until the next relevant sample is added. For example, coverage of recent immigrants declines from waves 1 to 5, is then restored at wave 6, declines again from waves 7 to 13 and is restored again at wave 14.

3 | RESPONSE MAXIMISATION AND NON-RESPONSE WEIGHTING

For all surveys, non-response introduces the potential for bias to affect estimation. This will occur if non-respondents are systematically different from respondents in relevant ways.[6] There is much that survey researchers can do to mitigate this through good design and effective survey procedures.[7] For longitudinal surveys, high wave-on-wave response rates are important independently of whether non-response introduces bias, as large sample sizes with repeated measures are required to support longitudinal analysis. For these reasons, building on the pioneering work of the British Household Panel Study (BHPS),[8] Understanding Society has always devoted considerable resources to maximising and equalising response rates. A wide range of techniques can be used to minimise panel attrition[9] and those implemented on Understanding Society include the use of respondent incentives, multiple modes of data collection including face-to-face, extensive reminders and call-backs, respondent-centred design of survey documents and branding,[10] and sensitivity to respondent concerns (e.g. use of self-completion for sensitive questions, a confidentiality pledge, professionalism and flexibility of interviewers to schedule/reschedule appointments). Procedures aimed primarily at equalising response propensity across subgroups include use of translation and multilingual interviews,[11] monitoring and management of fieldwork by key subgroups, targeting of incentives based on previous outcomes, targeting of invitation letters,[12] and targeted use of SMS messages.[13]

[5] Lynn, 2011.

[6] Groves et al., 2002; Lynn, 2008.

[7] Bethlehem, Cobben and Schouten, 2011.

[8] Laurie, Smith and Scott, 1999.

[9] Lynn, 2018.

[10] See, for example, Carpenter (2016).

[11] Lynn et al., 2018.

[12] Lynn, 2016.

[13] Cabrera-Álvarez and Lynn, 2023a.

Survey procedures have continually evolved in response to regular reviews of practice and evidence, much of which has come from the study's own extensive programme of methodological research into the nature of non-response and how it can be influenced through design, often involving large-scale experiments. Design features whose effects on non-response – either to the study as a whole or to study components – have been studied include the use of telephone interviewing,[14] introduction of web-based data collection,[15] interviewer continuity,[16] extended interviewer efforts,[17] interview length,[18] the inclusion of motivational statements in web questionnaires,[19] targeted respondent mailings,[20] use of negative framing in appeals to participate,[21] protocols for targeted assignment of respondents to data collection modes,[22] push-to-web recruitment methods,[23] contact sequences and early-bird incentives,[24] the timing of web survey invites,[25] ways of inviting respondents to a mobile app study,[26] the effects of providing personal feedback,[27] changing the value of incentives,[28] and the use of email messages[29] or SMS messages[30] for survey invitations and reminders.

Subsequent to each wave of data collection, sets of non-response-adjusted weights are produced and provided to users in each annual data release. The approach used to make the adjustments is to model each step in the attrition process (wave, instrument) sequentially and apply a series of multiplicative adjustments[31] to the design weights. A wide range of auxiliary variables is used at each level (household, individual, instrument) and at each stage in the survey process, with predictors selected through a stepwise procedure (aside from ethnic group, which is automatically retained in all models that include the EMB). The set of potential predictors available varies between steps in the process, generally increasing over waves as more survey data are collected. The broad strategy is summarised in Lynn and Kaminska (2010), though several extensions and amendments to the strategy have been developed over the years in response to changes in study design and data collection protocols and in patterns of non-response. Additionally, at each wave, subsequent to the non-response adjustment, post-stratification is carried out by age band, sex and country. This includes an adjustment for undetected mortality, essentially by applying published mortality rates by age and sex to the initial sample to produce the post-stratification targets.

A sizeable number of different sets of weights are produced, for use with a range of analysis bases defined by combinations of waves and survey instruments. Instruments applied annually include the household enumeration grid, household questionnaire, individual questionnaire, 'extra 5 minutes', proxy questionnaire, self-completion component and youth questionnaire. Occasional instruments include nurse assessment and blood samples.

[14] Lynn, 2013.

[15] Jäckle, Lynn and Burton, 2015; Bianchi, Biffignandi and Lynn, 2017.

[16] Lynn, Kaminska and Goldstein, 2014.

[17] Hall et al., 2013.

[18] Lynn, 2014.

[19] Al Baghal and Lynn, 2015.

[20] Fumagalli, Laurie and Lynn, 2013; Lynn, 2016, 2017.

[21] Lynn, 2019.

[22] Kaminska and Lynn, 2017.

[23] Lynn, 2020.

[24] Williams et al., 2022.

[25] Lynn, Bianchi and Gaia, 2023.

[26] Jäckle et al., 2022.

[27] Wenz et al., 2022; Benzeval et al., 2023.

[28] Cabrera-Álvarez and Lynn, 2023b.

[29] Cernat and Lynn, 2018.

[30] Cabrera-Álvarez and Lynn, 2023a.

[31] Lynn and Watson, 2021, section 18.4.

By wave 13, over 300 different weight variables had been provided in the main annual data releases. To enable users to easily identify the appropriate weight for a particular form of analysis, the weight variables are named following a convention that indicates the unit of analysis, survey instrument, component samples, and wave (cross-sectional analysis) or set of waves (longitudinal analysis). Extensive guidance is also provided through online user guidance[32] and training materials. Even with such a large number of weights made available, some users will still find that there is no weight tailored to the specific analysis they wish to carry out. To address this eventuality, guidance is provided on how to select an appropriate, albeit suboptimal, weight, or alternatively how a user can derive their own analysis-specific weight by adjusting one of the provided weights.[33]

4 | SAMPLE REPRESENTATIVENESS AT THE INITIAL WAVE

Understanding Society data allow the scientific community to address research questions related to social change, such as how behaviours and attributes evolve within individuals over time and the factors associated with these shifts. These analyses rely on the longitudinal design of the study, requiring repeated measures from the same individuals over time. However, some sample members did not respond to the initial interview, whilst others dropped out over subsequent waves, which could harm sample representativeness and affect the quality of survey estimates. In this section, we address the first concern, the lack of response at the initial wave, by providing some indicators of the representativeness of samples 1 and 5 (Table 1) – the BHPS original sample and the GPS – at their initial waves, respectively in 1991 and 2009–11. These indicators compare the survey profile of each sample to Census-based estimates from 1991 and 2011, respectively.

We study these two component samples because they constitute the largest parts of the sample and because they cover the household population of Great Britain, in the case of the BHPS, and of the United Kingdom, in the case of the GPS. Furthermore, the data collection of the 1991 Census, which took place in April, and the fieldwork of the initial wave of the BHPS, from September to December, occurred close to each other. Regarding the GPS, the time difference is greater because the first wave of fieldwork spanned from January 2009, with the last few cases being interviewed in March 2011, while 2011 Census fieldwork was in April.

Nonetheless, the interpretation of this analysis needs to consider some constraints. First, the Census covers people living in communal establishments[34] in addition to the household population, whilst the BHPS and GPS only cover the former. This inconsistency in the definition of the target populations could result in slightly different estimates, especially for some subgroups, such as older adults, who are more likely to live in communal establishments. Second, the Census and Understanding Society methodologies diverge from each other. For example, the wording of some questions, such as the one asking about ethnicity, differed. Also, the Census makes a more profuse use of proxy informants, where a household member reports information about the rest of the household, which could affect measurement. A more detailed description of the analysis methodology can be found in Lynn and Borkowska (2018).

Table 2 compares the Census-based estimates and the sample profile of wave 1 respondents for the BHPS original sample and the GPS regarding sex, age and ethnic background. The results show that the sample profile of the BHPS at the initial wave is broadly similar to the Census figures. The only slight difference is observed for the group aged 30–39, which is over-represented by 1.8 percentage

[32] See the Understanding Society webpage 'Selecting the correct weight for your analysis', https://www.understandingsociety.ac.uk/documentation/mainstage/user-guides/main-survey-user-guide/selecting-the-correct-weight-for-your-analysis.

[33] See the Understanding Society webpage 'Deriving your own weights', https://www.understandingsociety.ac.uk/documentation/mainstage/user-guides/main-survey-user-guide/deriving-your-own-weights.

[34] The Census estimate of people living in communal establishments was 2.8 per cent in 1991 in Great Britain and 1.8 per cent in 2011 in the United Kingdom.

TABLE 2 BHPS original sample and GPS wave 1 estimates compared to Census estimates for sex, age and ethnic group

	Population Census (1991)	BHPS wave 1 (1991)	Population Census (2011)	GPS wave 1 (2009–10)
Sex				
Male	47.7	46.4	48.6	45.4
Female	52.3	53.6	51.4	54.6
Age at wave 1				
16–19	6.6	6.6	6.3	6.2
20–29	19.3	18.9	16.8	14.8
30–39	17.4	19.2	16.2	16.9
40–49	16.9	18.5	18.1	18.9
50–59	13.3	12.7	15.0	15.7
60–69	12.7	12.0	13.3	14.4
70+	13.9	12.1	14.3	13.1
Ethnic group				
White	95.4	96.0	88.7	91.0
Black	1.4	1.4	2.7	2.3
South Asian	–	–	4.4	3.7
Other	3.2	2.6	4.2	3.0

Note: Column percentages. For the BHPS, the survey estimates are based on persons with a completed interview at wave 1 ($n = 9{,}897$). Ethnic group estimates based on $n = 9{,}878$ cases due to item missingness. Population figures combine data from the 1991 England and Wales Census and the 1991 Scottish Census (persons 16 and older). For the GPS, the survey estimates are based on adults (16 and older) who completed the adult individual or proxy interview at wave 1 ($n = 43{,}674$). The question about ethnic background was not in the proxy questionnaire, and the estimates were computed using only those who completed the individual interview ($n = 41{,}047$). Survey estimates are weighted by the design weight to account for unequal probabilities of selection. Population figures combine data from the 2011 England and Wales Census, 2011 Scottish Census and 2011 Northern Ireland Census. Sex estimates are limited to the resident population in Great Britain, as no sex breakdown was published for persons aged 16 and over in Northern Ireland.

points (p.p.), while those aged 70 or older are under-represented by the same amount. The sample distribution of the GPS is also close to the population estimates from the 2011 Census. In this case, males and those aged 20–29 are slightly under-represented by 3.2 p.p. and 2.0 p.p., respectively. An extended version of this analysis, including the variables Government Office Region (GOR), whether the person suffered a limiting long-term illness, economic activity and car or van ownership, only found a substantial difference for residents in Greater London, who were under-represented by 1.8 p.p. in the BHPS original sample and by 3.4 p.p. in the GPS.[35]

5 | ATTRITION

We now turn to panel attrition, which occurs when a sample member drops out of the study after responding to the initial interview. Panel attrition entails two main risks for survey estimates. First, some sample members might be more likely to drop out than others based on their characteristics, affecting the sample profile and increasing the risk of bias in the survey estimates. Second, the erosion of the sample size and the subsequent loss of statistical power can damage the precision of survey estimates, especially when analysing sample subgroups. Here we examine some indicators of the general level of attrition and representativeness of the samples forming Understanding Society over

[35] Lynn and Borkowska, 2018.

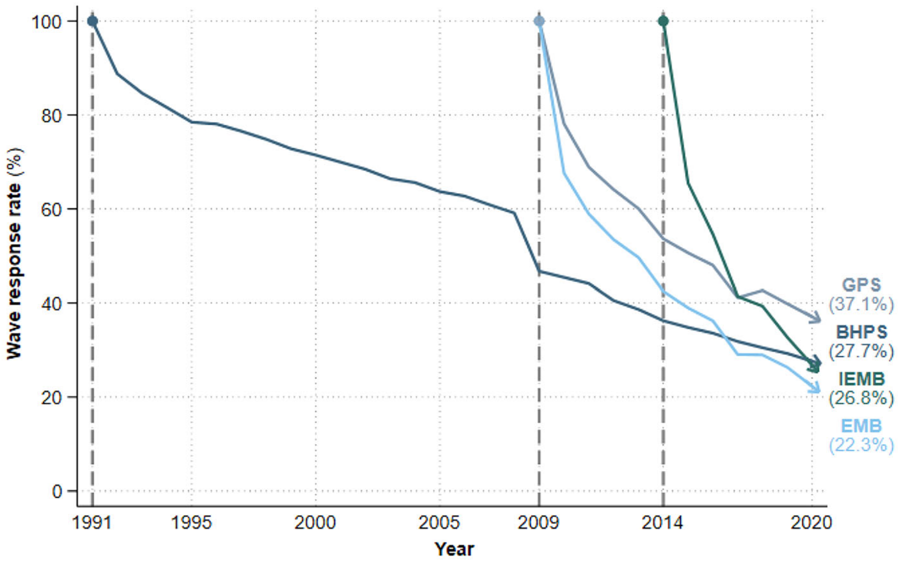

FIGURE 1 Wave response rates over time

Note: The base for the analysis is the respondents to the individual interview at the initial wave. Response rates are calculated for each wave accounting for changes in the eligibility status of panel members. Respondents include panel members completing the individual interview and proxy respondents.

time, and we provide some evidence of the ability of survey weights to mitigate the effect of non-response and attrition.

The first analysis we present in this section shows the wave response rates trend across the four main samples that form Understanding Society: the BHPS original sample, the GPS, the EMB sample, first interviewed in 2009–11, and the Immigrant and Ethnic Minority Boost (IEMB) sample, recruited in 2014–16. The computation of the wave response rates relied on several methodological decisions that should be considered when interpreting the analysis. First, the definition of respondent, which is relevant to establish the numerator of the response rate, encompasses both panel members completing the individual questionnaire and proxy respondents, for whom another household member responded. Second, the denominator of the rate consists of the adults – aged 16 and over – responding to the initial wave, to enable assessment of attrition over time. Then, in each wave, we excluded the panel members known to have become ineligible (moved abroad, moved into an institution or died). However, establishing eligibility status is challenging for panel members not contacted or those that dropped out of the study in the previous waves. To alleviate this issue, we relied on administrative records to detect hidden mortality in the four samples, and implemented an adjustment that captured the probability that a non-responding sample member was dead.[36] Third, these figures are sample-based estimates subjected to sampling error, and thus minor differences do not necessarily indicate bias.

Figure 1 presents cumulative response rates at each wave for the four samples. The BHPS original sample had a 78.1 per cent response rate after six years and 27.7 per cent in 2020. Compared to this, the GPS is eroding faster, with 53.7 per cent of sample members participating after six years and 37.1 per cent in 2020. For the boost samples (i.e. the EMB and IEMB), wave response rates are notably lower than those observed for the GPS. For instance, the EMB exhibits a 22.3 per cent response rate after 11 years, 14.8 p.p. lower than the GPS after the same time period. The most recent sample, the IEMB, has a 26.8 per cent response rate six years after the first interview.

[36] Kaminska, 2021.

TABLE 3 BHPS, GPS and EMB cumulative response: sex, age at wave 1 and ethnic group

	British Household Panel Survey original sample		**General Population Sample**		**Ethnic Minority Boost**	
	BHPS wave 1 (1991)	**wave 12 (29th wave) (2020–22)**	**wave 1 (2009–11)**	**wave 12 (2020–22)**	**wave 1 (2009–11)**	**wave 12 (2020–22)**
Full sample	10,264	27.7	43,673	37.1	6,623	22.3
Sex						
Male	4,833	26.1	19,773	35.7	3,128	20.7
Female	5,431	29.0	23,900	38.2	3,495	23.7
Age at wave 1						
16–19	696	24.4	2,700	22.3	647	15.7
20–29	1,960	28.4	6,389	26.5	1,602	18.1
30–39	1,972	33.1	7,408	35.7	1,727	23.6
40–49	1,877	35.4	8,266	41.0	1,277	27.6
50–59	1,298	23.1	6,891	48.1	712	24.7
60–69	1,213	7.8	6,287	46.5	363	27.5
70+	1,248	0.3	5,732	25.3	295	13.9
Ethnic group						
White	9,503	28.5	39,663	38.5	25	66.7
Black	138	9.5	940	17.8	1,880	18.6
South Asian			1,639	26.4	3,326	23.8
Other	252	25.5	1,385	23.9	1,392	23.0
Income						
Bottom	2,053	22.2	8,788	29.8	2,187	18.8
Second	2,053	20.6	8,729	32.6	1,296	20.3
Third	2,054	25.6	8,747	35.2	1,090	25.1
Fourth	2,052	30.2	8,720	39.7	1,084	24.5
Top	2,052	36.8	8,689	47.5	966	27.4

Note: The first column of each panel shows the panel members responding to wave 1 individual interview (full response or proxy) for each subgroup, while the second column shows the percentage of those interviewed at wave 12 of Understanding Society. Ethnic group was not included in the proxy questionnaire, and the analysis for these variables only used the sample of panel members completing the wave 1 interview. Income quintiles were calculated by combining the GPS and EMB sample and using the design weight, which explains the variation in the number of cases across quintile groups. For the EMB sample, the first column only includes the respondent to the initial interview who self-identified as part of an ethnic minority. The few sample members identified as White in the ethnic group variable ($n = 25$) correspond to individuals who changed their subjective self-identification after wave 1.

Table 3 explores the nature of attrition at wave 12 of Understanding Society (2020–22) across sample subgroups defined by sex, age, ethnicity and individual income for the BHPS original sample, the GPS and EMB sample. To be concise, we omit the IEMB from this analysis and point the reader to a more detailed study addressing attrition in this sample.[37] The response rates were similar for male and female sample members, with the latter showing only slightly higher rates. In the GPS, 38.2 per cent of females and 35.7 per cent of males were still participating after 12 annual waves, a difference of 2.5 p.p. Differences of a similar magnitude were observed in the BHPS (2.9 p.p.) sample after 29 years and in the EMB (3.0 p.p.) after 12 waves.

[37] Cabrera-Álvarez, James and Lynn, 2023.

Next, we examine the attrition rate based on age. It is important to note that age is a time-variant factor and the process of ageing itself may affect the likelihood of panel members dropping out of the study. For example, those aged 16–19 when the BHPS started in 1991 were 45–49 years old at the time of wave 12 (2020–22). In the cases of the GPS and EMB sample, panel members aged 16–19 in 2009, when the study started, were 27–30 years old in 2020–22. The decline in response rates was more pronounced for the younger and older groups of the sample at the initial wave, but for different reasons. For the youngest cohort (16–19 years old), the BHPS response rate after 29 waves was 24.4 per cent, 3.3 p.p. lower than the average. In the GPS, after 12 waves, this difference was more considerable, 14.8 p.p. below the average. In the EMB sample, the 16–19 cohort was also under-represented at wave 12, but the difference from the average response rate was 6.6 p.p. The lower response rate among younger members in the years following the initial interview finds two plausible explanations. First, they are more likely to move out of the original household, thereby complicating the location and contact processes. Second, younger adults have a lower propensity to respond to surveys in general.[38]

Sample members aged 70 and older at the initial wave also exhibit a lower response rate than average. In the case of the BHPS, the response rate is 0.3 per cent after 29 waves, whilst in the GPS, after 12 waves, it is 25.3 per cent, and 13.9 per cent in the EMB sample. The effect of ageing explains the lower response rate of this group. For example, in the BHPS, people aged 70 and older in 1991 would be 99 years old and over in 2020. Therefore, a substantial part of this age group has become ineligible over time – most will have died, and others will have moved to institutions – although these changes may have gone unnoticed. Also, the sample members still eligible for an interview are more likely to suffer health impairments that hinder survey participation.

The age group with the highest response rate varies across samples, but in each case broadly corresponds to those aged around 60–79 at the time of wave 12. In the BHPS, panel members aged between 30 and 49 in 1991 showed the highest response rates (aged 30–39, 33.1 per cent; aged 40–49, 35.4 per cent). In the GPS, panel members aged 50–69 at wave 1 have the lowest level of attrition by wave 12 (response rates: aged 50–59, 48.1 per cent; aged 60–69, 46.5 per cent). In the EMB sample, the cohorts aged 40–49 and 60–69 exhibit the highest response rates, 27.6 per cent and 27.5 per cent, respectively.

Panel members with an ethnic minority background have a consistently lower response rate compared to White people in the BHPS original sample and the GPS. The cumulative response rate of White panel members in the BHPS was 28.5 per cent, while for Black panel members it was 9.5 per cent. In the EMB sample, the ethnic minority groups present similar attrition to the same groups from the GPS. For example, the wave 12 response rate for Black panel members is 17.8 per cent in the GPS and 18.6 per cent in the EMB sample; in the case of South Asian panel members, the response rates are 26.4 per cent and 23.8 per cent, respectively.

Individuals in the top income quintile have a higher response rate than those in the bottom quintile across samples. In the BHPS sample, the bottom quintile had a 22.2 per cent response rate after 29 years, whilst the top quintile exhibits a 36.8 per cent response rate, 14.6 p.p. difference. In the case of the GPS, the difference between the top and bottom quintiles is 17.7 p.p.; for the EMB sample, the gap is narrower at 8.6 p.p.

In the third part of the analysis, we examine the performance of the longitudinal weights in adjusting the sample of panel members responding to all waves up to wave 12 for non-response and panel attrition with regard to sex, age, ethnicity and income. To do this, we separate the BHPS original sample, and the combination of the GPS and EMB sample, both first interviewed in the initial wave of Understanding Society (2009–11). For each sample, we present three columns in Table 4. The first column approximates how the sample initially interviewed would look at wave 12 of Understanding Society (2020–22) in the absence of panel attrition, the comparison standard for the analysis. The base

[38] Olson and Witt, 2011; Watson, 2003.

TABLE 4 Profiles for the BHPS and for the GPS and EMB sample combined, for wave 1 respondents eligible for an individual interview at wave 12 and respondents to wave 12 and all the previous waves, unweighted and weighted estimates

	Panel 1: BHPS original sample			**Panel 2: GPS and EMB sample combined**		
	BHPS wave 1 (1991)	**Understanding Society wave 12 (2020–22)**		**Understanding Society wave 1 (2009–11)**	**Understanding Society wave 12 (2020–22)**	
		Unweighted	**Weighted**		**Design weighted**	**Weighted**
Sex						
Male	47.8	43.3	48.2	48.2	42.3	48.2
Female	52.2	56.7	51.8	51.8	57.7	51.8
Age at wave 1						
16–19	8.7	5.7	9.9	6.9	2.2	5.9
20–29	26.0	23.0	29.5	18.0	9.6	18.4
30–39	22.6	28.3	26.7	17.5	15.9	18.1
40–49	19.5	30.2	21.7	19.8	21.5	21.2
50–59	12.2	10.7	9.7	15.8	22.8	17.3
60–69	7.0	2.1	2.6	12.9	21.0	13.1
70+	4.0	0.0	0.0	9.1	7.0	5.9
Ethnic group						
White	94.7	97.7	96.6	90.4	95.6	90.1
Black	2.0	0.2	0.2	2.3	0.8	2.6
South Asian				4.2	1.9	4.2
Other	3.3	2.1	3.2	3.1	1.7	3.1
Income						
Bottom	18.2	15.6	16.6	20.6	14.7	19.1
Second	15.9	11.8	13.2	18.2	17.1	17.3
Third	19.3	17.4	19.5	19.3	19.3	19.7
Fourth	23.1	24.1	24.7	20.5	22.0	21.1
Top	23.5	31.1	26.0	21.4	27.0	22.8
Unweighted base	7,363	1,373	1,373	44,970	10,924	10,924

Note: The first column of each panel shows the percentage of each group of the panel members responding to wave 1 individual interview (full response in case of the BHPS original sample and full response or proxy in the case of the GPS and EMB samples) who were eligible for an individual interview at wave 12. The second column shows the sample profile of those responding to wave 12 of Understanding Society who responded to all previous waves, which corresponds to the base used to generate the longitudinal weight. The third column shows the sample profile of the respondents to wave 12 and all the previous waves weighted by the longitudinal weight. Income quintiles were calculated by combining the GPS and EMB sample and using the design weight.

for these estimates is the sample members responding[39] to the initial wave of the study who are still eligible for an interview at wave 12. This means that panel members known to have moved abroad or died were excluded from the analysis. However, as described above, some moves into ineligibility were not detected and will therefore be erroneously included in the base. This might affect estimates

[39] The definition of respondent for the BHPS excludes the proxy responses as there is no weight available for this sample. Therefore, the analysis for this sample has been done using only the sample of respondents to the initial interview. The definition of respondent in the analysis of the GPS and EMB sample covers those who completed the adult interview and the proxy respondents.

for some groups, such as older adults. We use the wave 1 cross-sectional weight to adjust for the effect of non-response because this constitutes the base for the correction of attrition made by the longitudinal weight used in the analysis.

The second and third columns of each panel in Table 4 have a common base, sample members who responded to all waves up to wave 12 of Understanding Society (2020–22), the same base used to compute the longitudinal weight tested in this analysis. For the BHPS original sample, the second column shows the unweighted sample profile, and the third column contains the sample profile weighted by wave 12 longitudinal weight.[40] This weight adjusts the sample for non-response at wave 1 and attrition. For the GPS and EMB sample, the second column shows the sample profile weighted to account for the unequal probabilities of selection, while the third column shows the weighted estimates by the wave 12 longitudinal weight, which adjusts the sample for the unequal probabilities of selection, non-response at the initial wave and attrition.

The longitudinal weights effectively deal with non-response and panel attrition and restore the sample profile, although we observe some differences. In the BHPS original sample (Table 4, panel 1), the unweighted sample of those who participated in all 29 waves contains 43.3 per cent males whilst, in the absence of attrition, the proportion should have been closer to 47.8 per cent. This difference is due to the slightly higher response rate of females at the initial and successive waves. The longitudinal weight adjusts the sample for attrition and consequently restores the sample profile to 48.2 per cent males. Regarding age, the longitudinal weight is able to restore the sample profile. An apparent exception for those aged 60 and over is undoubtedly due to undetected changes in the eligibility status of the panel members. Panel members with an ethnic minority background are more likely to drop out. The longitudinal weight effectively adjusts the wave 12 sample, except for the case of Black panel members. The unweighted estimate for Black persons at wave 12 is 0.2 per cent, a figure that remains unaltered after applying the longitudinal weight versus the 2.0 per cent of the initial sample (though that estimate too could be upwardly biased due to undetected moves into ineligibility, as both mortality and emigration are differential by ethnic group). In terms of income, the wave 12 longitudinal weight broadly corrects the estimates for attrition, although panel members in the higher income quintiles remain slightly over-represented.

After 12 waves, attrition has changed the sample profile of the GPS and EMB sample with regard to the variables examined in Table 4 (panel 2). Yet, the wave 12 longitudinal weight successfully corrects the bias in the survey estimates. As with the BHPS original sample, the sample distribution of sex is affected by attrition – males are more likely to drop out of the study. While males are 48.2 per cent of the initial sample members assumed to be eligible for an interview at wave 12, this percentage is 42.3 per cent amongst those who responded at all 12 waves. The wave 12 longitudinal weight adjusts the percentage of males back to 48.2 per cent. The longitudinal weight is also effective in adjusting the sample distribution of age groups. However, for the oldest group (70 and over), the weighted estimate (5.9 per cent) is far from the percentage of this group in the initial sample (9.1 per cent). This difference is most likely mainly due to undetected ineligible cases in the initial sample. The longitudinal weight effectively adjusts the deviations due to attrition in the ethnicity and income variables.

[40] To correct the sample profile of the BHPS original sample, we used the ba_indin91_xw. Then, to correct the sample of respondents to all waves up to wave 12 we used the l_indin91_lw weight. The objective of this weight is to correct for non-response at wave 1 and attrition. Thus, the resulting sample profile should be comparable to the Great Britain population of 1991 that was alive in 2020–22. The analysis of the GPS and EMB sample uses the a_indpxus_xw weight to adjust the sample of respondents to the initial wave. Thus, the sample of respondents to all waves up to wave 12 is weighted by the design weight (a_psnenus_xd) to account for the differential probabilities of selection, and by the longitudinal weight (l_indpxus_lw).

6 | REFLECTIONS ON REPRESENTATIVENESS AND APPROACHES TO ESTIMATION

There are various approaches to data analysis when survey data are not representative of the population because of the sampling design and non-response. In this section, we summarise the issues as they apply to Understanding Society.

6.1 | Complex sampling designs

Complex sampling designs can complicate the process of analysing data by preventing us from treating the variables in the model as independent realisations of random variables following the unknown population data-generating process. Unless the individuals constituting the analysis sample were explicitly selected using simple random sampling – which is patently not the case for any subset of Understanding Society data – the argument below is relevant.

We argue that the simplest and most robust way to account for the effects of a complex sampling design is to use pseudo-likelihood estimation, which weights the analysis you would have carried out were the data drawn using a simple random sampling design, and then adjusts the standard error estimates for the effects of stratification and clustering. This approach has the advantages of:

- requiring no elaboration of the analyst's existing identification and modelling assumptions;
- handling the impact of both the complex sampling design and the survey non-response;
- being relatively simple to implement because the survey weights, strata identifiers and primary sampling unit (PSU) identifiers are routinely available from Understanding Society.

We illustrate the argument in the context of cross-sectional analysis based on data from one wave of Understanding Society. We defer discussion of how this generalises to longitudinal analysis to Section 6.3.

6.1.1 | The population and sampling design

Suppose our sample is drawn from population $P = \{1, \ldots, N\}$. This set indexes each of the N individuals (or units of analysis) present in the population at the time sampling took place. The sampling design of Understanding Society does not involve directly sampling individuals from a sampling frame listing the individuals in P. For instance, the GPS in Great Britain is (in the main) a stratified two-stage sample, with postcode sectors selected at stage one followed by residential addresses at stage two; if there are more than three dwellings or households at a single address, then a third sampling stage involves choosing three of these at random.

All individuals resident in the dwelling or household are included in the sample and, therefore, listed in the household grid. In Northern Ireland, the GPS is based on a simple random sample of residential addresses. Stratification of the British sample involves the nations of the United Kingdom and, within England, the Government Office Regions; further stratification of the postcode sectors is by the proportion of non-manual workers, population density, and the density of (non-White) minority ethnic groups. These variables, including the number of households/dwellings at the selected address, are known prior to the sample being drawn and are referred to as 'design variables'.

Finally, Understanding Society has available the sampling weights, or design weights, for each individual. These are based on the inverse of the selection probabilities, which are known or estimated for each individual in the population based on the design variables (see Section 2). Note that the design weight is a key component of the overall survey weight, which is what is used in practice in analysis, as discussed further on.

6.1.2 | Survey variables

Once the sample is drawn, the study measures a range of survey variables on the sample members. Analysts choose subsets of these variables for their analyses based on their research questions. In each case, the analyst partitions the chosen variables into an *outcome* Y and *predictor*s X; the predictors can include the characteristics of higher-level units defined by social, geographical, etc., structure as well as important identifying exogenous variables such as instrumental variables. The analyst will then make identification assumptions and, through the choice of model and estimation method, other econometric modelling assumptions to enable the inferences about the population parameters needed to answer the research questions.

The important thing to bear in mind is that the model specification and its attendant assumptions are all made at the population level not the sample level, so should ideally be unaffected by the sampling design.[41] For example, from any econometric textbook, consider the linear regression model

$$y_i = \mathbf{x}_i\boldsymbol{\beta} + e_i, \tag{1}$$

where row vector $\mathbf{x}_i$ includes a constant term and determines the effects of predictors X on outcome Y, and column vector $\boldsymbol{\beta}$ contains the coefficients of these effects for individual i. The basic modelling assumptions here are $E(y_i|\mathbf{x}_i) = \mathbf{x}_i\boldsymbol{\beta}$ and $E(e_i) = 0$ and both are defined at the population level. These are sufficient for descriptive analysis, but it is more usual to require that $E(e_i|\mathbf{x}_i) = 0$ for causal analysis, where $\mathbf{x}_i$ can be extended to include exogenous instrumental variables not specified in equation (1). The choice of ordinary or general least squares is made for efficiency reasons, with general least squares likely to give smaller standard errors if the residual variance $\text{var}(e_i|\mathbf{x}_i) \neq \text{var}(e_i)$, that is, another population-level assumption.

In practice, we use the sample data to make inferences about the population, and complex sampling designs complicate this such that we cannot use standard methods.[42]

6.1.3 | Pseudo-likelihood estimation

The approach to estimation we recommend is referred to as pseudo-likelihood in the survey sampling literature.[43] It works by combining the robust properties of the randomisation scheme used for sample selection together with the analyst's specific modelling assumptions. Pseudo-likelihood is appealing because:

- it requires the same identification and modelling assumptions as would have been required were the data drawn using a simple random sampling design;
- it requires only the sampling weights and information about strata and cluster membership;
- it is straightforwardly extended to accommodate adjustments for non-random non-response on the analysis through the use of the survey weights rather than sampling weights.

The principle is a simple one: the estimating equation based on the sample data is not unbiased for the population-level estimating equation, so it must be weighted to ensure that it is. The solution of

[41] In the survey sampling literature, this is called the superpopulation level where the parameters of the superpopulation model are distinguished from those of the finite population – here the UK population targeted by the sampling design. A population parameter in this case can be thought of as a summary statistic based on everyone sampled and everyone not sampled. This distinction is, however, not needed in this discussion.

[42] There are certain situations where explicit adjustments for the complex sampling design are unnecessary. One is that sample selection is conditionally independent of the design variables given the predictors in X and has no parameters in common with the conditional distribution of Y given X, and that likelihood-based estimation is used. However, we view this situation as very unlikely to arise in practice so do not consider it in detail.

[43] See section 3.4.4 of Skinner, Holt and Smith (1989).

the weighted estimating equation, the pseudo-likelihood estimate, is therefore consistent for the true population parameter.

Complications do arise when fitting multi-level, or hierarchical, cross-sectional random effects models that include random effects defined at higher levels. Pseudo-likelihood estimates of parameters such as random effect variances, or other parameters defined at higher levels (but not the regression coefficients even those for higher-level contextual effects), will generally be biased if these are structurally related to one of the stages of the sampling scheme. In such cases, bias can be corrected by incorporating selection probabilities for higher-level units in the sampling scheme into the estimation procedure.[44] However, no general solution has yet been proposed.

6.1.4 | Standard error estimation

Simply weighting the usual estimating equations is sufficient to adjust the parameter estimate for the complex design, but what about the estimated standard errors? A relatively simple way to do this is to use linearised variance estimation. This method does not work universally well for all surveys but is suitable for Understanding Society because the proportions of postcode sectors included in the sample within each of the British regions used to stratify the sample are relatively small.

The linearised estimator is based on weighted estimates of the two distinct components of the usual 'sandwich' variance estimator usually used to obtain standard error estimates of the model parameters.[45] Another plus of linearised estimation, for multi-stage stratified designs, like that of the GPS, is that accounting for clustering requires only including identifiers for the postcode-sector PSUs (the lower-level clustering induced by secondary, etc., sampling of addresses does not affect the first-order variance estimator).

6.2 | Non-response

Survey non-response includes unit non-response, wave non-response and item non-response, so can be viewed as the final stage of sample selection, albeit one that is not under the control of the sampler. This makes it straightforward to extend pseudo-likelihood estimation and linearised standard error estimation to include adjustments for complex sampling design and non-response. We just need estimates of the non-response probabilities in order to produce the survey weights to weight the estimating equations.[46]

As outlined in Section 3, Understanding Society weight adjustments are based on a sequence of models of non-response at each step in the process. Thus, for example, auxiliary variables to explain household-level non-response at wave 1 include a range of small-area indicators such as the percentage of households in the area that own their home, that have no car or that consist of a single pensioner and the percentage of persons with no qualifications, who are full-time employees or have a limiting long-term illness.[47] Auxiliary variables to explain individual participation conditional on household participation include individual-level indicators obtained in the household grid (e.g. age, marital status, employment status, ethnic group) and household-level measures from the household

[44] See, for example, Pfeffermann et al. (1998).

[45] More accurately, the estimate of the variance of the score (estimating equations), i.e., the 'filling' of the sandwich, is based on a design-consistent estimator of its population value. The 'bread' of the sandwich is the average derivative of the score function which is simply replaced by a weighted mean of the sample scores.

[46] The only scenario under which non-response weighting is unnecessary is if non-response among the sampled individuals is independent of the outcome and auxiliary variables used to estimate the non-response weights. We again consider such a scenario to be highly unlikely and so do not consider it further.

[47] Lynn et al., 2012.

questionnaire such as number of adults in the household, housing tenure, council tax band, and whether the household is behind with bill payments. For subsequent steps in the process, conditional on completing the wave 1 individual interview, all variables collected in that interview could be used as auxiliary variables.

In reality, the survey non-response processes will differ for the GPS, BHPS and boost samples, and by population subgroup and allocated data collection mode protocol. However, the focus of modelling is on the compound process alone: the issue for model building is to ensure predictors of all the processes are included in the model, even if the resulting model does not necessarily help us to understand the component processes.

Non-response weights are estimated, so, ideally, one should account for this when estimating standard errors.[48] However, this is rarely done in practice.

6.3 | Longitudinal

For longitudinal data, pseudo-likelihood estimation for analyses involving either (a) repeated measures of Y or (b) X that comprises variables from different study waves, or both, involves choosing the appropriate longitudinal weight. The discussion in Sections 6.1 and 6.2 applies directly to, for example, estimating linear panel models using feasible general least squares (for both random and fixed effects models) and marginal/population-average panel data models using generalised estimating equations, when the focus is on inference about the regression coefficients. (Note that random effects panel models with repeated measures nested within individuals are not problematic for pseudo-likelihood estimation of the random effect variance because, by definition, the random effect is defined at the individual level rather than a higher level.)

6.4 | Alternatives: joint modelling and multiple imputation

The other broad class of approaches for handling complex sampling designs involves incorporating the features of the design variables into the analytical model. Such approaches can bring efficiency gains (especially when used in conjunction with likelihood-based estimation techniques) and can address some of the blind spots of pseudo-likelihood modelling (e.g. the problems with multi-level random effects models including levels correlated with the stages of the complex sampling design). However, this is at the expense of a considerably more complex modelling task because of the further assumptions needed to incorporate the interaction between the variables in the analysis and those used in the design. The relative advantages and disadvantages of these approaches, referred to as disaggregated approaches in contrast to the aggregated ones we have discussed in Sections 6.1–6.3, are discussed in detail by Chambers and Skinner (2003, Introduction to Part A).

Particularly for panel data, complete-cases samples can sometimes tempt analysts to use a technique such as multiple imputation to impute missing values and thus increase the sample size.[49] Imputation requires the analyst to assume that the non-response mechanism is ignorable (sometimes referred to the data being ‘missing at random’). If, as is highly likely, the predictor variables are incompletely observed as well as the outcome, this involves analysts extending their modelling assumptions further to capture the distribution of, and relationships among, the predictors. Imputation also requires the analyst to specify an imputation model for the relationship between the missing and the observed variables for each incomplete case: if the imputation model is mis-specified, then bias will result even if the non-response is truly ignorable.

[48] See, for example, Robins, Rotnitzky and Zhao (1995) and Wooldridge (2007).

[49] Carpenter and Kenward, 2013.

Imputation can be attractive particularly if the number of complete cases is disproportionately small (that is, the total number of complete cases is a small fraction of the total number of complete and incomplete cases) relative to the proportion of observed variables among the incomplete cases. Otherwise, the gains from using imputation will be small: the apparently larger sample size is illusory, because any gains will be attenuated once uncertainty about the estimated imputation model is taken into account.

Quartango, Carpenter and Goldstein (2020) review issues surrounding how to combine multiple imputation with pseudo-likelihood estimation. They propose an approach for linear models that involves extending the regression model to include the survey weights and interactions with the predictors. More principled approaches that do not rely on pseudo-likelihood estimation require joint modelling of the analysis variables, sample selection and non-response, but these involve yet more assumptions and potential for model mis-specification, as discussed by Chambers and Skinner (2003). Future research into doubly robust estimators, which combine the relative robustness of pseudo-likelihood with the efficiency gains of multiple imputation (e.g. Seaman et al. 2012), thus appears to be a valuable way forward.

6.5 | Last words

We have discussed key design features and weighting issues that users of Understanding Society data should consider. For most analyses, we strongly recommend the use of the survey weights, available in the released data files, as a means of ensuring the absence of bias due to sample design and minimising the risk of bias due to differential non-response. If only design weights, rather than survey weights, are used, then much stronger assumptions regarding ignorability of the non-response process are needed and/or additional adjustment methods that account for the complexity of non-response in a hierarchical, longitudinal, multi-instrument study. Many issues with estimation are of course analysis/application specific, and effective statistical solutions are not yet available for some more complex forms of longitudinal or hierarchical analysis, such as how to obtain unbiased random effect variances in multi-level models with more than two levels when using complex survey data.

ACKNOWLEDGEMENTS

Understanding Society is an initiative funded by the Economic and Social Research Council (G2014-70) and various government departments, with scientific leadership by the Institute for Social and Economic Research, University of Essex, and survey delivery by NatCen Social Research and Kantar Public. The research data are distributed by the UK Data Service. We would like to thank the editors for their helpful comments, which greatly improved the presentation of the paper.

ORCID

Peter Lynn https://orcid.org/0000-0002-3493-5686
Pablo Cabrera-Álvarez https://orcid.org/0000-0001-8105-5908
Paul Clarke https://orcid.org/0000-0001-8238-3647

REFERENCES

Al Baghal, T., & Lynn, P. (2015), Using motivational statements in web instrument design to reduce item missing rates in a mixed mode context. *Public Opinion Quarterly*, 79, 568–79, https://doi.org/10.1093/poq/nfv023.

Benzeval, M., Andrayas, A., Mazza, J., Al Baghal, T., Burton, J., Crossley, T.F., & Kumari, M. (2023), Does the feedback of blood results in observational studies influence response and consent? A randomised study of the Understanding Society Innovation Panel. *BMC Medical Research Methodology*, 23, 134, https://doi.org/10.1186/s12874-023-01948-y.

Berthoud, R., Fumagalli, L., Lynn, P., & Platt, L. (2009), Design of the Understanding Society ethnic minority boost sample. Understanding Society Working Paper 2009-02, https://www.iser.essex.ac.uk/research/publications/working-papers/understanding-society/2009-02.

Bethlehem, J., Cobben, F., & Schouten, B. (2011), *Handbook of Nonresponse in Household Surveys*, Chichester: Wiley, https://doi.org/10.1002/9780470891056.

Bianchi, A., Biffignandi, S., & Lynn, P. (2017), Web-face-to-face mixed mode design in a longitudinal survey: effects on participation rates, sample composition and costs. *Journal of Official Statistics*, 33, 385–408, https://doi.org/10.1515/jos-2017-0019.

Cabrera-Álvarez P., James, N., & Lynn, P. (2023), Panel attrition in the General Population Sample and the Immigrant and Ethnic Minority Boost of Understanding Society. Understanding Society Working Paper 2023-03, https://www.iser.essex.ac.uk/research/publications/working-papers/understanding-society/2023-03.

Cabrera-Álvarez, P., & Lynn, P. (2023a), Text messages to incentivise response in a web-first sequential mixed-mode survey. *Social Science Computer Review,* forthcoming, https://doi.org/10.1177/08944393231213315.

Cabrera-Álvarez, P., & Lynn, P. (2023b), Short-term impact of increasing the value of unconditional and conditional incentives in Understanding Society. Understanding Society Working Paper 2023-08, https://www.iser.essex.ac.uk/research/publications/working-papers/understanding-society/2023-08.

Carpenter, H. (2016), UK Household Longitudinal Study: wave 6 technical report, https://www.understandingsociety.ac.uk/sites/default/files/downloads/documentation/mainstage/technical-reports/6614_wave6_technical_report.pdf.

Carpenter, J. R., & Kenward, M. (2013), *Multiple Imputation and its Application*. Chichester: Wiley.

Cernat, A., & Lynn, P. (2018), The role of email communications in determining response rates and mode of participation in a mixed-mode design. *Field Methods*, 30, 70–87, https://doi.org/10.1177/1525822X17726205.

Chambers, R. L., & Skinner, C. J. (eds) (2003), *Analysis of Survey Data*. Chichester: Wiley, https://doi.org/10.1002/0470867205.

Fumagalli, L., Laurie, H., & Lynn, P. (2013), Experiments with methods to reduce attrition in longitudinal surveys. *Journal of the Royal Statistical Society Series A (Statistics in Society)*, 176, 499–519, https://doi.org/10.1111/j.1467-985X.2012.01051.x.

Groves, R. M., Dillman, D. A., Eltinge, J., & Little, R. J. A. (2002), *Survey Nonresponse*. Hoboken, NJ: Wiley.

Hall, J., Brown, V., Nicolaas, G., & Lynn, P. (2013), Extended field efforts to reduce the risk of non-response bias: Have the effects changed over time? Can weighting achieve the same effects? *Bulletin of Sociological Methodology*, 117, 5–25, https://doi.org/10.1177/0759106312465545.

Jäckle, A., Lynn, P., & Burton, J. (2015), Going online with a face-to-face household panel: effects of a mixed mode design on item and unit nonresponse. *Survey Research Methods*, 9, 57–70, https://doi.org/10.18148/srm/2015.v9i1.5475.

Jäckle, A., Wenz, A., Burton, J., & Couper, M. P. (2022), Increasing participation in a mobile app study: the effects of a sequential mixed-mode design and in-interview invitation. *Journal of Survey Statistics and Methodology*, 10, 898–922, https://doi.org/10.1093/jssam/smac006.

Kaminska, O. (2021), Weighting for mortality in a longitudinal study. *European Survey Research Association Conference*, online, https://www.europeansurveyresearch.org/conf2021/prog.php?sess=45#198.

Kaminska, O., & Lynn, P. (2017), The implications of alternative allocation criteria in adaptive design for panel surveys. *Journal of Official Statistics*, 33, 781–99, https://doi.org/10.1515/jos-2017-0036.

Laurie, H., Smith, R., & Scott, L. (1999), Strategies for reducing nonresponse in a longitudinal panel study. *Journal of Official Statistics*, 15, 269–82.

Lynn, P. (2008), Chapter 3. The problem of non-response. In E. D. de Leeuw, J. J. Hox and D. A. Dillman (eds), *International Handbook of Survey Research Methods*, Mahwah, NJ: Lawrence Erlbaum Associates, 35–55.

Lynn, P. (2009), Sample design for Understanding Society. Understanding Society Working Paper 2009-01, https://www.understandingsociety.ac.uk/research/publications/514007.

Lynn, P. (2011), Maintaining cross-sectional representativeness in a longitudinal general population survey. Understanding Society Working Paper 2011-04, https://www.understandingsociety.ac.uk/research/publications/520028.

Lynn, P. (2013), Alternative sequential mixed mode designs: effects on attrition rates, attrition bias and costs. *Journal of Survey Statistics and Methodology*, 1, 183–205, https://doi.org/10.1093/jssam/smt015.

Lynn, P. (2014), Longer interviews may not affect subsequent survey participation propensity. *Public Opinion Quarterly*, 78, 500–09, https://doi.org/10.1093/poq/nfu015.

Lynn, P. (2016), Targeted appeals for participation in letters to panel survey members. *Public Opinion Quarterly*, 80, 771–82, https://doi.org/10.1093/poq/nfw024.

Lynn, P. (2017), From standardised to targeted survey procedures for tackling non-response and attrition. *Survey Research Methods*, 11, 93–103, http://doi.org/10.18148/srm/2017.v11i1.6734.

Lynn, P. (2018), Chapter 19. Tackling panel attrition. In D. L. Vannette and J. A. Krosnick (eds), *The Palgrave Handbook of Survey Research*, Palgrave Macmillan, 143–53, https://doi.org/10.1007/978-3-319-54395-6_19.

Lynn, P. (2019), Applying prospect theory to participation in a CAPI/web panel survey. *Public Opinion Quarterly*, 83, 559–67, https://doi.org/10.1093/poq/nfz030.

Lynn, P. (2020), Evaluating push-to-web methodology for mixed-mode surveys using address-based samples. *Survey Research Methods*, 14, 19–30, https://doi.org/10.18148/srm/2020.v14i1.7591.

Lynn, P., Bianchi, A., & Gaia, A. (2023), The impact of day of mailing on web survey response rate and response speed. *Social Science Computer Review,* forthcoming, https://doi.org/10.1177/08944393231173887.

Lynn, P., & Borkowska, M. (2018), Some indicators of sample representativeness and attrition bias for BHPS and Understanding Society. Understanding Society Working Paper 2018-01, https://www.understandingsociety.ac.uk/research/publications/524851.

Lynn, P., Burton, J., Kaminska, O., Knies, G., & Nandi, A. (2012), An initial look at non-response and attrition in Understanding Society. Understanding Society Working Paper 2012-02, https://www.understandingsociety.ac.uk/research/publications/520399.

Lynn, P., & Kaminska, O. (2010), Weighting strategy for Understanding Society. Understanding Society Working Paper 2010-05, https://www.understandingsociety.ac.uk/research/publications/514008.

Lynn, P., Kaminska, O., & Goldstein, H. (2014), Panel attrition: how important is interviewer continuity? *Journal of Official Statistics*, 30, 443–57, https://doi.org/10.2478/jos-2014-0028.

Lynn, P., Nandi, A., Parutis, V., & Platt, L. (2018), Design and implementation of a high-quality probability sample of immigrants and ethnic minorities: lessons learnt. *Demographic Research*, 38, 513–48, https://doi.org/10.4054/DemRes.2018.38.21.

Lynn, P., & Watson, N. (2021), Chapter 18. Issues in weighting for longitudinal surveys. In P. Lynn (ed.), *Advances in Longitudinal Survey Methodology*, New York: Wiley, 447–68.

Olson, K., & Witt, L. (2011), Are we keeping the people who used to stay? Changes in correlates of panel survey attrition over time. *Social Science Research*, 40, 1037–50, https://doi.org/10.1016/j.ssresearch.2011.03.001.

Pfeffermann, D., Skinner, C. J., Holmes, D. J., Goldstein, H., & Rasbash, J. (1998), Weighting for unequal selection probabilities in multilevel models. *Journal of the Royal Statistical Society* B60, 23–40, https://doi.org/10.1111/1467-9868.00106.

Quartango, M., Carpenter, J. R., & Goldstein, H. (2020), Multiple imputation with survey weights: a multilevel approach. *Journal of Survey Statistics & Methodology*, 8, 965–89, https://doi.org/10.1093/jssam/smz036.

Robins, J. M., Rotnitzky, A., & Zhao, L-P. (1995), Analysis of semiparametric regression models for repeated outcomes in the presence of missing data. *Journal of the American Statistical Association*, 90, 106–21.

Seaman, S. R., White, I. R., Copas, A. J., & Li, L. (2012), Combining multiple imputation and inverse probability weighting. *Biometrics*, 68, 129–37.

Skinner, C. J., Holt, D., & Smith, T. M. F. (eds) (1989), *Analysis of Complex Surveys*. Chichester: Wiley.

Watson, D. (2003), Sample attrition between waves 1 and 5 in the European Community Household Panel. *European Sociological Review*, 19, 361–78, https://doi.org/10.1093/esr/19.4.361.

Wenz, A., Jäckle, A., Burton, J., & Couper, M. P. (2022), The effects of personalized feedback on participation and reporting in mobile app data collection. *Social Science Computer Review*, 40, 165–78, https://doi.org/10.1177/0894439320914261.

Williams, J. et al. (2012), Understanding Society wave 14 boost trial: experiments with methods of recruiting a probability online boost sample. Understanding Society Working Paper 2012-07, https://www.understandingsociety.ac.uk/research/publications/547448.

Wooldridge, J. M. (2007), Inverse probability weighted estimation for general missing data problems. *Journal of Econometrics*, 141, 1281–301, https://doi.org/10.1016/j.jeconom.2007.02.002.

How to cite this article: Lynn, P., Cabrera-Álvarez, P., & Clarke, P. (2023), Sample composition and representativeness on Understanding Society. *Fiscal Studies*, 44, 341–59. https://doi.org/10.1111/1475-5890.12357