SURVEY FUTURES
SURVEY DATA COLLECTION METHODS COLLABORATION

Working Paper 20:

Do the effects of interviewer administered follow-ups of web non-respondents on dataset quality differ between overall datasets and population subgroup datasets?

Jamie C. Moore[1*], Gabriele Durrant[2,3], Pablo Cabrera Álvarez[1], Annette Jäckle[1], Peter W.F. Smith[2] and Jonathan Burton[1]

[1]Institute for Social and Economic Research, University of Essex, UK, [2]Department of Social Statistics and Demography, University of Southampton, UK, [3]National Centre for Research Methods, University of Southampton, UK.

June 2026

www.surveyfutures.net

Survey Futures is an Economic and Social Research Council (ESRC)-funded initiative (grant grant ES/X014150/1) aimed at bringing about a step change in survey research to ensure that high quality social survey research can continue in the UK. The initiative brings together social survey researchers, methodologists, commissioners and other stakeholders from across academia, government, private and not-for-profit sectors. Activities include an extensive programme of research, a training and capacity-building (TCB) stream, and dissemination and promotion of good practice. The research programme aims to assess the quality implications of the most important design choices relevant to future UK surveys, with a focus on inclusivity and representativeness, while the TCB stream aims to provide understanding of capacity and skills needs in the survey sector (both interviewers and research professionals), to identify promising ways to improve both, and to take steps towards making those improvements. Survey Futures is directed by Professor Peter Lynn, University of Essex, and is a collaboration of twelve organisations, benefitting from additional support from the Office for National Statistics and the ESRC National Centre for Research Methods. Further information can be found at www.surveyfutures.net.

Prior to citing this paper, please check whether a final version has been published in a journal. If so, please cite that version. In the meanwhile, the suggested form of citation for this working paper is:
Moore, J.C., Durrant, G, Cabrera-Alvarez, P., Jäckle, A., Smith, P.W.F. & Burton, J. (*2026*) 'Do the effects of interviewer administered follow-ups of web non-respondents on dataset quality differ between overall datasets and population subgroup datasets?', *Survey Futures Working Paper* no. 20. Colchester, UK: University of Essex. Available at https://surveyfutures.net/working-papers/.

Contents

Abstract……4

1. Introduction……5

1.1. Motivation: Survey dataset quality and the impacts of survey design changes ……………..6

1.2. Previous research relating to aims and research questions……………………………………….8

1.2.1. RQ1: Overall and subgroup respondent dataset sizes…………………………………………….8

1.2.2. RQ2: Overall and subgroup dataset representativeness………………………………………….9

1.2.3. RQ3: Under-represented (sub-) subgroups in overall and subgroup datasets…………….10

1.2.4. RQ4: Non-response biases after weighting in overall and subgroup datasets…………….10

2. Data………10

2.1. The *Understanding Society* survey………………………………………………………………………10

2.2. Overall analysis sample and subgroup sample definitions………………………………………12

3. Evaluation methods……………………………………………………………………………………………12

3.1. RQ1: Overall and subgroup respondent dataset sizes…………………………………………….12

3.2. RQ2: Overall and subgroup dataset representativeness………………………………………….13

3.2.1. Construction of sample inclusion weights……………………………………………………………13

3.2.2. Auxiliary covariates used in weight estimation………………………………………………………14

3.2.3. Methods to evaluate the representativeness of respondent samples……………………….15

3.3. RQ3: Under-represented (sub-) subgroups in overall and subgroup datasets………………..17

3.4. RQ4: Non-response biases after weighting in overall and subgroup datasets………………..17

4. Results………19

4.1. RQ1: Overall and subgroup respondent dataset sizes…………………………………………….19

4.2. RQ2: Overall and subgroup dataset representativeness………………………………………….20

4.3. RQ3: Under-represented (sub-) subgroups in overall and subgroup datasets………………..21

4.4. RQ4: Non-response biases after weighting in overall and subgroup datasets………………..24

5. Conclusions and implications for survey practice…………………………………………………….25

6. References………29

Appendix A1: Estimation and evaluation of analysis dataset sample inclusion weights………..45

Appendix A2. Missing values and imputation strategy……………………………………………………56

Appendix A3: Partial unconditional CV (CV_u) derivation…………………………………………………67

Appendix A4: Supplementary tables…………………………………………………………………………….69

Abstract

Many social surveys have adopted web first sequential mixed mode designs in which web mode is offered, then non-respondents followed up in interviewer administered face-to-face (Computer Assisted Personal Interviewing (CAPI) or telephone (Computer Assisted Telephone Interviewing (CATI) modes. These designs cost less than CAPI or CATI only designs, and can produce higher quality datasets than web only designs, though research suggests that benefits have declined over time. However, the impacts of such follow-ups on dataset quality for important population subgroups are unknown. We address this question by using Understanding Society: the UK Household Longitudinal Study data to investigate the impact of CATI follow ups of web non-respondents on overall and Young adult, Older adult, Ethnic minority and Low income subgroup dataset quality. We find that follow-ups: 1) increase dataset sizes by more than for the overall dataset for some subgroups but not others; 2) improve both overall and subgroup dataset representativeness; 3) improve the representation of some under-represented subgroups but not others in both overall and subgroup datasets; and 4) reduce non-response biases remaining after non-response weighting for some subgroups but not for others or for the overall dataset. We then discuss the implications of our findings for survey practice.

1. Introduction

Two key components of dataset quality in general population surveys are that they support unbiased estimation of population parameters of interest (Davis-Kean et al. 2017), and similar estimation of parameters for population subgroups, which are often the focus of substantive research (Benzeval et al. 2020). The latter component can be understudied in the context of survey design: evidence upon which decisions are based is often at the level of the entire dataset and ignores implications for subgroups. One example of this concerns sequential mixed mode designs in which first web mode is offered, then non-respondents followed up in interviewer administered face-to-face (Computer Assisted Personal Interviewing, CAPI) or telephone (Computer Assisted Telephone Interviewing, CATI) modes. These designs are now used in many surveys. They cost less than CAPI or CATI only designs (Dilman 2009, p. 401). They can also increase dataset size and resemblance to study populations (representativeness) and reduce non-response bias compared to web only designs (De Leeuw 2018), though notably work suggests that benefits are declining over time (Cabrera-Alvarez et al. 2025; Moore et al. 2025). However, little is known about the impacts of follow-ups in population subgroups.

In this paper, we address this knowledge gap by using data from Understanding Society: the UK Household Longitudinal Study (UKHLS) to quantify the impacts of CATI follow-ups of web non-respondents on data quality for the substantively important Young adult (age 16-34), Older adult (age 75+), Ethnic minority, and Low income (lowest household income quintile) population subgroups. We compare findings to those for the overall population dataset. We treat each subgroup as a population, with members with differing characteristics, instead of as monotypic as in other work. We define dataset quality in terms of response rates and dataset sizes, representativeness, and non-response biases remaining after non-response weighting. We do not quantify measurement differences, where respondent answers depend on mode used, potentially affecting survey estimates (De Leeuw 2018; Burton & Jäckle 2020): these are left to future work. We consider the following four research questions: how do follow-ups of web non-respondents with CATI impact on each subgroup in terms of

RQ1: response rates and dataset sizes, and do findings differ from the overall dataset?

RQ2: dataset representativeness, and do findings differ from the overall dataset?

RQ3: under-represented (sub-) population subgroups, and do findings differ from the overall dataset?

RQ4: non-response biases remaining after non-response weighting, and do findings differ from the overall dataset?

1.1. Motivation: Survey dataset quality and the impacts of survey design changes

General population surveys should support unbiased estimation of study population parameters of interest, and also, as they are often the focus of substantive research, similar estimation of these parameters for population subgroups (Davis-Kean et al. 2017; Benzeval et al. 2020). Estimable parameters should include summary values, gradients, associations and causal relationships (Lynn 2015). Such components of dataset quality, however, are affected by declining survey response rates (Luiten et al. 2020). These reduce dataset size, increasing sampling error, and, if non-respondents and respondents differ, can also cause non-response biases, deviations in estimates from population values that can lead to invalid inference. In response, survey designers seek to maximise dataset quality during data collection by using bias prevention measures such as offering incentives or following up non-respondents to increase response in under-represented population subgroups (e.g. Groves et al. 2001; Groves & Heeringa 2006). They also seek to adjust for remaining biases post collection by producing non-response weights or imputing responses for non-respondents (e.g. Valliant & Dever 2013; Little & Rubin 2014). In addition, the success of bias prevention measures can increase the effectiveness of bias adjustments (Schouten et al. 2016; Moore et al. 2024).

Surveys also face cost pressures. One approach to minimising the increases caused by efforts to maximise dataset quality is to replace CAPI or CATI with less costly modes such as web (Couper et al. 2007; Schonlau et al. 2009; Baker et al. 2010). Beyond reducing costs, web may increase response by some population subgroups compared to other modes (McGonagle & Sastry 2023). However, overall response rates are often lower (Daikeler et al. 2020; Wu et al. 2022), and dataset quality tends to be maximised by mixed mode designs combining web and other modes (Cornese & Bosnjak 2018). Hence, many surveys now utilise sequential mixed mode designs that follow up web non-respondents by CAPI or CATI (see Moore et al. 2025 and references therein). These designs are less costly than CAPI or CATI only designs (see Lipps & Pekari 2021; McGonagle et al. 2023 for recent evidence). They can also lead to higher quality datasets than web only designs (Dillman et al. 2009; Klausch et el. 2015; Lipps & Pekari 2021; Mackeben & Sakshaug 2023; McGonagle & Sastry 2023; Moore et al. 2024; Moore & Durrant 2025). However, recent work suggests that benefits are declining over time as levels of internet access in populations increase (Cabrera-Alvarez et al. 2025; Moore et al. 2025).

A limitation of the previous research on impacts of CAPI / CATI follow-ups of web non-respondents on dataset quality, however, is that effects have mainly been studied for overall datasets that sample entire populations. Little is known about impacts in substantively important population subgroups such as young or older adults, ethnic minorities or low income individuals (see Longhi 2020; de Gessa et al. 2022; Hu & Coulter 2025; Olsen et al. 2024; Wells et al. 2025; Zhang et al. 2025 for research on these subgroups using UKHLS datasets). Moreover, when subgroups are considered, they are treated as monotypic (i.e. with the same characteristics) rather than as populations themselves, with individuals with differing characteristics who may respond differently to the survey design. This makes it difficult to evaluate whether a dataset enables unbiased estimation in subgroup datasets, especially regarding multivariate relationships. Hence, for a full understanding of the impacts of CAPI / CATI follow-ups of web non-respondents on survey dataset quality, and to properly inform survey design decisions, such subgroups must be studied in a similar manner to overall datasets.

1.2. Previous research relating to aims and research questions

1.2.1. RQ1: Overall and subgroup respondent dataset sizes

Limited previous work exists relating to our research questions, and some of it is not useful for our purposes as it considers only combined web plus CAPI / CATI respondent datasets (e.g. Kappelhof 2015). CAPI / CATI follow-ups increase overall dataset size by 3 to 85% (Dillman et al. 2009; Klausch et el. 2015; Lipps & Pekari 2021; McGonagle & Sastry 2023; Mackeben & Sakshaug 2023; Moore et al. 2024, 2025; Moore & Durrant 2025). However, evidence from datasets spanning the period 2012-2023 from the UKHLS Innovation Panel (IP), a survey that for a random subset of participants utilises a web first with CAPI / CATI follow-up design, shows that increases are declining over time as the percentage of the population with internet access increases (Moore et al. 2025: see also Cabrera-Alvarez et al. 2025). Some information also exists on the impacts of follow ups on dataset size for some of the population subgroups examined in this paper, from the UKHLS COVID-19 Study, a survey in which at two waves ca.1/3 of web non-respondents were followed up by CATI (Moore & Durrant 2025; Moore et al. 2024). As calculated from reported overall dataset sizes and subgroup respondent proportions, follow-ups increase dataset size for 20-29 year olds (the youngest age group studied) by 1.3 and 1.9%, for those aged 70+ (the oldest) by 11.9 and 15.1%, and for ethnic minorities by 4.3 and 5.0%.

1.2.2. RQ2: Overall and subgroup dataset representativeness

CAPI / CATI follow-ups tend to increase overall dataset representativeness (Lipps & Pekari 2021; Mackeben & Sakshaug 2023; Moore et al. 2024, 2025; Moore & Durrant 2025: but see Klausch et el. 2015), though in the previously mentioned UKHLS IP datasets increases have declined over time to the point of now being negligible (Moore et al. 2025). Regarding subgroups of interest, the only information available is from breakdowns of analyses of overall datasets in which groups are treated as monotypic, i.e. concerns representation in terms of the closeness of overall proportions of cases to study population proportions. By this definition, in the UKHLS COVID-19 Study datasets, CATI follow-ups improve representation of ethnic minority and aged 70+ subgroups but not that of the 20-29 year old subgroup (Moore & Durrant 2025: Moore et al. 2024). In the UKHLS IP datasets, CAPI / CATI follow-ups improve aged 75+ subgroup representation but at a declining rate over time, have a similar impact on the lowest income quintile subgroup representation but only early in the study period, and have no impact on 16-34 year old subgroup representation (Moore et al. 2025). In addition, in a UKHLS main survey dataset from 2020 in which a web-first with CATI follow-up design was used, follow-ups improve representation of the aged 75+ subgroup, but not that of the 16-34 year old or lowest income quintile subgroups (Moore et al. 2025).

1.2.3. RQ3: Under-represented (sub-) subgroups in overall and subgroup datasets

As discussed in section 1.2.2, CAPI / CATI follow-ups have mixed impacts on the representation of subgroups in overall datasets. No information exists on impacts on the representativeness of subgroups within each of these population subgroups.

1.2.4. RQ4: Non-response biases after weighting in overall and subgroup datasets

In the overall UKHLS COVID-19 Study and UKHLS main survey 2020 datasets, CATI follow-ups have little impact on non-response biases remaining after non-response weighting (Moore et al. 2024, 2025; Moore & Durrant 2025). In the overall UKHLS IP datasets, CAPI / CATI follow-ups reduce biases early in the study period, but impacts are negligible by its end (Moore et al. 2025). No information exists on similar biases within the population subgroups of interest.

2 Data

2.1. The *Understanding Society* survey

The UKHLS survey is a household panel study that began in 2009 and follows people living in a sample of UK households (University of Essex, Institute for Social and Economic Research 2024a, b). All household members aged 16+ are eligible for annual interviews. The first individual in a household to complete their survey is asked to complete a household grid identifying and giving basic sociodemographic information about co-residents. The person responsible for household bills, or their spouse, is also asked to complete a household questionnaire giving information about the household and housing. The sample, which includes respondents from the preceding British Household Panel Survey that began in 1991, is a probability sample selected using a clustered and stratified design (Lynn 2009). All individuals in sample households are followed even if they move within the UK, only becoming ineligible if they emigrate or die. Non-responding households at the first wave of issue and households with no responses for two waves are not re-issued. High dataset representativeness and low non-response imply that the survey supports valid population inference (Benzeval et al. 2020).

The survey began with a CAPI-only design. From wave 7 (2017), an increasing proportion of households have been issued to a web-first with CAPI follow up design. During the COVID-19 pandemic, CAPI interviewing was suspended and a web-first with CATI follow up design introduced. In this paper, we analyze data from the first eight months of this period, April to December 2020, corresponding to survey wave 11, quarters 6-8 (data collection for a wave lasts two years, and each year is split into quarters), and wave 12, quarters 2-4. During the period studied, all sample members were invited to complete online. This included households that had already been invited to complete the survey online in previous years, and households that had always been issued to CAPI-first. A reminder letter and email (where an email address was available) were sent a week later, a second email after two weeks, and, to non-respondents, another letter and email after three weeks and a further email after a month. After five weeks, though the online interview remained available, any non-respondents were issued to CATI interviewers.

2.2. Overall analysis sample and subgroup sample definitions

The overall analysis sample consists of all adults eligible for annual interviews from April to December 2020. No individuals appear in both the wave 11 and 12 samples. Those who had withdrawn from the survey or were found to have died or moved out of scope during the household interview are removed from the sample. We use household grid information from the survey to identify the age and ethnicity of individuals: we could not use information from the individual interviews because it only exists for survey respondents. Even given this, some individuals lack relevant information due to a lack of household grid information for the wave or missing responses. In these cases, where it exists we use information from previous or future waves. We remove individuals for whom information is still not available (905 in total) from the sample. The resulting overall sample includes 30,848 individuals. We define the Young adult population subgroup as those aged 16-34 at time of interview (8,664 individuals), the Older adult subgroup as those aged 75+ (3,092 individuals), the Ethnic minority subgroup as those not of White British ethnicity (7,376 individuals), and the Low income subgroup as those in the 1st quintile of household income (6,593 individuals).

3. Evaluation methods

3.1. RQ1: Overall and subgroup respondent dataset sizes

Our evaluation methods broadly follow those of Moore et al. (2025). We address RQ1 by quantifying the numbers and percentages of overall and subgroup sample members responding by web, CATI, and web plus CATI. The key comparisons when addressing our research questions are between web and web plus CATI (hereafter referred to as ‘combined’) respondent datasets for overall and subgroup samples.

3.2. RQ2: Overall and subgroup dataset representativeness

We address RQ2 by considering the representativeness of web and combined respondents to the adult questionnaire (i.e. unit respondents) compared to the relevant (overall or subgroup) study population. In these analyses we use the issued sample of households, weighted to adjust for non-response and attrition over time, as the study populations. In the following, we report how the weights are calculated, the auxiliary variables used in their creation, and the methods used to assess representativeness.

3.2.1. Construction of sample inclusion weights

We create sample inclusion weights to adjust the issued sample to the composition of the study population. The creation of the weights and evaluations of their performance are fully described in Appendix A1. The following is a summary. We begin with the wave 6 inclusion enumeration weights released with the publicly available data. These weight all members of households with a completed household grid at wave 6, and serve as sample inclusion weights at wave 7.

The mentioned weights are then adjusted for sample attrition (those in households not responding for two waves are excluded from issued samples for following waves) up until 2020. For this adjustment, we multiply them with the inverses of the probabilities of inclusion in the 2020 issued sample as estimated using logistic regression with a set of auxiliary covariates likely to be correlated with both inclusion and survey answers (see section 3.2.2 for covariates). There are, however, some individuals for whom weights cannot be calculated in this way, because they were in households that did not respond at wave 6 or joined households after wave 6 and therefore have no inclusion enumeration weight. Hence, if other household members have an adjusted weight, at the next stage such weights are shared with them to provide a valid weight. Then, to weight remaining unweighted individuals (those in households where no-one has an adjusted weight), the weights of individuals with similar covariate values in other households are shared. Finally, the weights are post-stratified to the cross-tabulation of sex (male, female), age (0-15, 16-34, 35-54, 55-74, 75+), and region (north, east, south, west) of estimated 2020 population totals. Weights are also produced for under 16-year-olds so that they are available if they turn 16 and become eligible for the adult interviews by 2020.

3.2.2. Auxiliary covariates used in weight estimation

As noted in section 2.1, we cannot use information from the adult questionnaires because it only exists for survey respondents. Hence, we instead use the following household grid and household questionnaire covariates: sex (male, female), age (0-15, 16-34, 35-54, 55-74, 75+), ethnicity (white British, south Asian, Black, Other); activity last week (in work, not in work), housing tenure (owner occupied, mortgage, rented / other), household structure (single adult; single adult, kids; couple, no kids; couple, kids; other), region (north, east, south, west), behind with paying bills (no, yes), behind with paying council tax (no, yes), household location (urban, rural), equivalised household income (quintiles), number of rooms in household (continuous), household size (continuous), and whether the individual is part of one of the two ethnic minority boost samples included in the overall sample, which oversample such minorities but include some White British individuals (no, yes). Covariate missingness rates can be up to ~35%, due to item and household non-response. We use information from previous or future waves to reduce these rates to 0% to ~20%. We then utilise imputation given values of other individuals with similar characteristics as measured by the weighting model auxiliary covariates listed in the previous paragraph to replace remaining missing covariate values. After this, we remove individuals with imputed missing age, or household income values from the analysis sample (we took this approach to maximise use of information in the dataset: see section 2.2 for analysis sample details). These procedures are fully outlined in Appendix A2.

3.2.3. Methods to evaluate the representativeness of respondent samples

We evaluate the representativeness of respondents using Coefficients of Variation of response propensities (CVs) (Schouten et al. 2012). These quantify variation in response propensities across sample members, estimated by logistic regression using auxiliary covariates available for both respondents and non-respondents. Assuming covariates are correlated with survey variables, low propensity variation (high representativeness) implies low non-response bias risk (see Schouten et al. 2016; Nishimura et al. 2017 for empirical support).

To quantify overall dataset representativeness, we calculate the overall sample CV, which for sample size *n* and auxiliary covariate set $\boldsymbol{x}$, producing the propensity vector $p_{\boldsymbol{x}}$ is:

$$\widehat{CV}(p_{\boldsymbol{x}}) = \frac{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(\hat{p}_i - \hat{\bar{p}}\right)^2}}{\hat{\bar{p}}}, \quad (1)$$

where $\hat{p}_i$ is the response propensity of subject *i* and $\hat{\bar{p}}$ average response propensity. Weights can be applied when estimating propensities to map sample members to the population. The less propensities differ, the smaller the overall CV and the greater dataset representativeness.

Specifically, we use sample bias adjusted CVs, which account for bias caused by estimating response propensities (see Shlomo et al. 2012; de Heij et al. 2015), to quantify dataset representativeness. We cannot conduct exact tests of whether CVs are statistically significant or differ from one another because work does not exist on the distribution of such differences and is beyond the scope of this paper. We instead compute CV standard errors (see Shlomo et al. 2012; de Heij et al. 2015) for conversion into 95% confidence intervals (CIs). This approach provides accurate inference for CV significance (indicated by CIs that do not overlap zero), and significant differences between CVs (non-overlapping CIs). However, a small amount of overlap between CV CIs will be equivocal regarding significance, that is, tests of differences will be slightly conservative (see Goldstein & Healy 1995 for discussion).

We use nine of the covariates listed in section 3.2.2 in regression models estimating response propensities: Sex, Age (excluding those aged 0-15 as they are not eligible for the adult interviews), Activity last week, Ethnicity, Housing tenure, Household structure, Region, Behind with paying bills and Household income. Note however, that we do not use all covariates in each analysis. For Young adults, we exclude Age from the covariate list because no individuals in the other categories are in the dataset. For Older adults, we exclude Age for the same reasons, Activity last week as no individuals report being In Work, and collapse Household Structure to Single, with or without kids, Couple, with or without kids and Other categories as only three individuals report living with kids. For Ethnic minorities, we exclude Ethnicity as there are no white British individuals in the dataset. For low income individuals, we exclude household income. Proper comparison of CVs between datasets requires that the same response propensity models be used (Schouten et al. 2012), so for each subgroup dataset we report a comparator analysis of the overall dataset using the same covariate (sub)-set. We compute CVs and 95% CIs for web and combined respondent datasets, using the R code released by de Heij et al. (2015) in R.4.4.3 (other analyses in this paper are conducted in Stata 19.2). This code is included in the files accompanying this paper. It allows survey strata, but not primary sampling unit (PSU), to be accounted for in analyses. The impact of not accounting for PSU on CVs is unknown, but more generally it has been found that parameter estimates are unaffected while standard errors are reduced. Hence, the null hypothesis, in this case that web and combined datasets do not differ, will be rejected too often (e.g. Lynn 2019).

3.3. RQ3: Under-represented (sub-) subgroups in overall and subgroup datasets

To quantify the impacts of follow-ups on under-represented subgroups, we compute covariate category unconditional CVs (CV_us) (Schouten et al. 2012; see Appendix A3 for details). These indicators decompose response propensity variation into that associated with specific auxiliary covariate categories. CV_us may be negative, indicating under-representation of the covariate category, or positive, indicating over-representation. Our analyses follow the same approach as used for overall CVs. We compute sample bias adjusted CV_us and their standard errors for conversion into 95% CIs (see Shlomo & Schouten 2013; Shlomo et al. 2013; de Heij et al. 2015 for details). For web and combined respondent datasets, we compute indicators for all categories of the auxiliary covariates included in the response propensity models, using the R code of de Heij et al. (2015). For each subgroup, we also undertake comparator analyses of the overall dataset using the same auxiliary covariates.

3.4. RQ4: Non-response biases after weighting in overall and subgroup datasets

Most surveys including the UKHLS provide non-response weights to adjust for the impact of unit non-response on estimates. We investigate the impact of follow ups on the quality of non-response weighted estimates and whether they differ between overall and subgroup datasets. First, we create non-response weights separately for the overall web and combined respondent datasets, using the methods used to create the sample inclusion weights. We use the covariates listed in section 3.2.2 (excluding those aged 0-15) to calculate predicted probabilities of response, then multiply their inverse with the earlier created sample inclusion weights.

Then, we quantify non-response biases in non-response weighted estimates of survey measured respondent characteristics. As is often the case (Hand 2018), population values for considered characteristics are mostly unavailable, so we quantify biases compared to equivalent benchmark estimates of characteristics weighted using the earlier created sample inclusion weights. We statistically compare web and combined dataset quality overall by for each characteristic computing the absolute standardised bias as the absolute difference between the non-response weighted estimate and the benchmark sample inclusion weighted estimate, standardised by the benchmark estimate standard deviation. We then use paired Wilcoxon sign tests to compare web and combined dataset biases across all considered characteristics (we use this test due to the small sample sizes). These comparisons, and for subgroups how they compare to similar for the overall dataset, are our primary focus when investigating this research question. In addition, we evaluate the significance of individual untransformed biases, i.e. differences between non-response and benchmark sample inclusion weighted estimates. Respondents are a subset of the issued sample (i.e. there are partial dependencies between datasets), so to undertake this we use the test of Moore et al. (2024). We also incorporate survey strata and PSU into these latter analyses.

We replicate these analyses for both the overall dataset and each of the four subgroups. For the overall dataset, we use as characteristics the categories of the covariates listed in section 3.2.3. For the subgroups, we use categories of the covariate subsets listed in the same section.

4. Results

4.1. RQ1: Overall and subgroup respondent dataset sizes

Table 1 documents the response rates by mode for the data set overall and each of the subsets. In the overall sample, 59% of individuals respond by web and 14% by CATI, for a combined rate of 73%. Follow-ups therefore increase overall respondent dataset size by 14 / 59 = 23.7%.

Among Young adults, 51% respond by web and 10% by CATI, for a combined rate of 61%. Follow-ups therefore increase respondent dataset size by 19.6%. Among Older adults, 45% respond by web and 31% by CATI, for a combined rate of 76%. Follow-ups therefore increase respondent dataset size by 68.9%. Among Ethnic minority individuals, 47% respond by web and 16% by CATI, for a combined rate of 63%. Follow-ups therefore increase respondent dataset size by 34.0%. Among Low-income individuals, 47% respond by web and 21% by CATI, for a combined rate of 68%. Follow-ups therefore increase respondent dataset size by 44.7%. In sum, CATI follow-ups have more impact on respondent dataset size than for the overall dataset for some subgroups but not others, with the impact largest for Older adults.

4.2. RQ2: Overall and subgroup dataset representativeness

Table 2 documents the overall CVs for each subgroup and for each comparator overall dataset analysis. For Young adults, the overall CV for web respondents is statistically significantly larger than zero (= 0.282, 95% confidence interval (CI) 0.263 – 0.301), implying non-representativeness. It is significantly reduced (i.e. CV 95% CIs do not overlap) in the combined respondent dataset (= 0.242, 95% CI 0.226 – 0.257), implying that CATI follow ups improve representativeness. A similar improvement is found in the comparator overall dataset analysis (web respondent overall CV = 0.217, 95% CI 0.208 – 0.226; combined respondent overall CV = 0.163, 95% CI 0.156 – 0.170).

For Older adults, web respondents are significantly non-representative (overall CV = 0.409, 95% CI 0.370 – 0.448). CATI follow ups significantly improve representativeness (combined respondent overall CV = 0.127, 95% CI 0.106 – 0.148). A smaller, but still significant, similar improvement is found in the comparator overall dataset analysis (web respondent overall CV = 0.205, 95% CI 0.196 – 0.214; combined respondent overall CV = 0.150, 95% CI 0.143 – 0.157).

For Ethnic minorities, web respondents are significantly non-representative (overall CV = 0.263, 95% CI 0.241 – 0.285). CATI follow ups significantly improve representativeness (combined respondent overall CV = 0.200, 95% CI 0.183 – 0.216). A similar improvement is found in the comparator overall dataset analysis (web respondent overall CV = 0.249, 95% CI 0.240 – 0.258; combined respondent overall CV = 0.178, 95% CI 0.171 – 0.185).

For Low-income individuals, web respondents are significantly non-representative (overall CV = 0.351, 95% CI 0.327 – 0.375). CATI follow ups significantly improve representativeness (combined respondent overall CV = 0.202, 95% CI 0.186 – 0.218). A smaller, but still significant, similar improvement is found in the comparator overall dataset analysis (web respondent overall CV = 0.237, 95% CI 0.228 – 0.246; combined respondent overall CV = 0.164, 95% CI 0.157 – 0.171).

In sum, follow ups improve dataset representativeness for all four subgroups as well as in each comparator overall dataset analysis. Compared to in the overall dataset analyses, improvements are greatest for the Low income and especially the Older adult subgroups.

4.3. RQ3: Under-represented (sub-) subgroups in overall and subgroup datasets

Each subgroup analysis is again accompanied by a comparator overall dataset analysis using the same response propensity model. A negative unconditional CV (CV_u) with 95% CIs that do not overlap zero indicates significant under-representation of a covariate category. A combined dataset CV_u of smaller magnitude, with 95% CIs that do not overlap those for the web dataset, indicates significantly reduced under-representation due to follow ups.

For Young adults, there are 8 auxiliary covariates with 27 categories in the response propensity model of which 8 categories are significantly under-represented in the web dataset: Males; South Asian; Not in work; Rented / Other accommodation; Single adult, kids household; Other household; Behind with bills; and 1st quintile household Income (Figure 1a; see Appendix A4 Table 1 for tabulated values). Follow-ups do not significantly reduce under-representation for any categories, but do cause significant under-representation of the North region. Impacts of follow-ups are greater in the comparator overall dataset analysis where 11 categories are significantly under-represented in the web dataset: those listed above plus Black ethnicity, Single adult households and North region (Figure 1b; see Appendix A4 Table 1 for tabulated values and 95% CIs). Follow-ups significantly reduce under-representation of 4 categories: South Asian, Not in work, Rented / Other accommodation and 1st quintile household income. In addition, Single adult households become significantly over-represented, i.e. have a positive CV_u with 95% CIs that do not overlap zero.

For Older adults, there are 7 auxiliary covariates with 23 categories in the response propensity model of which 9 categories are significantly under-represented in the web dataset: Females, Black ethnicity, Rented / Other accommodation, Single adult households, Behind with bills, 1st quintile household income, Other households, North region and South Asian (Figure 2a; see Appendix A4 Table 2 for tabulated values). Follow-ups significantly reduce under-representation of the first 6 categories, and, although web and combined dataset 95% CIs still overlap, the next 2 become representative (i.e. have CV_u 95% CIs that overlap zero). In addition, Other ethnicity becomes significantly under-represented. Impacts of follow ups are less in the comparator overall dataset analysis where 9 categories are significantly under-represented in the web dataset: those listed above excluding Female, plus Male (Figure 2b; see Appendix A4 Table 2 for tabulated values). Follow-ups significantly reduce under-representation of 2 categories: Rented / Other accommodation and 1st quintile household Income. In addition, Other Ethnicity becomes significantly under-represented, and Single adult households becomes significantly over-represented.

For Ethnic minorities, there are 8 auxiliary covariates with 27 categories in the response propensity model of which 10 categories are significantly under-represented in the web dataset: Males, Aged 16-34, Aged 75+, Not in work, Rented / Other accommodation, Single adult, kids household, Other household, South region, Behind with bills and 1st quintile household income (Figure 3a; see Appendix A4 Table 3 for tabulated values). Follow-ups significantly reduce under-representation of the category Aged 75+, and Not in work, Single adult, kids household and South region become representative. Impacts of follow-ups are similar in the comparator overall dataset analysis with 11 categories significantly under-represented in the web dataset: those listed above excluding South region, plus Single adult households and North region (Figure 3b; see Appendix A4 Table 3 for tabulated values). Follow-ups significantly reduce under-representation of 4 categories: Aged 75+, Not in work, Rented / Other accommodation and 1st quintile household Income. In addition, Single adult households become significantly over-represented.

For Low-income individuals, there are 8 auxiliary covariates with 26 categories in the response propensity model of which 9 categories are significantly under-represented in the web dataset: Aged 75+, Not in work, Rented / Other accommodation, Males, South Asian, Black ethnicity, Single adult, kids household, Other household, North Region and Behind with bills (Figure 4a; see Appendix A4 Table 4 for tabulated values). Follow-ups significantly reduce under-representation of the first 3 of these categories, and Black ethnicity and North region become representative. In addition, Aged 16-34 and Couple, kids households become significantly under-represented. Impacts of follow-ups are similar in the comparator overall dataset analysis with 12 categories significantly under-represented in the web dataset: those listed above, plus Aged 16-34, Other ethnicity and Single adult household (Figure 4b; see Appendix A4 Table 4 for tabulated values). Follow-ups significantly reduce under-representation of 4 categories: Aged 75+, Black ethnicity, Not in work, and Rented / Other accommodation. In addition, Other ethnicity become significantly under-represented.

In sum, under-representation of population subgroups is reduced by follow-ups for some subgroups but not others. Comparing the results for each subgroup to the corresponding overall dataset analyses shows that the follow-ups have the largest effects on under-representation for the Older adult and Low-income subgroups.

4.4. RQ4: Non-response biases after weighting in overall and subgroup datasets

Table 3 reports the means of absolute standardised biases (MASBs) comparing non-response weighted estimates to sample inclusion weighted benchmarks and their 95% Confidence Intervals for overall and subgroup web and combined datasets, and also for overall and subgroup datasets *p* values from paired Wilcoxon sign tests comparing web and combined dataset biases. For the overall dataset, the web dataset (MASB) is 0.003 (95% CI 0.002 – 0.004). Follow-ups do not significantly reduce biases (combined dataset MASB = 0.002; 95% CI 0.001 – 0.003): the distribution of differences is symmetric about zero (Wilcoxon sign test $p = 0.856$). Unstandardised biases are reported in Appendix A4 Table 5. All estimates are prevalences, so can be multiplied by 100 to give percentages. No web or combined dataset biases are significant or above 1 percentage point.

For Young adults, the web dataset MASB is 0.031 (95% CI 0.022 – 0.041). Follow-ups significantly reduce biases (combined dataset MASB = 0.020; 95% CI 0.015 – 0.025; $p = 0.029$). Four web dataset unstandardized biases and one combined dataset bias are significant, with the largest above 3 percentage points (Appendix A4 Table 5).

For Older adults the web dataset MASB is 0.069 (95% CI 0.045 – 0.093). Follow-ups significantly reduce biases (combined dataset MASB = 0.022; 95% CI 0.016 – 0.028; $p = 0.011$). Eight unstandardised web dataset biases and no combined dataset biases are significant, with the largest above 9 percentage points (Appendix A4 Table 5).

For Ethnic minorities, the web dataset MASB is 0.026 (95% CI 0.019 – 0.033). Follow-ups do not significantly reduce biases (combined dataset MASB = 0.019; 95% CI 0.014 – 0.023; $p = 0.557$). No unstandardized web or combined dataset biases are significant, though the largest is above 3 percentage points (Appendix A4 Table 5).

For Low-income individuals, the web dataset MASB is 0.037 (95% CI 0.023 – 0.051). Follow-ups significantly reduce biases (combined dataset MASB = 0.018; 95% CI 0.011 – 0.025; $p = 0.015$). Seven unstandardized web dataset biases and two combined dataset biases are significant, with the largest above 4 percentage points (Appendix A4 Table 5). Hence, CATI follow-up has no impact on non-response biases remaining after non-response weighting for the overall dataset or for the Ethnic minority subgroup. It does though, reduce biases for the Young adult, Older adult and Low-income subgroups.

5. Conclusions and implications for survey practice

Summary: This paper builds on previous research examining the impact of following up web non-respondents with Computer Aided Telephone interviews (CATI) on data quality. However, in this paper, we focus on the effects of this survey design on four population subgroups that are often analysed by substantive researchers: Young adult, Older adult, Ethnic minority, and Low-income subgroups. We used data from Understanding Society: The UK Household Longitudinal Study (UKHLS) collected in 2020, in which all sample members were issued to web with non-respondents followed up in CATI. We quantified the effects of the CATI follow-ups on response rates and dataset size, dataset representativeness, and non-response biases remaining after non-response weighting for each of the subgroups as well as for the overall dataset.

Key findings: Our results are summarized in Table 4. They show that CATI follow-ups of web non-respondents improved dataset size both for subgroups and the overall dataset. Impacts were mostly, but not always, greater for the subgroups than for the overall dataset, with the greatest impact on Older adults. Our results also show that in each case follow-ups improved respondent dataset representativeness, again with the greatest impact on Older adults. In addition, follow-ups improved representation of some under-represented groups within subgroups but not others, with impacts greatest on Older adults. However, analyses also show that follow-ups reduced non-response biases remaining after non-response weighting for the Young adult, Older adult and Low income subgroups, but not for the Ethnic minority subgroup or the overall dataset.

The findings on overall dataset quality are comparable to those of previous research (Dillman et al. 2009; Klausch et el. 2015; Lipps & Pekari 2021; McGonagle & Sastry 2023; Mackeben & Sakshaug 2023; Moore et al. 2024, 2025; Moore & Durrant 2025; see also Introduction). Our results on subgroup dataset quality, with the exception of some limited work on dataset size and representativeness when subgroup members are treated as monotypic (i.e. with unvarying characteristics: see Introduction), are novel, as far as we are aware.

Implications of findings for survey practice: Our findings imply that follow-ups may have greater impacts on dataset quality for subgroups than for the overall dataset. This finding is particularly important given the recent work of Moore et al. (2025: see also Cabrera-Alvarez et al. 2025), who showed that in the UKHLS Innovation Panel survey the impacts of CAPI / CATI follow-ups on overall datasets have declined over time as levels of internet access increase, and concluded that the point may come soon when they are no longer needed to maximise dataset quality. Our results in this paper suggest that even if follow-ups no longer make a difference to the overall dataset, survey designers should be cautious of discontinuing follow-ups of web non-respondents to reduce survey costs without explicit evidence of a similar lack of impact on dataset quality for subgroups.

Limitations: One limitation of our research is that the non-response weighted estimates of respondent characteristics were not compared to population values, i.e. actual non-response biases were not calculated. This is because, as in many studies (see Hand 2018), population values for most considered characteristics do not exist. Instead, we took the approach of comparing estimates to similar benchmark estimates calculated using the weighted issued samples: Benzeval et al. (2020) provide evidence that similar UKHLS estimates approximate population values. A second limitation is that the characteristics we consider are restricted to mainly sociodemographic and often household level characteristics from the survey household grid and questionnaire, which were the only ones available for all sample members; findings may differ for more substantive variables. A third limitation is that we study a panel survey in which sample members are repeatedly interviewed and in many cases have previously been exposed to web mode; findings may differ in cross-sectional surveys that only interview individuals once.

Future research: Our findings indicate a number of avenues for future research. The first is to repeat our analyses on future UKHLS datasets, to assess whether, given increasing levels of internet access (see Ofcom 2025), a point is reached where web only designs produce both overall and population subgroup datasets of similar quality to those that follow up web non-respondents by CAPI or CATI (see also Cabrera-Alvarez et al. 2025; Moore et al. 2025). This work could potentially also include other population subgroups, both those that are currently substantively important such as individuals living in rented accommodation (e.g. Clair et al. 2024). The other avenues concern investigating whether comparable findings exist with other surveys. Both other UK surveys and surveys in other countries should be considered: different results between countries may be expected due to differences in levels and patterns of internet access (e.g. Eurostat 2024) and also differences in the ethnicities of minorities present (see Kappelhof 2015). In addition, cross-sectional surveys should be studied, as well as surveys where web non-respondents are followed up in other modes such as paper (see Olson et al. 2021). Moreover, it would be worthwhile to consider similar research questions in surveys that employ 'knock-to-nudge' methods in which non-respondents are contacted to encourage completion by web (e.g. Kastberg & Siegler 2022; Kunz et al. 2024). We end by noting that this research would also benefit if it could consider more substantive variables than was possible in this paper.

6. References

Baker, R., Blumberg, S., Brick, M.J., Couper, M.P., Courtright, M., Dennis, M., Dillman, M., Frankel, M.R., Garland, P., Groves, R.M., Kennedy, C., Krosnick, J., Lee, S., Lavrakas, P.J., Link, M., Piekarski, L., Rao, K., Rivers, D., Thomas, R.K. & Zahs, D. (2010) *AAPOR report on online panels.* American Association for Public Opinion Research.

Benzeval, M., Bollinger, C. R., Burton, J., Crossley, T.F. & Lynn, P. (2020) *The representativeness of Understanding Society.* Understanding Society Working Paper Series 2020–08, Institute for Social and Economic Research.

Brown, M. & Calderwood, L. (2020) *Mixing modes in longitudinal surveys: an overview*. Centre for Longitudinal Studies.

Burton, J. & Jäckle, A. (2020) *Mode Effects*, Understanding Society Working Paper 2020- 05. Colchester: University of Essex.

Alvarez, P.C., Moore, J.C., Jäckle, A., Durrant, G., Burton, J. & Smith, P.W. (2025) *Are we stepping into the future? Exploring the representativeness of web-only surveys of the general population.* Understanding Society Working Paper Series No. 2025-16, Colchester: University of Essex

Carpenter, J.R. & Kenward, M.G. (2013) *Multiple Imputation and Its Application.* John Wiley & Sons, Ltd., Chichester. https://doi.org/10.1002/9781119942283

Clair, A., Baker, E. & Kumari, M. (2024) Are housing circumstances associated with faster epigenetic ageing? J. Epidemiol. Community Health 2024, 78: 40–46. https://doi.org.10.1136/jech-2023-220523

Cornesse, C. & Bosnjak, M. (2018) Is there an association between survey characteristics and representativeness? A meta-analysis. *Surv. Res. Meth.*, 12: 1-13. https://doi.org/10.18148/srm/2018.v12i1.7205

Couper, M. P., Kapteyn, A., Schonlau, M. & Winter, J. (2007) Noncoverage and nonresponse in an Internet survey. *Soc. Sci. Res.* 36: 131–148. https://doi.org/10.1016/j.ssresearch.2005.10.002

Daikeler, J., Bosnjak, M., & Lozar Manfreda, K. (2020). Web versus other survey modes: An updated and extended meta-analysis comparing response rates. *J. Surv. Stat. Meth.,* **8**, 513–539. https://doi.org/10.1093/jssam/smz008

Davis-Kean, P., Chambers, R.L., Davidson, L.l., Kleinert, C., Ren, Q. and S. Tang (2017) Longitudinal Studies Strategic Review: 2017 Report to the Economic and Social Research Council Swindon: ESRC.

de Heij, V., Schouten, B. and Shlomo, N. (2015) RISQ Manual 2.1: Tools in SAS and R for the computation of R indicators and partial R indicators. Available from: www.risq-project.eu.

de Leeuw, E. (2018), Mixed-Mode: Past, Present, and Future. *Survey Research Methods*, 12: 75 – 89.

Di Gessa, G., Xue, B., Lacey, R. & McMunn, A. (2022) Young adult carers in the UK—New evidence from the UK Household Longitudinal Study. *International Journal of Environmental Research and Public Health*, **19**, 14076.

Dillman, D. A., Phelps, G., Tortora, R., Swift, K., Kohrell, J., Berck, J. & Messer, B. L. (2009) Response rate and measurement differences in mixed-mode surveys using mail, telephone, interactive voice response (IVR) and the Internet. Soc. Sci. Res 38: 1-18.

Eurostat (2024) *Digital economy and society statistics - households and individuals.* https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Digital_economy_and_society_statistics_-_households_and_individuals#:~:text=In%20the%20EU%2C%20the%20share,or%20goods%20online%20in%202024.

Groves, R. M., Dillman, D. A., Eltinge, J. L., & Little, R. J. (eds.) (2001) *Survey Nonresponse*. Wiley Series in Survey Methodology.

Groves, R. M. & Heeringa, S. (2006) Responsive design for household surveys: tools for actively controlling survey errors and costs. *J. Roy. Stat. Soc. A.*, **169**, 439-457. https://doi.org/10.1111/j.1467-985X.2006.00423.x

Hand, D. J. (2018) Statistical Challenges of Administrative and Transaction Data, *Journal of the Royal Statistical Society Series A: Statistics in Society*, 181, 555–605, https://doi.org/10.1111/rssa.12315

Hu, Y. & Coulter, R. (2025) Living Apart Together and Older Adults' Mental Health in the United Kingdom. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, *80*(3), p.gbae192.

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 Non-Response, *Surv. Res. Meth.* 9: 57-70. https://doi.org/10.18148/srm/2015.v9i1.5475

Kappelhof, J.W.S. (2015) *Surveying ethnic minorities: The impact of survey design on data quality*. SCP.

Kastberg, S. & Siegler, V. (2022) *Impact of COVID-19 on ONS social survey data collection*. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/methodologies/impactofcovid19ononssocialsurveydatacollection

Klausch, T., Hox, J. & Schouten, B. (2015) *Assessing the mode dependency of sample selectivity across the survey response process.* Statistics Netherlands, The Hague.

Kunz, T., Daikeler, J. & Ackermann-Piek, D. (2024) Interviewer-observed paradata in mixed-mode and innovative data collection, *International Journal of Market Research*, **66**, 14-26.

Lipps, O. & Pekari, N. (2021) Sequentially mixing modes in an election survey. *Survey Methods: Insights from the Field.* Retrieved from https://surveyinsights.org/?p=15281

Little, R. J., & Rubin, D. B. (2014) *Statistical analysis with missing data*, Wiley: New York. https://doi.org/10.1002/9781119013563

Longhi, S. (2020) A longitudinal analysis of ethnic unemployment differentials in the UK. *Journal of Ethnic and Migration Studies*, **46**, 879-892.

Luiten, A., Hox, J., & de Leeuw, E. (2020) Survey Nonresponse Trends and Fieldwork Effort in the 21st Century: Results of an International Study across Countries and Surveys. *Journal of Official Statistics*, *36*, 469-487. https://doi.org/10.2478/jos-2020-0025

Lynn, P. (2009) *Sample Design for Understanding Society.* Understanding Society Working Paper Series, 2009–01, Institute for Social and Economic Research.

Lynn, P. (2015) The Need for Representative Samples. *Longitudinal and Life Course Studies*, 6, 447 – 475. http://dx.doi.org/10.14301/llcs.v6i4.345.

Mackeben J. & Sakshaug J.W. (2023) Transitioning an employee panel survey from telephone to online and mixed-mode data collection. *Stat. J. IAOS*. 39: 213-232. https://doi.org/10.3233/SJI-220088

McGonagle K.A. & Sastry N. (2023) Transitioning to a Mixed-Mode Study Design in a National Household Panel Study: Effects on Fieldwork Outcomes, Sample Composition and Costs. *Surv Res Methods*. 17: 411-427. https://doi.org/10.18148/srm/2023.v17i4.8172

Moore, J.C., Durrant, G.B. & Smith, P.W. (2018) Data set representativeness during data collection in three UK social surveys: generalizability and the effects of auxiliary covariate choice. *J. Roy. Stat. Soc. Ser. A,* 181: 229-248. https://doi.org/0964–1998/18/181229

Moore, J.C., Burton, J., Crossley, T. F., Fisher, P., Gardiner, C., Jäckle, A., & Benzeval, M. (2024) *Assessing Bias Prevention and Bias Adjustment in a Sub-Annual Online Panel Survey*. Understanding Society Working Papers Series 2024-04. Colchester: University of Essex

Moore, J.C. & Durrant, G. (2025) *Quantifying the impacts of web-first sequential mixed mode survey design on UKHLS COVID-19 Study dataset quality.* Understanding Society Working Paper 2025-04. Colchester: University of Essex

Moore, J.C., Cabrera Álvarez, P., Durrant, G., Jäckle, A., Smith, P.W.F. & Burton, J. (2025) *Are interviewer administered follow-ups of web non-respondents still needed to maximise respondent dataset quality? Evidence using Understanding Society: the UK Household Longitudinal Study.* Understanding Society Working Paper 2025-02. Colchester: University of Essex

Nishimura, R., Wagner, J. & Elliott, M. (2016). Alternative indicators for the risk of non-response bias: a simulation study. *Int. Stat. Rev.,* 84, 43-62. https://doi.org/10.1111/insr.12100

Olsen, J.R., Whitley, E., Long, E., Rigby, B.P., Macdonald, L., Dibben, G.O., Palmer, V.J., Benzeval, M., Mitchell, K., McCann, M. and Anderson, M., (2024) Individual, social and area level factors associated with older people's walking: Analysis of an UK household panel study (Understanding Society). *Social Science & Medicine*, **358,** 117083.

Olson, K., Smyth, J.D., Horwitz, R., Keeter, S., Lesser, V., Marken, S., Mathiowetz, N.A., McCarthy, J.S., O'Brien, E., Opsomer, J.D. & Steiger, D. (2021) Transitions from telephone surveys to self-administered and mixed-mode surveys: AAPOR task force report. *J. Surv. Stat. Meth.* 9: 381-411. https://doi.org/10.1093/jssam/smz062

Schonlau, M., Soest, A. van, Kapteyn, A. & Couper, M. (2009) Selection Bias in Web Surveys and the Use of Propensity Scores: *Soc. Meth. Res.* 37: 291-318. https://doi.org/10.1177/0049124108327128

Schouten, B., Cobben, F. and Bethlehem, J. (2009) Indicators for the representativeness of survey response. *Survey Methodology,* 35, 101-113.

Schouten, B., Shlomo, N. and Skinner, C. (2011) Indicators for monitoring and improving representativeness of response. *Journal of Official Statistics*, 27, 231-253.

Schouten, B., Bethlehem, J., Beullens, K., Kleven, Ø., Loosveldt, G., Luiten, A., Rutar, K., Shlomo, N. and Skinner, C. (2012) Evaluating, comparing, monitoring, and improving representativeness of survey response through R-indicators and partial R-indicators. *Int. Statist. Rev.*, 80, 382-399. https://doi.org/10.1111/j.1751-5823.2012.00189.x

Schouten, B., Cobben, F., Lundquist, P. & Wagner, J. (2016) Does more balanced survey response imply less non-response bias? *J. Roy. Stat. Soc. A,* **179**, 727-748. https://doi.org/10.1111/rssa.12152

Shlomo, N., Skinner, C., & Schouten, B. (2012) Estimation of an indicator of the representativeness of survey response. *J. Stat. Plan. Inf.*, 142, 201-211.

Shlomo, N. and Schouten, B. (2013) *Theoretical Properties of Partial Indicators for Representative Response.* Technical Report, University of Southampton.

University of Essex, Institute for Social and Economic Research. (2024a). *Understanding Society: Waves 1-14, 2009-2023 and Harmonised BHPS: Waves 1-18, 1991-2009*. [data collection]. *19th Edition.* UK Data Service. SN: 6614, DOI: http://doi.org/10.5255/UKDA-SN-6614-20

University of Essex, Institute for Social and Economic Research (2024b) *Understanding Society: Waves 1-14, 2009-2023 and Harmonised BHPS: Waves 1-18, 1991-2009, User Guide*. Colchester: University of Essex.

Valliant, R., Dever, J. A., & Kreuter, F. (2013). *Practical tools for designing and weighting survey samples (Vol. 1)*. New York: Springer. https://doi.org/10.1007/978-1-4614-6449-5

Wells, W., Xue, B., Lacey, R. & McMunn, A. (2025) Differences by ethnicity in the association between unpaid caring and health trajectories over 10 years in the UK Household Longitudinal Study. *J Epidemiol Community Health*, **79,** 94-101.

Wu M-J., Zhao, K. & Fils-Aime, F. (2022) Response rates of online surveys in published research: A meta-analysis. *Comp. Human Behav. Rep.*, **7**, 100206. https://doi.org/10.1016/j.chbr.2022.100206.

Zhang, L., Gagné, T. & McMunn, A. (2023). Changes in economic activity and mental distress among young adults during the COVID-19 pandemic: Differences between the first and second infection waves in the UK. *Plos one*, **18,** e0292540.

Table 1: Overall and subgroup dataset issued sample sizes, and respondent dataset sizes and response rates by mode.

	Overall	16-34 years	Aged 75+	Ethnic Minority	Low income
Issued sample N	30,848	8,664	3,092	7,376	6,593
Combined respondent N	22,445	5,326	2,341	4,613	4,512
Combined response rate (%)	73	61	76	63	68
Web respondent N	18,022	4,416	1,378	3,443	3,097
Web response rate (%)	59	51	45	47	47
CATI respondent N	4,423	910	963	1,170	1,415
CATI response rate (%)	14	11	31	16	21

Table 2. Overall CVs and their 95% confidence Intervals (in brackets) for the subgroup web and combined respondent datasets. As response propensity models underlying CVs differ between subgroups, in each case a comparator analysis of the overall dataset with the same response propensity model is also reported on: see paper text for further explanation.

	16-34 year old	Age 75+	Ethnic minority	Low income
Subgroup web dataset	0.282 (0.263 – 0.301)	0.409 (0.370 – 0.448)	0.263 (0.241 – 0.285)	0.351 (0.327 – 0.375)
Subgroup combined dataset	0.242 (0.226 – 0.257)	0.127 (0.106 – 0.148)	0.200 (0.183 – 0.216)	0.202 (0.186 – 0.218)
Overall web dataset, same response propensity model	0.217 (0.208 – 0.226)	0.205 (0.196 – 0.214)	0.249 (0.240 – 0.258)	0.237 (0.228 – 0.246)
Overall combined dataset, same response propensity model	0.163 (0.156 – 0.170)	0.150 (0.143 – 0.157)	0.178 (0.171 – 0.185)	0.164 (0.157 – 0.171)

Table 3. Means of absolute standardised non-response biases (MASBs) in non-response weighted respondent characteristics compared to benchmark sample inclusion weighted equivalents and their 95% confidence Intervals (in brackets) for overall and subgroup web and combined datasets. For the overall dataset and each subgroup, we also report *p* values from paired Wilcoxon sign tests evaluating the significance of differences in biases between web and combined datasets.

	Overall	16-34 year old	Age 75+	Ethnic minority	Low income
Web	0.003	0.031	0.069	0.026	0.037
	(0.002 – 0.004)	(0.022 – 0.041)	(0.045 – 0.093)	(0.019 – 0.033)	(0.023 – 0.051)
Combined	0.002	0.020	0.022	0.019	0.018
	(0.001 – 0.003)	(0.015 – 0.025)	(0.016 – 0.028)	(0.014 – 0.023)	(0.011 – 0.025)
Paired Wilcoxon sign test *p*	0.856	0.029	0.011	0.557	0.015

Table 4: Summary of evaluation results. * indicates the subgroup for which follow-ups have the greatest impact on the dataset quality components.

Quality component	Overall dataset	16-34 year olds	Aged 75+	Ethnic minority	Low income
RQ1: Response rates	Follow-ups increase dataset size by 23.7%	Follow-ups increase dataset size less than for the overall dataset (= 19.6%)	Follow-ups increase dataset size more than for the overall dataset (= 68.9%)*	Follow-ups increase dataset size more than for the overall dataset (= 34.0%)	Follow-ups increase dataset size more than for the overall dataset (= 44.7%)
RQ2: Dataset representativeness	Not computed	Improved by follow-ups in subgroup and comparator overall dataset analyses	Improved by follow-ups in subgroup and comparator overall dataset analyses*	Improved by follow-ups in subgroup and comparator overall dataset analyses	Improved by follow-ups in subgroup and comparator overall dataset analyses
RQ3: Under-represented subgroups	Not computed	Less impact of follow-ups in subgroup than comparator overall dataset analysis	Greater impact of follow-ups in subgroup than comparator overall dataset analysis*	Similar impact of follow-ups in subgroup and comparator overall dataset analysis	Similar impact of follow-ups in subgroup and comparator overall dataset analyses
RQ4: Biases after non-response weighting	Follow-ups do not reduce biases	Follow-ups reduce biases	Follow-ups reduce biases	Follow-ups do not reduce biases	Follow-ups reduce biases

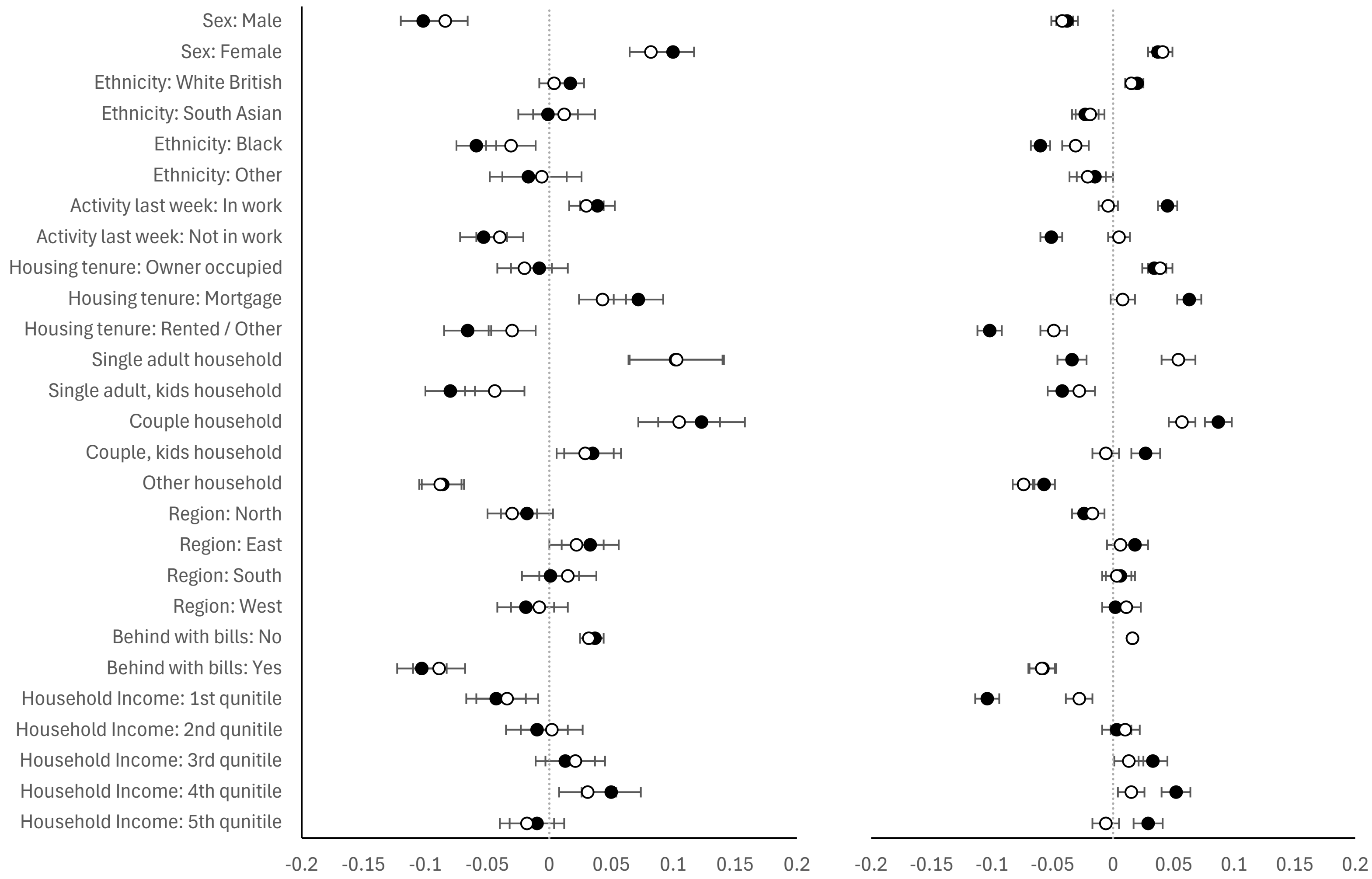
Figure 1. Unconditional covariate category CVs for: a) the Young adult subgroup dataset; and b) using the same covariates in the response propensity model, the overall dataset. Web respondents are represented by black circles, and combined (web plus CATI) respondents by white circles. Error bars indicate 95% CIs.

Figure 2. Unconditional covariate category CVs for: a) the Older adult subgroup dataset; and b) using the same covariates in the response propensity model, the overall dataset. Web respondents are represented by black circles, and combined (web plus CATI) respondents by white circles. Error bars indicate 95% CIs.

Figure 3. Unconditional covariate category CVs for: a) the Ethnic minority subgroup dataset; and b) using the same covariates in the response propensity model, the overall dataset. Web respondents are represented by black circles, and combined (web plus CATI) respondents by white circles. Error bars indicate 95% CIs.

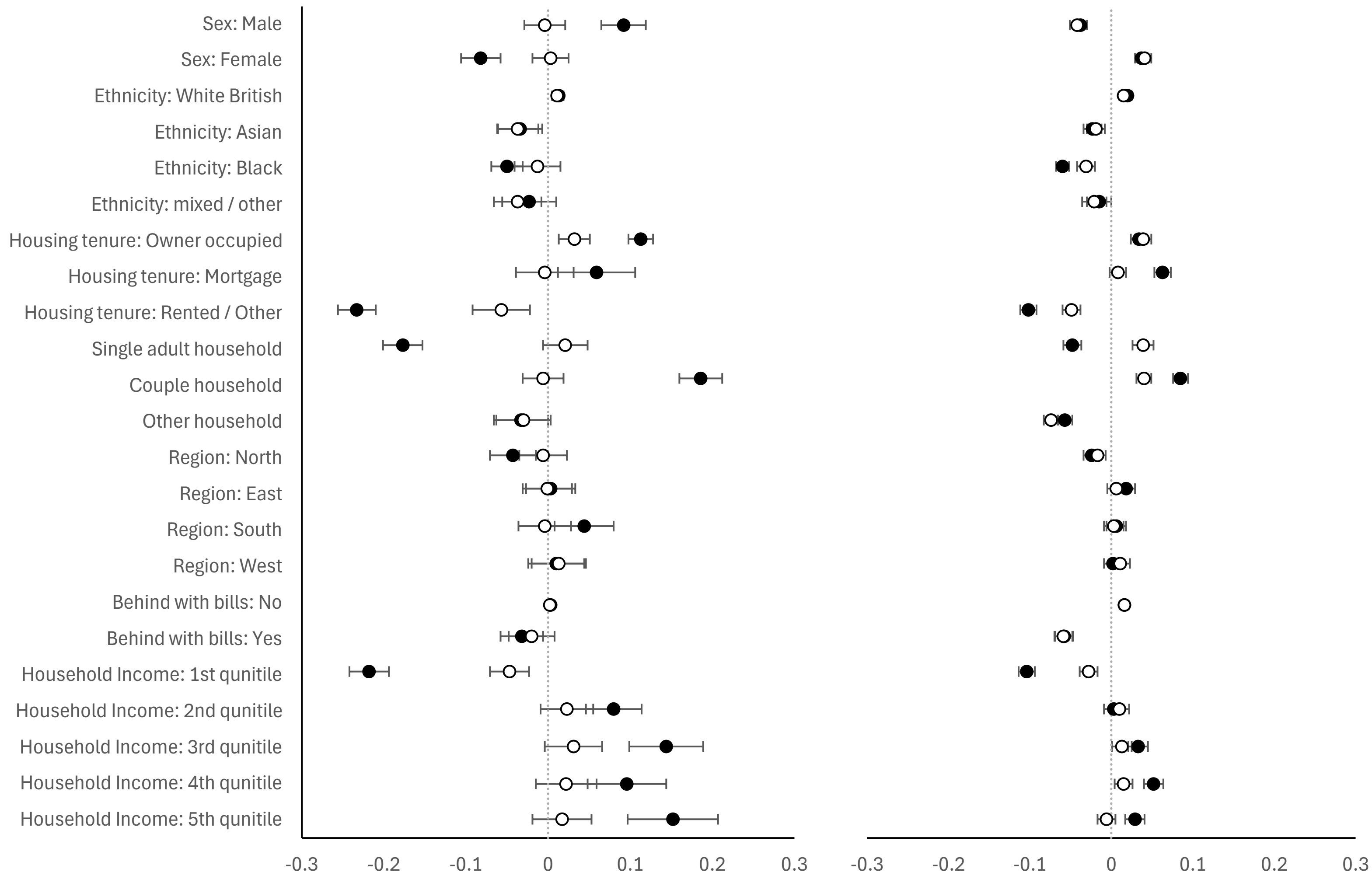
Figure 4. Unconditional covariate category CVs for: a) the Low income subgroup dataset; and b) using the same covariates in the response propensity model, the overall dataset. Web respondents are represented by black circles, and combined (web plus CATI) respondents by white circles. Error bars indicate 95% CIs.

Fig 1.



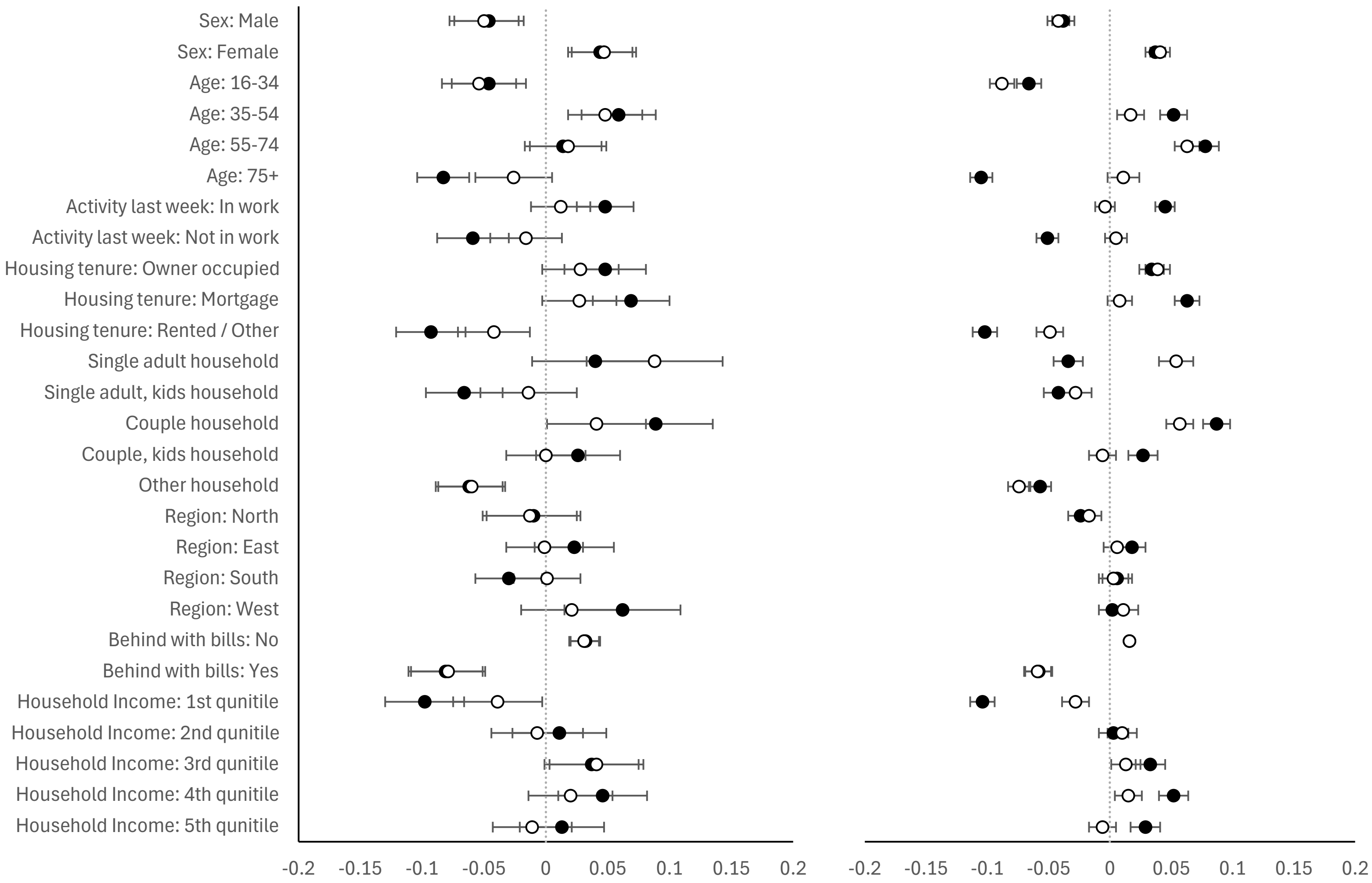
Sex: Male
Sex: Female
Ethnicity: White British
Ethnicity: South Asian
Ethnicity: Black
Ethnicity: Other
Activity last week: In work
Activity last week: Not in work
Housing tenure: Owner occupied
Housing tenure: Mortgage
Housing tenure: Rented / Other
Single adult household
Single adult, kids household
Couple household
Couple, kids household
Other household
Region: North
Region: East
Region: South
Region: West
Behind with bills: No
Behind with bills: Yes
Household Income: 1st qunitile
Household Income: 2nd qunitile
Household Income: 3rd qunitile
Household Income: 4th qunitile
Household Income: 5th qunitile
-0.2 -0.15 -0.1 -0.05 0 0.05 0.1 0.15 0.2
-0.2 -0.15 -0.1 -0.05 0 0.05 0.1 0.15 0.2

Fig 2.



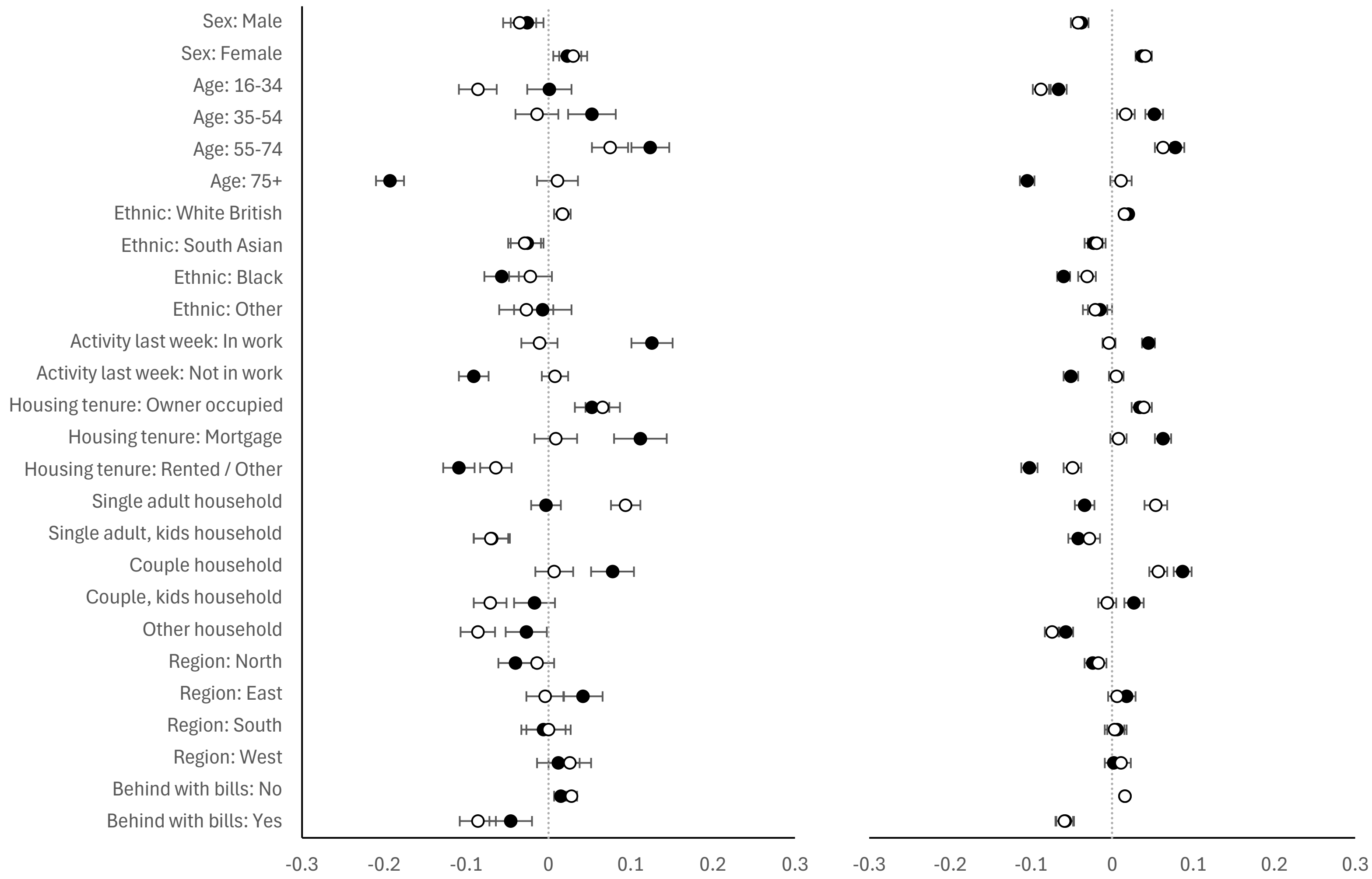
Sex: Male
Sex: Female
Ethnicity: White British
Ethnicity: Asian
Ethnicity: Black
Ethnicity: mixed / other
Housing tenure: Owner occupied
Housing tenure: Mortgage
Housing tenure: Rented / Other
Single adult household
Couple household
Other household
Region: North
Region: East
Region: South
Region: West
Behind with bills: No
Behind with bills: Yes
Household Income: 1st qunitile
Household Income: 2nd qunitile
Household Income: 3rd qunitile
Household Income: 4th qunitile
Household Income: 5th qunitile
-0.3
-0.2
-0.1
0
0.1
0.2
0.3
-0.3
-0.2
-0.1
0
0.1
0.2
0.3

Fig 3.



Sex: Male
Sex: Female
Age: 16-34
Age: 35-54
Age: 55-74
Age: 75+
Activity last week: In work
Activity last week: Not in work
Housing tenure: Owner occupied
Housing tenure: Mortgage
Housing tenure: Rented / Other
Single adult household
Single adult, kids household
Couple household
Couple, kids household
Other household
Region: North
Region: East
Region: South
Region: West
Behind with bills: No
Behind with bills: Yes
Household Income: 1st qunitile
Household Income: 2nd qunitile
Household Income: 3rd qunitile
Household Income: 4th qunitile
Household Income: 5th qunitile
-0.2 -0.15 -0.1 -0.05 0 0.05 0.1 0.15 0.2
-0.2 -0.15 -0.1 -0.05 0 0.05 0.1 0.15 0.2

Fig 4:


Sex: Male
Sex: Female
Age: 16-34
Age: 35-54
Age: 55-74
Age: 75+
Ethnic: White British
Ethnic: South Asian
Ethnic: Black
Ethnic: Other
Activity last week: In work
Activity last week: Not in work
Housing tenure: Owner occupied
Housing tenure: Mortgage
Housing tenure: Rented / Other
Single adult household
Single adult, kids household
Couple household
Couple, kids household
Other household
Region: North
Region: East
Region: South
Region: West
Behind with bills: No
Behind with bills: Yes
-0.3 -0.2 -0.1 0 0.1 0.2 0.3
-0.3 -0.2 -0.1 0 0.1 0.2 0.3

Appendix A1: Estimation and evaluation of analysis dataset sample inclusion weights

Sample inclusion weights for each wave are not released with the UKHLS datasets, nor are weights for web mode first sample members. Hence, sets of customised weights were created for use in the substantive analyses in the paper. These weights, termed cross-sectional (see below) sample inclusion weights, adjust the released sample inclusion enumeration weights for attrition from the sample (HHs not responding for two waves are not issued, nor are individuals who ask to be removed from the survey: see also main paper section 3.1). They also adjust for selection into the web first sample (individuals can request another mode). HH weight-sharing methods are then used to share these weights with unweighted HH members (those who enter HHs at waves after the inclusion enumeration weight was constructed), split (divide) these shared weights with remaining unweighted sample members (those in HHs where there is no weight to be shared) with similar characteristics, and post-stratify the split weights to relevant population estimates. The procedure is described below. Following this, evaluations of weight performance are reported.

A1 1. Weight estimation

The dataset of interest consists of UKHLS sample members offered web first interviewing at wave 11 or 12 during COVID-19 pandemic in 2020. The input to the weight produced is the released main survey wave 6 inclusion enumeration weight. Since only those enumerated at wave 6 have this weight, and only these individuals are included in the sample at wave 7, it is also the wave 7 sample inclusion weight for these individuals. For each of the wave 11 and wave 12 samples, this weight is adjusted for wave 11 / 12 sample non-inclusion, i.e. multiplied by 1 / the estimated probability of wave 11 / 12 sample inclusion given inclusion in the wave 7 sample, with inclusion probabilities estimated using regression modelling with 11 survey-measured, auxiliary covariates (see Table 1 for covariates, and section A1.2. for details of methods used to select final models and estimate inclusion probabilities).

To produce the final weight to be used in the substantive analyses, the post-stratified cross-sectional sample inclusion weight, four further steps are then undertaken. First, the weights estimated above are shared with unweighted wave 11/12 web first HH members, i.e. (final) weights in HH including unweighted web first members equal the sum of existing weights across web first HH members divided by the number of web first HH members with or without existing weights. This method provides unbiased estimates from HH probability samples and is used in a number of panel surveys (Ernst 1983; Lavallee 1995, 2007; Schonlau et al. 2013; Zhang 2022).

Second, each of these shared weights is split with remaining unweighted sample members (those in HHs without a sample inclusion weight to share) with similar survey measured characteristics using the procedure of Moore & Clarke (2025). This procedure: 1) uses regression modelling of existing weights to predict ‘synthetic’ weights for unweighted individuals given their survey measured characteristics; 2) matches / clusters existing and synthetic weights; and 3) splits (divides) the existing weights in each cluster with the unweighted individuals in the cluster. It will produce unbiased estimates of the population estimated by the shared weights assuming that shared weighted and unweighted sample members are exchangeable given the same characteristics, and that clusters of sample members with similar characteristics are identified adequately.

Third, these two split weights are combined, rescaled to have a mean of one, and then post-stratified to Great Britain population estimates for 2020 (Office for National Statistics 2024; National Records of Scotland 2024; Northern Ireland Statistics and Research Agency 2024). Population estimates for the cross tabulation of Region (4 categories), Sex (2 categories) and Age (5 categories) are used in the post-stratification (a total of 40 strata). Fourth, in the final step, the post-stratified weights are again rescaled to have a mean of one.

A1 1.1.3. Regression model selection methods

To identify the final regression models used to estimate inclusion / selection probabilities, Lasso procedures are used. Lasso procedures (Tibshirani 1996; Steyerberg et al. 2001) are regularised regression methods. As with other regularised regression methods for binary data (i.e. 0 = sample non-inclusion, 1 = sample inclusion), they maximise the joint probability of the model parameters given the observed data similar to maximum likelihood methods, but in addition impose a regularisation penalty on model complexity (Ahrens et al. 2020). Due to the imposition of this penalty, such methods tend to outperform maximum likelihood methods in terms of out of sample prediction, as reducing model complexity and inducing shrinkage bias decreases prediction error. In doing so, they also address the problem of model overfitting: high in-sample fit, but poor prediction performance on unseen data.

Regularised regression methods incorporate tuning parameters that determine the amount and form of regularisation penalty. Several techniques exist to choose the value of these parameters. The first is cross-validation, which explicitly evaluates out of sample prediction performance. The data are split into training and validation datasets. The models for different values of the tuning parameters are then estimated and variables selected using the training dataset. Next, they are applied to the validation dataset, and performance quantified (Ahrens et al. 2020). The second technique is the use of information criteria. These are interpretable as likelihood methods that penalise the number of parameters in models. Again, models for different tuning parameters are estimated and variables selected, then the best performing is chosen based on information criteria value. When producing the sample inclusion weights, we use information criteria techniques to choose tuning parameter values and identify models for estimating inclusion probabilities. Specifically, we utilise the Extended Bayesian Information Criterion (EBIC: Chen & Chen 2008), because simulations show that in the majority of scenarios they perform better than other similar options in terms of model identification (see Ahrens et al. 2020). We do not use cross validation methods because the size of analysis datasets prevents their division into training and validation datasets (see Moore et al. 2024 for further justification of these methods in the current context). We use the Stata 18 package ‘lassologit’ (Ahrens et al. 2020) to perform analyses.

The above techniques require that predictors are standardised so that they have unit variance. Hence, when modelling inclusion probabilities for weight estimation we first convert all multi-category predictors into dummy variables. Once the selected model is identified, we then extend it to all selected covariate categories whether they were selected or not: in previous work, we have found that this approach reduces biases (relative to benchmarks) in weighted estimates (unpublished results). After final model identification, we use post-Lasso estimation to estimate inclusion / selection probabilities for weight estimation, because Lasso estimated coefficients are subject to attenuation bias (Ahrens et al. 2020). Specifically, we use probit models, with inclusion probabilities predicted using model coefficients and sample member characteristics.

A1 2. Evaluations of weight performance

A1 2.1. Evaluation methods

The custom sample inclusion weights are evaluated in two ways. First, they are evaluated against internal benchmarks. This approach involves using weighted estimates of survey measured characteristics for the sample from a given survey wave as benchmarks and evaluating the performance of equivalent weighted estimates from a comparator dataset (from the same or following waves) in recovering them (see Moore et al. 2024 for an example of this approach). Note that in this instance the comparisons possible are restricted because different weighted datasets often represent different populations. We cannot evaluate cross-sectional sample inclusion weights from a given wave using the previous wave sample and its equivalent weights as the benchmark due to HH joiners and refreshment samples entering the former dataset. Similarly, concerning interim weights estimated in the course of final weight estimation, shared weight datasets cannot be compared to unshared weight (i.e. longitudinal sample inclusion weight) datasets due to the former including HH joiners, nor can we compare longitudinal sample inclusion-weighted datasets to previous wave equivalent benchmarks because there are no previous wave benchmarks.

We can, however, compare the split weighted dataset to the (benchmark) shared weight dataset, because the weight splitting procedure divides the shared weights with unweighted sample members with similar characteristics, so such weights produce estimates for the same population. In this evaluation, weighted estimates of incidence in the dataset for each of the categories of the 11 covariates included in models estimating the sample inclusion / selection probabilities underlying weight estimation (see Table 1) are computed for benchmark and comparator datasets and compared. Despite partial dependencies between datasets (the benchmark dataset is a subset of the comparator dataset), is independent samples T tests are used for statistical inference because tests accounting for such dependencies (see Moore et al. (2024) and Appendix A3 can only be justified in a design-based framework when the comparator dataset is instead a subset of the benchmark dataset. Note that these latter tests will be anti-conservative. Survey design (Primary Sampling Unit and Strata) is accounted for in estimation and testing. In addition, as overall performance measures, we report means across all considered covariate categories of absolute differences between comparator and benchmark estimates standardised by benchmark estimate standard deviations (MASBs) and their 95% CIs.

The second approach to evaluating weight performance involves comparing weighted estimates of survey measured characteristics to external benchmarks. In these evaluations, as external benchmarks the Region × Sex × Age cross-tabulation of UK population estimates for the given year that are utilised to post stratify the (split) cross-sectional sample inclusion weights to produce the final weight used in substantive analyses (see previously) are used. The population totals in each of the cells are converted to incidences (= cell total / overall population total) and compared to equivalent incidences computed using survey measured characteristics for (split) cross-sectional sample inclusion weighted and final post-stratified weighted comparator datasets. One sample T tests are used for statistical inference, with survey design (Primary Sampling Unit and Strata for each weight) accounted for in estimation and testing. As overall performance measures, we also report the means of absolute differences (MABs) across all crosstabulation cells and their 95% confidence intervals (CIs).

A1 2.2. Evaluation results

The evaluations of split, cross-sectional sample inclusion weighted estimates compared to shared cross-sectional weighted benchmarks indicate that the means of the absolute biases standardised by benchmark estimate standard deviations (MASB) is 0.003 (95% CI 0.002 – 0.004). No significant differences between estimates for individual characteristics are observed (Table A2 2). Hence, the split weights perform well at recovering same wave shared weight benchmarks.

The evaluations of the split, cross-sectional sample inclusion weights compared to external population estimate benchmarks indicate that the mean absolute difference (MAB) is 0.004 (95% CI 0.003 – 0.005). 34 significant differences between estimates are observed, but the largest is 0.01 (Table A2 3). Note also that significant differences are to be expected given the number of tests and the likelihood of type 1 errors. The evaluations of the post-stratified, split, cross-sectional sample inclusion weights compared to external population estimate benchmarks indicate that the MAB is 0.001 (95% CI 0.001 – 0.001). No significant differences found between estimates are observed (Table A2 3). Hence, both the split, cross-sectional sample inclusion weights and the final post-stratified weights used in substantive analyses perform well at recovering relevant external population estimate benchmarks.

A1 3. References

Ahrens, A., Hansen, C. B., & Schaffer, M. E. (2020) lassopack: Model selection and prediction with regularized regression in Stata. *The Stata Journal*, 20*:* 176-235.

Chen, J. & Chen, Z. (2008) Extended Bayesian information criteria for model selection with large model spaces. *Biometrika*, 95: 759–771. DOI: 10.1093/biomet/asn034

Ernst, L.R. (1989) Weighting issues for longitudinal household and family estimates. In *Panel Surveys* (Kasprzyk, D., Duncan, G., Kalton, G. & Singh, M.P. (eds), p. 135–159. Wiley & Sons, New York,

Lavallee, P. (1995) Cross-sectional weighting of longitudinal surveys of individuals and households using the weight share method*. Survey Methodology*, 21:25–32.

Lavallée, P. (2007) *Indirect Sampling,* New York: Springer

Moore, J.C., Burton, J., Crossley, T. F., Fisher, P., Gardiner, C., Jäckle, A., & Benzeval, M. (2024) *Assessing Bias Prevention and Bias Adjustment in a Sub-Annual Online Panel Survey*. Understanding Society Working Papers Series 2024-04.

Moore, J.C. & Clarke, P. (2025) Two new solutions to the zero survey weights problem. Understanding Society Working Papers Series 2025-??.

National Records of Scotland (2024) *Mid Year Population Estimates Time Series Data.* https://www.nrscotland.gov.uk/publications/population-estimates-time-series-data/

Northern Ireland Statistics and Research Agency (2024) *2023 Mid-year Population Estimates for Northern Ireland*. https://www.nisra.gov.uk/system/files/statistics/MYE23-bulletin_1.pdf

Office for National Statistics (2024) *Mid 2011 to mid-2023 detailed time series edition of population estimates for England and Wales.* https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/estimatesofthepopulationforenglandandwales

Schonlau, M., Kroh, M., & Watson, N. (2013) The implementation of cross-sectional weights in household panel surveys. *Statistics Surveys*, 7: 37-57.

StataCorp (2023) Stata Lasso Reference Manual Release 18. StataCorp LLC, College Station, Texas.

Steyerberg, E.W., Eijkemans, M.J.C. & Habbema, J.D.F. (2001) Application of shrinkage techniques in Logistic regression analysis: a case study. *Statistica Neerlandica,* 55: 76–88. DOI: 10.1111/1467-9574.00157.

Tibshirani, R. (1996) Regression and shrinkage via the Lasso, *Journal of the Royal Statistical Society, Series B*, 58, 267-288.

Zhang, L. C. (2022). *Graph sampling*. Chapman and Hall/CRC.

Table A1 1: Auxiliary covariates used in weight estimation and evaluations of weight performance and their categorisations.

	Categorisation
Sex	1) Male; 2) Female
Age	1) 16-34; 2) 35-54; 3) 55-74; 4) 75+
Ethnicity	1) White British; 2) Black; 3) South Asian; 4) Other.
Activity last week	1) In work; 2) Not in Work
Tenure	1) Owner occupied; 2) Mortgage; 3) Rented / Other.
HH Structure	1) 1 adult; 2) 1 adult, kids; 3) Couple, no kids; 4) Couple, kids; 5) Other
Region	1) North; 2) East; 3) South; 4) West.
Behind with bills	1) No; 2) Yes
HH income	1) 1st quintile; 2) 2nd quintile; 3) 3rd quintile; 4) 4th quintile; 5) 5th quintile
Behind with Council Tax	1) Yes; 2) No
Urban	1) Urban; 2) Rural
Nos. rooms in HH	Continuous
Household size	
Member of Ethnic minority boost sample	1) No; 2) Yes

Table A1 2: Split cross-sectional sample inclusion weighted estimates of member survey measured characteristics compared to equivalent shared cross-sectional sample inclusion weighted benchmarks. 'Bch.' is the benchmark estimate (standard error in brackets). 'Diff' is the difference between the sample inclusion weighted estimate and the benchmark estimate. * equals *P*<0.05.

Variable	Bch. (se)	Diff
Sex: Male	0.483 (0.003)	-0.001
Age: 0-15	0.135 (0.002)	-0.001
Age: 16-34	0.225 (0.002)	-0.000
Age: 35-54	0.245 (0.002)	-0.000
Age: 55-74	0.268 (0.002)	0.003
Age: 75+	0.127 (0.002)	-0.002
Ethnicity: white British	0.862 (0.002)	0.000
Ethnicity: Black	0.053 (0.001)	0.000
Ethnicity: south Asian	0.023 (0.001)	0.000
Ethnicity: Other	0.063 (0.001)	-0.000
Activity Last Week: In work	0.450 (0.003)	0.001
Housing tenure: Owner occupied	0.338 (0.003)	0.003
Housing tenure: Mortgage	0.351 (0.003)	0.002
Housing tenure: Rented/Other	0.311 (0.003)	-0.005
Household Structure: 1 adult	0.161 (0.002)	-0.002
Household Structure: 1 adult, kids	0.050 (0.001)	-0.000
Household Structure: Couple	0.244 (0.002)	0.002
Household Structure: Couple, kids	0.231 (0.002)	-0.000
Household Structure: Other	0.315 (0.003)	0.000
Region: North	0.329 (0.003)	-0.000
Region: South	0.263 (0.002)	0.000
Region: East	0.246 (0.002)	-0.002
Region: West	0.161 (0.002)	0.002
Household Location: Urban	0.751 (0.002)	-0.001
Behind with bills: No	0.928 (0.001)	0.001
Behind Council Tax: Yes	0.071 (0.001)	-0.001
Household income: 1st quintile	0.236 (0.002)	-0.003
Household income: 2nd quintile	0.211 (0.002)	-0.000
Household income: 3rd quintile	0.189 (0.002)	0.000
Household income: 4th quintile	0.180 (0.002)	0.001
Household income: 5th quintile	0.184 (0.002)	0.002

Table A1 3: Split ('SP') and post-stratified (PS) cross-sectional sample inclusion weighted estimates of issued sample survey measured characteristics compared to equivalent UK estimated population incidence benchmarks. 'Bch.' is the benchmark population estimate. 'SP. diff' is the difference between the split weight estimate and the benchmark estimate. 'PS. diff' is the difference between the split weight estimate and the benchmark estimate. * equals $P<0.05$.

Strata	Bch.	SP. diff	PS. Diff
Region1_Sex1_Age1	0.030	-0.007*	-0.001
Region1_Sex1_Age2	0.036	0.001	-0.001
Region1_Sex1_Age3	0.037	0.003*	-0.001
Region1_Sex1_Age4	0.035	0.009*	-0.001
Region1_Sex2_Age5	0.016	0.002*	-0.000
Region1_Sex2_Age1	0.029	-0.008*	-0.001
Region1_Sex2_Age2	0.036	-0.000	-0.001
Region1_Sex2_Age3	0.039	0.004*	-0.001
Region1_Sex2_Age4	0.036	0.010*	-0.001
Region1_Sex2_Age5	0.019	0.003*	-0.001
Region2_Sex1_Age1	0.026	-0.007*	-0.001
Region2_Sex1_Age2	0.031	-0.002*	-0.001
Region2_Sex1_Age3	0.034	-0.004*	-0.001
Region2_Sex1_Age4	0.030	0.003*	-0.001
Region2_Sex2_Age5	0.010	0.005*	-0.000
Region2_Sex2_Age1	0.025	-0.005*	-0.001
Region2_Sex2_Age2	0.031	-0.002	-0.001
Region2_Sex2_Age3	0.035	-0.001	-0.001
Region2_Sex2_Age4	0.031	0.003*	-0.001
Region2_Sex2_Age5	0.014	0.005*	-0.000
Region3_Sex1_Age1	0.028	-0.010*	-0.001
Region3_Sex1_Age2	0.036	-0.004*	-0.001
Region3_Sex1_Age3	0.039	-0.010*	-0.001
Region3_Sex1_Age4	0.027	0.003*	-0.001
Region3_Sex2_Age5	0.009	0.005*	-0.000
Region3_Sex2_Age1	0.026	-0.010*	-0.001
Region3_Sex2_Age2	0.037	-0.007*	-0.001
Region3_Sex2_Age3	0.040	-0.007*	-0.001
Region3_Sex2_Age4	0.029	0.003*	-0.001
Region3_Sex2_Age5	0.012	0.005*	-0.000
Region4_Sex1_Age1	0.015	-0.005*	-0.000
Region4_Sex1_Age2	0.019	-0.002*	-0.001
Region4_Sex1_Age3	0.021	-0.003*	-0.001
Region4_Sex1_Age4	0.020	0.001	-0.001
Region4_Sex2_Age5	0.007	0.002*	-0.000
Region4_Sex2_Age1	0.015	-0.004*	-0.000
Region4_Sex2_Age2	0.019	-0.002*	-0.001
Region4_Sex2_Age3	0.021	-0.001	-0.001
Region4_Sex2_Age4	0.021	0.002*	-0.001
Region4_Sex2_Age5	0.009	0.003*	-0.000

Appendix A2. Missing values and imputation strategy

This appendix provides a detailed description of the missing data in the UKHLS main survey datasets used in the analyses. We also document the imputation process using information from previous and following waves and a chained equation imputation model for the remaining missingness.

A2 1. Pre-imputation datasets

This section documents the generation of the initial datasets prior to the imputation. The following paragraphs detail the exclusion of ineligible cases and other individuals not issued to the field from the sample for each relevant wave of the UKHLS main survey. Although the analysis of the paper is based on the adults, i.e., sample members aged 16 and over, who were invited to complete the individual questionnaire, the tables provided in this appendix include all sample members regardless of their age or whether they were issued to web-first or not. This is because the datasets with all sample members were required to compute the sample inclusion weights (see Appendix A1).

For each wave to be used in the weighting or analysis, we excluded from the sample the individuals who had become ineligible and those who could not be issued to the field or were not assigned to a household. The ineligible sample members were those found to have died before the fieldwork or moved out of scope, i.e., relocated abroad. In addition, some individuals, primarily the households participating online, who had not provided a valid address could not be issued to the interviewers and were excluded. Finally, a small group of sample members were not assigned to a household and were dropped from the datasets.

Table A2 1 shows the number of individuals excluded because they were ineligible or because they could not be issued to the field in the main survey.

Table A2 1. UKHLS main survey: Exclusion of sample members of the analysis by wave

	Initial sample (n)	Ineligible[1]		Not issued/no household assigned	
		n	%	n	%
Wave 7 (2015-17)	90,021	786	0.9	6,815	7.6
Wave 11 (2019-21)	58,499	469	0.8	2,485	4.2
Wave 12 (2020-22)	55,891	446	0.8	2,479	4.4

Note: [1]Sample members ineligible for an interview are those who died before the fieldwork or moved abroad. [2]Not issued includes households for which there was no address and could not be issued to the interviewers, plus the individuals who were not assigned to a household.

A2 2. Missing values and imputation strategy

Auxiliary information from respondents and non-respondents is necessary to compute the survey weights and conduct the representativeness analysis. *Understanding Society* is a household survey where household members are invited to complete a household grid and a questionnaire, and each adult is asked to complete an individual questionnaire at each wave. Therefore, for each wave, we have some information available from all household members from responding households, i.e., households that completed the household grid and questionnaire. However, for sample members at a given wave there is no information available if no one in the household participated in the survey. In addition, survey respondents can refuse to answer a particular question in the household questionnaire (item non-response). Table A2 2 (column 2) shows the percentage of missing values (item non-response) for the variables for each wave of the main survey involved in the weights and analysis. The level of missingness varies between 15.4% of sex at wave 11 and 33.7% of being behind with council tax bills at wave 7. The percentage of sample members with all the information missing is 27.3% at wave 7, 15.4% at wave 11, and 17.8% at wave 12.

Table A2 2. UKHLS main survey: Percentage of missing values for each variable as observed (Obs.) and after the imputation from previous and following waves (IPFW) by wave

	Wave 7		Wave 11		Wave 12	
	Obs.	IPFW	Obs.	IPFW	Obs.	IPFW
All variables missing	27.3	0.4	15.4	0.0	17.8	0.0
Sex	27.3	0.5	15.5	0.0	17.8	0.0
Age	27.4	0.6	15.5	0.1	17.8	0.1
Ethnic background	28.7	2.6	17.3	2.5	19.8	2.7
Region (GOR)	27.3	13.9	15.5	2.1	17.8	3.1
Urbanicity	27.3	13.0	17.2	1.3	17.8	2.2
Employment status	27.6	13.4	15.8	1.6	18.1	2.5
Household income	28.8	13.8	19.7	3.0	22.6	3.9
Behind with bills	29.2	13.9	20.1	3.2	23.1	4.1
Behind with council tax	33.7	19.4	25.3	9.2	28.0	10.0
Household type	28.8	13.8	19.7	3.0	22.6	3.9
Tenure status	29.3	14.1	20.7	3.2	23.1	4.2
Number of rooms	28.9	13.9	19.8	3.1	22.8	4.0
Base (n)	82,420	82,420	55,545	55,545	52,966	52,966

We developed a two-step strategy to impute the missing values. This strategy consisted of 1) an imputation based on information from the previous and following waves (IPFW), and 2) a model-based imputation using a multivariate imputation chained equations (MICE) model for the remaining missing values.

A 2.1. Imputation from previous and following waves (IPFW)

An advantage of working with a longitudinal survey is the availability of information for those who are part of the panel but do not participate in a given wave. The missing values can be imputed using a last observation carried forward, or baseline observation carried forward. However, this assumes that the state of sample members has not changed between the observation and the wave with a missing value, which might introduce some bias in the estimates (Kenward & Molenberghs, 2009; Saha & Jones, 2009). We limited the number of previous and following waves used in the imputation to enhance the use of the panel information while minimising the risks associated with this. For sex, ethnicity and age, we used the information from the previous or following six waves since these characteristics will not change. For the rest of the household and individual level characteristics, we limited the range of values used for the imputation to the previous and following two waves.

Table A2 2 (columns IPFW) presents the remaining missing values if the information is imputed using the previous and following waves for the datasets of the main survey. The table show that the level of missingness decreases substantially when information from the previous and following waves is considered. For a comparison of the distribution of the observed and the imputed values using the IPFW, see Table A2 5.

After the first imputation stage, sample members for whom all the information remained missing were excluded from the analysis and, therefore, were not included in the model-based imputation. Table A2 3 presents the number of sample members excluded by wave in the main survey datasets.

Table A2 3. UKHLS main survey: Sample members for whom no information was available

		No information	
Wave	**Base (n)**	**n**	**%**
Wave 7	82,420	356	0.4
Wave 11	55,545	16	0.0
Wave 12	52,966	15	0.0

A2 2.2. Model-based imputation strategy

The remaining missing cases were imputed using a model-based strategy. Given the relatively low level of missing information after the first stage of imputation at most waves, a single imputation using multivariate imputation by chained equations (MICE) was considered. The imputation strategy accounted for the multilevel structure of the data that involves sample members (level-1) nested in households (level-2). This is because using a single-level imputation method that ignores the hierarchical structure of the data can result in a conceptual problem where sample members from the same household exhibit different household characteristics (Van Buuren, 2018). Thus, we implemented a two-step imputation to produce a consistent imputation that considered the multilevel data structure (Grund et al., 2018). First, a dataset at the household level that included the household characteristics and an average of the level-1 predictors (i.e., sex, age, employment status and ethnic background) was used to impute the missing values of the household variables. Second, we produced an individual-level dataset that included the individual-level predictors and the household-level variables imputed in the household-level dataset. This dataset was used to impute the individual-level variables. Table A2 4 presents the variables used in the imputation models. We used a propensity mean matching technique for the imputation of continuous, semi-continuous, and dummy variables (Austin & van Buuren, 2023; Vink et al., 2014). The categorical variables were imputed using multinomial logistic regression models. After these imputations, individuals with missing Age, Ethnic background or Income covariate values after the IPFW stage were removed from the analysis sample. This approach was taken to ensure that only individuals for which we had reported information for these subgroup defining covariates were considered, but also to make use of the information on other covariates for these individuals in the MICE imputations (see main paper text for numbers of individuals removed from the analysis sample).

Table A2 4. Specification of the MICE models

Household-level dataset imputation	
Variables	**Role and model**
Average of individual characteristics: female, age, being employed, white, Asian and black. Urbanicity, household income, being behind with bills, being behind with council tax and number of rooms.	**Imputed** using a propensity mean matching technique with 10 donors (nearest neighbor).
Region (GOR), household type and tenure status.	**Imputed** using a multinomial logistic regression model.
Household size, ethnic sample.	**Regular**. Not imputed but included as a predictor in the models.
Individual-level dataset imputation	
Variables	**Role and Model**
Sex and age	**Imputed** using a propensity mean matching technique with 10 donors (nearest neighbor).
Ethnic background	**Imputed** using a multinomial logistic regression model.
From the household-level dataset imputation: urbanicity, household income, being behind with bills, being behind with council tax, number of rooms, region (GOR), household type, tenure status, Household size, ethnic sample.	**Regular.** Not imputed but included as a predictor in the models.

Finally, we provide a summary of the differences between the observed and imputed values at the different stages of the imputation. For this purpose, Table A2 5 presents the distribution of the variables involved in the imputation at the different stages of the process for the main survey. In most cases, the imputation, both the IPFW and the model-based, had a minimal impact on the distribution of the variables. The difference between the observed and complete distribution of the variables results in less than a percentage point. However, there are a few exceptions. At waves that registered a higher level of missingness (e.g., main survey wave 7), the variables tenure status and household type showed some differences. Regarding tenure status, the distribution of imputed, both IPFW and MICE, exhibit a higher proportion of individuals renting their accommodation and a lower prevalence of people who own their houses compared to the observed distribution. In terms of household type, the imputed values show a lower proportion of one couple with no children households and a higher proportion of other households. Finally, in the main survey datasets, the imputed values show a higher proportion of persons with an ethnic minority background compared to the observed ones. These differences indicate that, in some instances, the imputation has increased the presence of sub-groups that are usually affected by attrition to a greater extent (Cabrera-Álvarez et al., 2023), suggesting that the imputation might have improved the overall representativeness of the datasets.

Table A2 5. UKHLS main survey: Distributions of the imputed variables by wave. Percentages of individuals without values for Age, Ethnic background or Income covariates after the IPFW procedure are excluded from the dataset, so percentage imputed by the MICE procedure are not reported.

	(a) Observed	**(b) IPFW**	**(a) + (b)**	**(c) MICE**	**Complete (a) + (b) + (c)**
Wave 7 (2015-17)					
Sex					
Male	48.0	48.8	48.1	27.3	48.1
Female	52.0	51.2	51.9	72.7	51.9
Base (n)	58,624	10,824	69,448	11	69,459
Age					
16-24	43.9	54.8	45.6		45.6
25-44	33.9	31.8	33.6		33.6
45-64	15.5	8.3	14.3		14.3
65+	6.7	5.1	6.4		6.4
Base (n)	58,618	10,841	69,459		69,459
Ethnic background					
White	72.9	60.8	71.0		71.0
Asian	13.9	17.7	14.5		14.5
Black	5.4	9.9	6.1		6.1
Mixed & Other	7.8	11.6	8.4		8.4
Base (n)	58,625	10,834	69,459		69,459
Region (GOR)					
North East	3.3	2.5	3.2	5.2	3.2
North West	10.4	9.9	10.4	7.2	10.3
Yorkshire and the Humber	8.8	7.1	8.6	7.5	8.6
East Midlands	6.7	7.1	6.8	9.9	6.8
West Midlands	9.0	8.7	8.9	8.8	8.9
East of England	8.0	7.2	7.9	7.9	7.9
London	15.1	26.2	16.7	5.9	16.6
South East	11.4	10.1	11.2	12.7	11.2
South West	7.2	5.9	7.1	6.6	7.1
Wales	6.2	8.3	6.5	9.5	6.5
Scotland	7.7	7.1	7.6	8.5	7.7
Northern Ireland	6.0	0.0	5.1	10.2	5.2
Base (n)	58,610	10,155	68,765	694	69,459
Urbanicity					
Urban	78.0	83.4	78.8	100.0	78.8
Rural	22.0	16.6	21.2	0.0	21.2
Base (n)	58,610	10,845	69,455	4	69,459
Employment status					
Yes	45.2	47.0	45.5	34.0	45.5
No	54.8	53.0	54.5	66.0	54.5
Base (n)	58,453	10,818	69,271	188	69,459
Income					
Up to £1,999	25.3	54.7	30.3		30.3
£2,000-£2,999	24.8	25.1	24.8		24.8
£3,000-£4,999	32.8	16.0	29.9		29.9
£5,000 and over	17.1	4.3	14.9		14.9
Base (n)	57,495	11,964	69,459		69,459
Behind with bills					
Up to date with bills	94.6	90.4	93.9	95.6	93.9
Behind with some bills	5.4	9.6	6.1	4.4	6.1
Base (n)	57,231	12,092	69,323	136	69,459
Behind with Council Tax					
Yes	7.0	11.8	7.8	5.1	7.6
No	93.0	88.2	92.2	94.9	92.4
Base (n)	53,626	11,387	65,013	4,446	69,459
Household type					

	(a) Observed	(b) IPFW	(a) + (b)	(c) MICE	Complete (a) + (b) + (c)
1 adult, no children	10.4	8.5	10.0	0.0	10.0
1 adult, children	6.1	9.1	6.6	0.0	6.6
Couple, no children	20.6	13.4	19.4	0.0	19.4
Couple with children	32.3	30.6	32.0	0.0	32.0
Other households	30.7	38.4	32.1	0.0	32.1
Base (n)	57,495	11,964	69,459	0	69,459
Tenure status					
Owned outright	28.3	18.0	26.5	17.2	26.4
Owned on mortgage	40.1	36.5	39.5	28.5	39.5
Rented and others	31.6	45.5	34.0	54.3	34.1
Base (n)	57,095	12,108	69,203	256	69,459
Number of rooms					
1	41.3	48.1	42.4	70.5	42.5
2	36.4	32.6	35.7	19.4	35.7
3	15.5	12.8	15.0	9.3	15.0
4	4.7	4.9	4.8	0.8	4.8
5 or more	2.1	1.6	2.0	0.0	2.0
Base (n)	57,421	11,909	69,330	129	69,459
Wave 11 (2019-21)					
Sex					
Male	47.5	48.5	47.6	50.0	47.6
Female	52.5	51.5	52.4	50.0	52.4
Base (n)	45,596	7,015	52,611	2	52,613
Age					
16-24	40.3	52.4	41.9		41.9
25-44	33.7	31.4	33.4		33.4
45-64	18.1	9.2	16.9		16.9
65+	7.9	7.1	7.8		7.8
Base (n)	45,593	7,020	52,613		52,613
Ethnic background					
White	75.9	62.0	74.1		74.1
Asian	12.9	18.5	13.6		13.6
Black	4.0	9.9	4.8		4.8
Mixed & Other	7.2	9.7	7.5		7.5
Base (n)	45,598	7,015	52,613		52,613
Region (GOR)					
North East	3.5	3.1	3.4	4.3	3.4
North West	10.3	11.3	10.5	6.1	10.4
Yorkshire and the Humber	8.9	10.0	9.0	14.4	9.0
East Midlands	7.1	5.9	6.9	8.4	7.0
West Midlands	8.8	11.8	9.1	7.8	9.1
East of England	8.6	7.5	8.5	7.6	8.5
London	13.2	18.2	13.8	3.3	13.8
South East	11.7	11.4	11.6	11.6	11.6
South West	7.7	6.2	7.5	7.3	7.5
Wales	6.2	6.8	6.3	10.6	6.3
Scotland	7.9	7.8	7.9	8.9	7.9
Northern Ireland	6.2	0.0	5.4	9.6	5.5
Base (n)	45,578	6,640	52,218	395	52,613
Urbanicity					
Urban	76.9	72.0	76.1	75.0	76.1
Rural	23.1	28.0	23.9	25.0	23.9
Base (n)	44,656	7,953	52,609	4	52,613
Employment status					
Yes	46.3	45.1	46.2	42.0	46.2
No	53.7	54.9	53.8	58.0	53.8
Base (n)	45,410	7,091	52,501	112	52,613
Income					
Up to £1,999	21.5	39.4	24.5		24.5

	(a) Observed	(b) IPFW	(a) + (b)	(c) MICE	Complete (a) + (b) + (c)
£2,000-£2,999	21.0	26.7	22.0		22.0
£3,000-£4,999	34.7	24.7	33.0		33.0
£5,000 and over	22.8	9.3	20.5		20.5
Base (n)	43,678	8,935	52,613		52,613
Behind with bills					
Yes	93.2	87.2	92.2	84.5	92.2
No	6.8	12.8	7.8	15.5	7.8
Base (n)	43,460	9,043	52,503	110	52,613
Behind with Council Tax					
Yes	6.8	12.5	7.8	23.4	8.8
No	93.2	87.5	92.2	76.6	91.2
Base (n)	40,622	8,678	49,300	3,313	52,613
Household type					
1 adult, no children	11.5	9.2	11.1	0.0	11.1
1 adult, children	4.4	8.7	5.2	0.0	5.2
Couple, no children	22.7	10.8	20.7	0.0	20.7
Couple with children	28.7	30.7	29.0	0.0	29.0
Other households	32.7	40.6	34.1	0.0	34.1
Base (n)	43,678	8,935	52,613	0	52,613
Tenure status					
Owned outright	32.7	20.9	30.6	23.4	30.6
Owned on mortgage	40.6	38.4	40.2	39.4	40.2
Rented and others	26.6	40.6	29.1	37.2	29.1
Base (n)	43,151	9,368	52,519	94	52,613
Number of rooms					
1	38.2	44.8	39.4	46.7	39.4
2	36.5	35.7	36.4	33.3	36.4
3	16.7	13.8	16.2	15.6	16.2
4	5.8	3.7	5.5	0.0	5.5
5 or more	2.6	2.0	2.5	4.4	2.5
Base (n)	43,595	8,973	52,568	45	52,613
Wave 12 (2020-22)					
Sex					
Male	47.4	48.8	47.6	0.0	47.6
Female	52.6	51.2	52.4	100.0	52.4
Base (n)	42,111	7,461	49,572	1	49,573
Age					
16-24	39.5	50.3	41.1		41.1
25-44	33.5	30.5	33.1		33.1
45-64	18.8	10.5	17.6		17.6
65+	8.2	8.7	8.2		8.2
Base (n)	42,109	7,464	49,573		49,573
Ethnic background					
White	77.1	61.3	74.8		74.8
Asian	12.2	19.9	13.3		13.3
Black	3.8	8.9	4.5		4.5
Mixed & Other	6.9	9.9	7.4		7.4
Base (n)	42,115	7,458	49,573		49,573
Region (GOR)					
North East	3.5	2.7	3.4	3.6	3.4
North West	9.9	14.6	10.6	11.9	10.6
Yorkshire and the Humber	8.7	11.2	9.0	12.3	9.1
East Midlands	7.2	6.3	7.1	9.6	7.1
West Midlands	8.8	10.4	9.0	7.6	9.0
East of England	8.8	7.1	8.6	7.6	8.5
London	12.6	18.4	13.4	2.0	13.3
South East	12.0	9.7	11.7	6.3	11.6
South West	7.9	5.9	7.6	5.8	7.6
Wales	6.3	6.2	6.3	10.8	6.3

	(a) Observed	(b) IPFW	(a) + (b)	(c) MICE	Complete (a) + (b) + (c)
Scotland	8.1	7.4	8.0	11.9	8.0
Northern Ireland	6.2	0.0	5.3	10.5	5.3
Base (n)	42,096	7,031	49,127	446	49,573
Urbanicity					
Urban	75.2	81.7	76.1	42.9	76.1
Rural	24.8	18.3	23.9	57.1	23.9
Base (n)	42,096	7,470	49,566	7	49,573
Employment status					
Yes	46.2	43.0	45.7	46.4	45.7
No	53.8	57.0	54.3	53.6	54.3
Base (n)	41,941	7,492	49,433	140	49,573
Income					
Up to £1,999	19.9	36.2	23.1		23.1
£2,000-£2,999	20.0	24.5	20.9		20.9
£3,000-£4,999	34.4	27.2	33.0		33.0
£5,000 and over	25.7	12.1	23.1		23.1
Base (n)	40,001	9,572	49,573		49,573
Behind with bills					
Yes	93.5	86.5	92.2	67.3	92.1
No	6.5	13.5	7.8	32.7	7.9
Base (n)	39,751	9,724	49,475	98	49,573
Behind with Council Tax					
Yes	6.0	13.4	7.5	2.0	7.1
No	94.0	86.6	92.5	98.0	92.9
Base (n)	37,214	9,216	46,430	3,143	49,573
Household type					
1 adult, no children	11.7	10.0	11.4	0.0	11.4
1 adult, children	4.1	8.1	4.9	0.0	4.9
Couple, no children	23.4	11.9	21.2	0.0	21.2
Couple with children	27.6	29.9	28.1	0.0	28.1
Other households	33.1	40.1	34.5	0.0	34.5
Base (n)	40,004	9,569	49,573	0	49,573
Tenure status					
Owned outright	34.0	21.8	31.7	20.5	31.6
Owned on mortgage	40.9	36.6	40.0	36.3	40.0
Rented and others	25.1	41.6	28.3	43.2	28.4
Base (n)	39,736	9,691	49,427	146	49,573
Number of rooms					
1	37.3	43.1	38.4	50.0	38.4
2	36.4	36.9	36.5	46.3	36.5
3	17.7	12.8	16.7	3.7	16.7
4	6.0	5.1	5.8	0.0	5.8
5 or more	2.7	2.2	2.6	0.0	2.6
Base (n)	39,913	9,606	49,519	54	49,573

A2 3. References

Austin, P. C., & van Buuren, S. (2023) Logistic regression vs. Predictive mean matching for imputing binary covariates. *Stat. Meth. Med. Res.*, 32, 2172–2183. https://doi.org/10.1177/09622802231198795

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 Papers, 2023–03. https://doi.org/10.5255/UKDA-SN-6614-15

Grund, S., Lüdtke, O., & Robitzsch, A. (2018) Multiple Imputation of Missing Data at Level 2: A Comparison of Fully Conditional and Joint Modeling in Multilevel Designs. *J. Educ. Behav. Stats.,* 43, 316–353. https://doi.org/10.3102/1076998617738087

Kenward, M. G., & Molenberghs, G. (2009) Last Observation Carried Forward: A Crystal Ball? *J. Biopharm. Stat.,* 19, 872–888. https://doi.org/10.1080/10543400903105406

Saha, C., & Jones, M. P. (2009) Bias in the last observation carried forward method under informative dropout. *J. Stat. Plan. Inf.*, 139, 246–255. https://doi.org/10.1016/j.jspi.2008.04.017

Van Buuren, S. (2018). *Flexible imputation of missing data (Second edition)*. CRC Press, Taylor & Francis Group.

Vink, G., Frank, L. E., Pannekoek, J., & van Buuren, S. (2014) Predictive mean matching imputation of semicontinuous variables. *Stat. Neerl.*, 68, 61–90. https://doi.org/10.1111/stan.12023

Appendix A3: Partial unconditional CV (CV_u) derivation

CV_us, which are bounded by the overall CV, quantify univariate associations with response propensity variation (Schouten et al. 2011). Using the same notation as in the main text, the CV_u for covariate Z with *K* categories is

$$\widehat{CV}_u(Z, p_x) = \frac{\sqrt{\frac{1}{n}\sum_{k=1}^{K} n_k\left(\hat{\bar{p}}_k - \hat{\bar{p}}\right)^2}}{\hat{\bar{p}}}, \quad (1)$$

where n_k is the number of observations and $\hat{\bar{p}}_k$ is the mean response propensity in covariate category k. Large values suggest substantial between category variability and non-representativeness associated with Z. Category CV_us decompose and are bounded by covariate CV_us. The CV_u for category *k* of Z is

$$\widehat{CV}_u(Z_k, p_x) = \frac{\sqrt{\frac{n_k}{n}}\left(\hat{\bar{p}}_k - \hat{\bar{p}}\right)}{\hat{\bar{p}}}. \quad (2)$$

Values can be positive or negative, implying respectively over- or under-representation.

In addition, adjustments to correct biases caused by estimating propensities exist, as do approximate standard errors that when converted into 95% confidence intervals (CV ± 1.96 × SE) enable inference regarding (comparative) representativeness or otherwise (de Heij et al. 2015). Population level analysis is also possible by applying sample weights.

A3 3. References

de Heij, V., Schouten, B. and Shlomo, N. (2015) RISQ Manual 2.1: Tools in SAS and R for the computation of R indicators and partial R indicators. Available from: www.risq-project.eu.

Schouten, B., Shlomo, N. and Skinner, C. (2011) Indicators for monitoring and improving representativeness of response. *Journal of Official Statistics*, 27, 231-253.

Appendix A4: Supplementary tables

Table A4 1: Young adult subgroup and (using the same response propensity model) overall dataset web and web plus CATI respondent dataset covariate category CV_us and 95% CIs. * indicates significance.

	Young adult						Overall dataset					
	Web			Combined			Web			Combined		
		95% CIs			95% CIs			95% CIs			95% CIs	
	CV_u	CI -	CI +	CV_u	CI -	CI +	CV_u	CI -	CI +	CV_u	CI -	CI +
Sex: Male	-0.102*	-0.119	-0.084	-0.084*	-0.101	-0.066	-0.102*	-0.119	-0.084	-0.084*	-0.101	-0.066
Sex: Female	0.100*	0.083	0.118	0.082*	0.065	0.100	0.100*	0.083	0.118	0.082*	0.065	0.100
Ethnicity: whte British	0.017*	0.006	0.028	0.004	-0.007	0.016	0.017*	0.006	0.028	0.004	-0.007	0.016
Ethnicity: South Asian	-0.001	-0.025	0.022	0.012	-0.013	0.036	-0.001	-0.025	0.022	0.012	-0.013	0.036
Ethnicity: black	-0.059*	-0.076	-0.043	-0.031*	-0.051	-0.011	-0.059*	-0.076	-0.043	-0.031*	-0.051	-0.011
Ethnicity: other	-0.017	-0.048	0.014	-0.006	-0.038	0.026	-0.017	-0.048	0.014	-0.006	-0.038	0.026
Activity Last Week: In work	0.039*	0.025	0.053	0.030*	0.015	0.044	0.039*	0.025	0.053	0.030*	0.015	0.044
Activity Last Week: Not in work	-0.053*	-0.072	-0.034	-0.040*	-0.060	-0.021	-0.053*	-0.072	-0.034	-0.040*	-0.060	-0.021
Housing tenure: Owner occupied	-0.008	-0.031	0.015	-0.020	-0.043	0.002	-0.008	-0.031	0.015	-0.020	-0.043	0.002
Housing tenure: Mortgage	0.072*	0.053	0.092	0.043*	0.024	0.063	0.072*	0.053	0.092	0.043*	0.024	0.063
Housing tenure: Rented/Other	-0.066*	-0.085	-0.046	-0.030*	-0.049	-0.010	-0.066*	-0.085	-0.046	-0.030*	-0.049	-0.010
Single adult household	0.102*	0.065	0.140	0.103*	0.066	0.141	0.102*	0.065	0.140	0.103*	0.066	0.141
Single adult, kids household	-0.080*	-0.100	-0.061	-0.044*	-0.068	-0.021	-0.080*	-0.100	-0.061	-0.044*	-0.068	-0.021
Couple household	0.123*	0.088	0.158	0.105*	0.071	0.138	0.123*	0.088	0.158	0.105*	0.071	0.138
Couple, kids household	0.035*	0.012	0.058	0.029*	0.007	0.052	0.035*	0.012	0.058	0.029*	0.007	0.052
Other household	-0.086*	-0.103	-0.069	-0.088*	-0.105	-0.070	-0.086*	-0.103	-0.069	-0.088*	-0.105	-0.070
Region: North	-0.018	-0.038	0.003	-0.030*	-0.050	-0.009	-0.018	-0.038	0.003	-0.030*	-0.050	-0.009
Region: East	0.033*	0.010	0.055	0.022	-0.000	0.044	0.033*	0.010	0.055	0.022	-0.000	0.044
Region: South	0.001	-0.022	0.025	0.015	-0.008	0.038	0.001	-0.022	0.025	0.015	-0.008	0.038
Region: West	-0.019	-0.042	0.004	-0.008	-0.031	0.015	-0.019	-0.042	0.004	-0.008	-0.031	0.015
Behind with bills: No	0.037*	0.030	0.044	0.032*	0.024	0.039	0.037*	0.030	0.044	0.032*	0.024	0.039
Behind with bills: Yes	-0.103*	-0.123	-0.083	-0.089*	-0.110	-0.068	-0.103*	-0.123	-0.083	-0.089*	-0.110	-0.068
Household income: 1st quintile	-0.043*	-0.067	-0.018	-0.034*	-0.059	-0.009	-0.043*	-0.067	-0.018	-0.034*	-0.059	-0.009
Household income: 2nd quintile	-0.010	-0.035	0.015	0.002	-0.023	0.027	-0.010	-0.035	0.015	0.002	-0.023	0.027
Household income: 3rd quintile	0.013	-0.011	0.037	0.021	-0.004	0.045	0.013	-0.011	0.037	0.021	-0.004	0.045
Household income: 4th quintile	0.050*	0.026	0.074	0.031*	0.008	0.055	0.050*	0.026	0.074	0.031*	0.008	0.055
Household income: 5th quintile	-0.010	-0.032	0.012	-0.018	-0.040	0.003	-0.010	-0.032	0.012	-0.018	-0.040	0.003

Table A4 2: Older adult subgroup and (using the same response propensity model) overall dataset web and combined (web plus CATI) respondent dataset covariate category CV_us and 95% CIs. * indicates significance.

	Older adult						Overall dataset					
	Web			Combined			Web			Combined		
		95% CIs			95% CIs			95% CIs			95% CIs	
	CV_u	CI -	CI +	CV_u	CI -	CI +	CV_u	CI -	CI +	CV_u	CI -	CI +
Sex: Male	0.092*	0.066	0.119	-0.004	-0.029	0.021	-0.038*	-0.047	-0.030	-0.042*	-0.051	-0.034
Sex: Female	-0.082*	-0.106	-0.058	0.003	-0.019	0.026	0.037*	0.029	0.046	0.041*	0.033	0.049
Ethnicity: whte British	0.013*	0.007	0.018	0.011*	0.005	0.017	0.020*	0.015	0.024	0.015*	0.011	0.020
Ethnicity: South Asian	-0.034*	-0.061	-0.007	-0.037*	-0.062	-0.012	-0.023*	-0.034	-0.012	-0.019*	-0.031	-0.008
Ethnicity: black	-0.050*	-0.069	-0.032	-0.013	-0.041	0.016	-0.060*	-0.068	-0.052	-0.031*	-0.042	-0.020
Ethnicity: other	-0.023	-0.056	0.010	-0.037*	-0.066	-0.008	-0.015	-0.030	-0.000	-0.021*	-0.035	-0.006
Housing tenure: Owner occupied	0.113*	0.098	0.128	0.032*	0.013	0.051	0.034*	0.024	0.044	0.039*	0.029	0.049
Housing tenure: Mortgage	0.059*	0.013	0.106	-0.004	-0.038	0.031	0.063*	0.053	0.073	0.008	-0.002	0.018
Housing tenure: Rented/Other	-0.233*	-0.256	-0.210	-0.057*	-0.092	-0.023	-0.102*	-0.112	-0.092	-0.049*	-0.060	-0.039
Single adult household	-0.177*	-0.201	-0.152	0.021	-0.006	0.048	-0.048*	-0.059	-0.037	0.039*	0.025	0.052
Couple household	0.186*	0.160	0.211	-0.006	-0.031	0.018	0.085*	0.076	0.094	0.040*	0.031	0.049
Other household	-0.033*	-0.066	-0.001	-0.030	-0.063	0.004	-0.057*	-0.066	-0.047	-0.074*	-0.083	-0.065
Region: North	-0.043*	-0.071	-0.015	-0.006	-0.034	0.023	-0.024*	-0.033	-0.014	-0.017*	-0.027	-0.007
Region: East	0.003	-0.027	0.034	-0.001	-0.031	0.029	0.018*	0.007	0.028	0.006	-0.004	0.017
Region: South	0.044*	0.009	0.080	-0.004	-0.036	0.029	0.006	-0.005	0.018	0.003	-0.008	0.015
Region: West	0.010	-0.024	0.043	0.013	-0.019	0.046	0.002	-0.009	0.014	0.011	-0.001	0.022
Behind with bills: No	0.003*	0.001	0.005	0.002	-0.001	0.004	0.016*	0.012	0.019	0.016*	0.013	0.019
Behind with bills: Yes	-0.032*	-0.058	-0.006	-0.020	-0.049	0.008	-0.058*	-0.069	-0.046	-0.059*	-0.070	-0.047
Household income: 1st quintile	-0.218*	-0.242	-0.194	-0.047*	-0.071	-0.022	-0.104*	-0.114	-0.095	-0.028*	-0.040	-0.017
Household income: 2nd quintile	0.080*	0.046	0.114	0.023	-0.008	0.055	0.003	-0.008	0.015	0.010	-0.001	0.022
Household income: 3rd quintile	0.144*	0.099	0.188	0.031	-0.004	0.066	0.033*	0.021	0.046	0.013*	0.001	0.025
Household income: 4th quintile	0.096*	0.048	0.143	0.022	-0.015	0.059	0.052*	0.040	0.065	0.015*	0.003	0.026
Household income: 5th quintile	0.152*	0.098	0.207	0.017	-0.019	0.052	0.029*	0.017	0.041	-0.006	-0.017	0.005

Table A4 3: Ethnic minority subgroup and (using the same response propensity model) overall dataset web and combined (web plus CATI) respondent dataset covariate category CV_us and 95% CIs. * indicates significance.

	Ethnic minority						Overall dataset					
	Web			Combined			Web			Combined		
		95% CIs			95% CIs			95% CIs			95% CIs	
	CV_u	CI -	CI +	CV_u	CI -	CI +	CV_u	CI -	CI +	CV_u	CI -	CI +
Sex: Male	-0.046*	-0.074	-0.018	-0.050*	-0.078	-0.022	-0.038*	-0.047	-0.030	-0.042*	-0.051	-0.034
Sex: Female	0.044*	0.018	0.070	0.047*	0.021	0.074	0.037*	0.029	0.046	0.041*	0.033	0.049
Age: 16-34	-0.046*	-0.076	-0.016	-0.054*	-0.084	-0.024	-0.066*	-0.076	-0.056	-0.088*	-0.098	-0.078
Age: 35-54	0.059*	0.029	0.089	0.048*	0.018	0.078	0.052*	0.041	0.063	0.017*	0.007	0.028
Age: 55-74	0.014	-0.017	0.046	0.018	-0.014	0.049	0.078*	0.068	0.089	0.063*	0.053	0.073
Age: 75+	-0.083*	-0.104	-0.062	-0.026	-0.058	0.005	-0.105*	-0.114	-0.096	0.011	-0.002	0.024
Activity Last Week: In work	0.048*	0.025	0.070	0.012	-0.011	0.036	0.045*	0.038	0.053	-0.004	-0.012	0.004
Activity Last Week: Not in work	-0.059*	-0.088	-0.031	-0.016	-0.045	0.014	-0.051*	-0.060	-0.043	0.005	-0.004	0.014
Housing tenure: Owner occupied	0.048*	0.016	0.081	0.028	-0.003	0.059	0.034*	0.024	0.044	0.039*	0.029	0.049
Housing tenure: Mortgage	0.069*	0.038	0.100	0.027	-0.004	0.057	0.063*	0.053	0.073	0.008	-0.002	0.018
Housing tenure: Rented/Other	-0.093*	-0.122	-0.065	-0.042*	-0.071	-0.013	-0.102*	-0.112	-0.092	-0.049*	-0.060	-0.039
Single adult household	0.040	-0.011	0.091	0.088*	0.032	0.143	-0.034*	-0.046	-0.022	0.054*	0.040	0.068
Single adult, kids household	-0.066*	-0.097	-0.035	-0.014	-0.053	0.025	-0.042*	-0.054	-0.031	-0.028*	-0.041	-0.015
Couple household	0.089*	0.043	0.135	0.041*	0.001	0.081	0.087*	0.076	0.099	0.057*	0.046	0.068
Couple, kids household	0.026	-0.008	0.060	-0.000	-0.033	0.032	0.027*	0.015	0.039	-0.006	-0.017	0.006
Other household	-0.062*	-0.090	-0.035	-0.060*	-0.087	-0.033	-0.057*	-0.066	-0.047	-0.074*	-0.083	-0.065
Region: North	-0.010	-0.048	0.028	-0.013	-0.052	0.025	-0.024*	-0.033	-0.014	-0.017*	-0.027	-0.007
Region: East	0.023	-0.009	0.055	-0.001	-0.032	0.030	0.018*	0.007	0.028	0.006	-0.004	0.017
Region: South	-0.030*	-0.056	-0.003	0.001	-0.025	0.028	0.006	-0.005	0.018	0.003	-0.008	0.015
Region: West	0.062*	0.014	0.109	0.021	-0.020	0.063	0.002	-0.009	0.014	0.011	-0.001	0.022
Behind with bills: No	0.032*	0.020	0.044	0.031*	0.019	0.043	0.016*	0.012	0.019	0.016*	0.013	0.019
Behind with bills: Yes	-0.081*	-0.111	-0.051	-0.079*	-0.109	-0.049	-0.058*	-0.069	-0.046	-0.059*	-0.070	-0.047
Household income: 1st quintile	-0.098*	-0.130	-0.067	-0.039*	-0.074	-0.003	-0.104*	-0.114	-0.095	-0.028*	-0.040	-0.017
Household income: 2nd quintile	0.011	-0.027	0.049	-0.007	-0.045	0.030	0.003	-0.008	0.015	0.010	-0.001	0.022
Household income: 3rd quintile	0.037	-0.001	0.075	0.041*	0.003	0.080	0.033*	0.021	0.046	0.013*	0.001	0.025
Household income: 4th quintile	0.046*	0.010	0.082	0.020	-0.014	0.055	0.052*	0.040	0.065	0.015*	0.003	0.026
Household income: 5th quintile	0.013	-0.021	0.047	-0.011	-0.043	0.022	0.029*	0.017	0.041	-0.006	-0.017	0.005

Table A4 4: Low income subgroup and (using the same response propensity model) overall dataset web and combined (web plus CATI) respondent dataset covariate category CV$_u$s and 95% CIs. * indicates significance.

	Low income						Overall dataset					
	Web			Combined			Web			Combined		
		95% CIs			95% CIs			95% CIs			95% CIs	
	CV$_u$	CI -	CI +	CV$_u$	CI -	CI +	CV$_u$	CI -	CI +	CV$_u$	CI -	CI +
Sex: Male	-0.026*	-0.046	-0.006	-0.035*	-0.055	-0.015	-0.038*	-0.047	-0.030	-0.042*	-0.051	-0.034
Sex: Female	0.023*	0.005	0.040	0.030*	0.013	0.047	0.037*	0.029	0.046	0.041*	0.033	0.049
Age: 16-34	0.001	-0.025	0.028	-0.086*	-0.109	-0.064	-0.066*	-0.076	-0.056	-0.088*	-0.098	-0.078
Age: 35-54	0.053*	0.025	0.082	-0.014	-0.040	0.011	0.052*	0.041	0.063	0.017*	0.007	0.028
Age: 55-74	0.124*	0.101	0.147	0.075*	0.053	0.096	0.078*	0.068	0.089	0.063*	0.053	0.073
Age: 75+	-0.193*	-0.211	-0.176	0.011	-0.014	0.036	-0.105*	-0.114	-0.096	0.011	-0.002	0.024
Ethnicity: White British	0.017*	0.007	0.027	0.017*	0.007	0.027	0.020*	0.015	0.024	0.015*	0.011	0.020
Ethnicity: South Asian	-0.026*	-0.046	-0.006	-0.029*	-0.049	-0.009	-0.023*	-0.034	-0.012	-0.019*	-0.031	-0.008
Ethnicity: Black	-0.057*	-0.078	-0.036	-0.022	-0.048	0.005	-0.060*	-0.068	-0.052	-0.031*	-0.042	-0.020
Ethnicity: Other	-0.007	-0.042	0.028	-0.027	-0.060	0.006	-0.015	-0.030	-0.000	-0.021*	-0.035	-0.006
Activity Last Week: In work	0.126*	0.101	0.150	-0.011	-0.032	0.011	0.045*	0.038	0.053	-0.004	-0.012	0.004
Activity Last Week: Not in work	-0.091*	-0.109	-0.073	0.008	-0.008	0.023	-0.051*	-0.060	-0.043	0.005	-0.004	0.014
Housing tenure: Owner occupied	0.053*	0.032	0.074	0.066*	0.045	0.087	0.034*	0.024	0.044	0.039*	0.029	0.049
Housing tenure: Mortgage	0.112*	0.080	0.143	0.009	-0.016	0.035	0.063*	0.053	0.073	0.008	-0.002	0.018
Housing tenure: Rented/Other	-0.109*	-0.128	-0.089	-0.064*	-0.083	-0.045	-0.102*	-0.112	-0.092	-0.049*	-0.060	-0.039
Single adult household	-0.003	-0.021	0.016	0.094*	0.076	0.111	-0.034*	-0.046	-0.022	0.054*	0.040	0.068
Single adult, kids household	-0.069*	-0.091	-0.047	-0.070*	-0.092	-0.049	-0.042*	-0.054	-0.031	-0.028*	-0.041	-0.015
Couple household	0.078*	0.052	0.104	0.007	-0.016	0.029	0.087*	0.076	0.099	0.057*	0.046	0.068
Couple, kids household	-0.017	-0.042	0.008	-0.071*	-0.091	-0.052	0.027*	0.015	0.039	-0.006	-0.017	0.006
Other household	-0.027*	-0.052	-0.002	-0.086*	-0.107	-0.065	-0.057*	-0.066	-0.047	-0.074*	-0.083	-0.065
Region: North	-0.040*	-0.061	-0.019	-0.014	-0.035	0.007	-0.024*	-0.033	-0.014	-0.017*	-0.027	-0.007
Region: East	0.042*	0.017	0.066	-0.004	-0.027	0.019	0.018*	0.007	0.028	0.006	-0.004	0.017
Region: South	-0.006	-0.033	0.021	-0.000	-0.027	0.026	0.006	-0.005	0.018	0.003	-0.008	0.015
Region: West	0.012	-0.014	0.038	0.026	-0.000	0.051	0.002	-0.009	0.014	0.011	-0.001	0.022
Behind with bills: No	0.015*	0.007	0.024	0.028*	0.021	0.036	0.016*	0.012	0.019	0.016*	0.013	0.019
Behind with bills: Yes	-0.046*	-0.072	-0.021	-0.086*	-0.108	-0.064	-0.058*	-0.069	-0.046	-0.059*	-0.070	-0.047

Table A4 5: Overall and subgroup dataset web and combined (web plus CATI) respondent non-response weighted dataset biases in survey measured characteristics compared to issued sample weighted benchmark equivalents. 'Bch.' is the benchmark estimate (standard error in brackets). 'Diff' is the difference between the non-response weighted estimate and the benchmark estimate. * equals *P*<0.05. Missing benchmark / test values indicate characteristics not included among those utilised in that web / combined dataset comparison. Note that with the Older adult subgroup, Single adult households include those with or without kids, and Couple households include those with or without kids.

	Overall			Young adult			Older adult			Ethnic minority			Low income		
		Web	Combined		Web	Combined		Web	Combined		Web	Combined		Web	Combined
Variable	Bch. (se)	Diff	Diff	Bch. (se)	Diff	Diff	Bch. (se)	Diff	Diff	Bch. (se)	Diff	Diff	Bch. (se)	Diff	Diff
Sex: Male	0.485 (0.003)	-0.002	0.001	0.492 (0.005)	0.033*	0.013	0.442 (0.009)	-0.094*	-0.025	0.474 (0.006)	-0.003	-0.005	0.427 (0.006)	-0.025	-0.015
Age: 16-34	0.281 (0.003)	-0.000	0.002							0.379 (0.006)	-0.031	-0.027	0.220 (0.005)	-0.040*	-0.014
Age: 35-54	0.324 (0.003)	-0.001	-0.000							0.429 (0.006)	0.013	0.005	0.223 (0.005)	-0.005	0.000
Age: 55-74	0.277 (0.003)	-0.001	-0.001							0.155 (0.004)	0.014	0.015	0.314 (0.006)	0.008	0.006
Age: 75+	0.118 (0.002)	0.002	-0.001							0.038 (0.002)	0.004	0.007	0.243 (0.005)	0.036*	0.009
Ethnicity: White British	0.864 (0.002)	0.002	0.001	0.818 (0.004)	0.017	0.016	0.956 (0.004)	-0.004	-0.007				0.874 (0.004)	0.007	0.004
Ethnicity: Black	0.054 (0.001)	-0.001	-0.001	0.075 (0.003)	-0.010	-0.008	0.016 (0.002)	0.002	0.003				0.037 (0.002)	0.002	0.001
Ethnicity: South Asian	0.023 (0.001)	-0.001	-0.000	0.030 (0.002)	-0.007	-0.004	0.010 (0.002)	0.001	0.000				0.027 (0.002)	-0.008	-0.005
Ethnicity: Other	0.058 (0.001)	0.000	-0.000	0.077 (0.003)	0.000	-0.004	0.018 (0.002)	0.001	0.005				0.061 (0.003)	-0.001	0.001
Activity Last Week: In work	0.564 (0.003)	-0.001	0.002	0.648 (0.005)	0.003	0.000				0.608 (0.006)	0.005	0.006	0.344 (0.006)	-0.040*	-0.004
Housing tenure: Owner occupied	0.342 (0.003)	-0.001	-0.000	0.169 (0.004)	0.018*	0.010	0.747 (0.008)	-0.046	-0.017	0.169 (0.004)	0.003	-0.001	0.376 (0.006)	0.004	-0.008
Housing tenure: Mortgage	0.346 (0.003)	-0.001	0.001	0.406 (0.005)	0.011	0.004	0.035 (0.003)	-0.016*	-0.001	0.382 (0.006)	0.014	0.009	0.146 (0.004)	-0.019*	-0.008
Housing tenure: Rented/Other	0.311 (0.003)	0.002	-0.001	0.424 (0.005)	-0.028	-0.013	0.218 (0.007)	0.062*	0.018	0.449 (0.006)	-0.017	-0.008	0.478 (0.006)	0.015	0.016
Single adult household	0.178 (0.002)	0.005	0.000	0.109 (0.003)	-0.027*	-0.006	0.432 (0.009)	0.093*	0.015	0.126 (0.004)	-0.013	-0.005	0.351 (0.006)	0.000	-0.006
Single adult, kids household	0.032 (0.001)	0.001	0.001	0.056 (0.002)	0.007	0.005				0.038 (0.002)	0.005	-0.001	0.264 (0.005)	-0.006	0.007
Couple household	0.258 (0.002)	-0.003	-0.001	0.091 (0.003)	-0.005	-0.010*	0.465 (0.009)	-0.066*	-0.003	0.108 (0.004)	0.003	0.006	0.217 (0.005)	0.005	0.000
Couple, kids household	0.175 (0.002)	-0.000	-0.000	0.160 (0.004)	-0.005	-0.008				0.247 (0.005)	0.017	0.014	0.169 (0.005)	0.000	-0.001
Other household	0.356	-0.003	-0.000	0.584	0.031	0.019	0.103	-0.027	-0.012	0.482	-0.013	-0.013	0.516	0.045*	0.021*

	(0.003)			(0.005)			(0.005)			(0.006)			(0.006)		
Region: North	0.305	-0.002	-0.001	0.296	-0.005	0.002	0.367	-0.006	-0.014	0.195	-0.011	-0.010	0.064	0.006	0.005
	(0.003)			(0.005)			(0.009)			(0.005)			(0.003)		
Region: South	0.258	0.000	0.000	0.251	-0.006	-0.008	0.248	0.011	0.006	0.220	0.001	0.004	0.196	-0.008	0.001
	(0.002)			(0.005)			(0.008)			(0.005)			(0.005)		
Region: East	0.273	-0.000	-0.000	0.299	0.000	-0.003	0.212	-0.003	0.008	0.523	0.016	0.005	0.069	0.000	0.003
	(0.003)			(0.005)			(0.007)			(0.006)			(0.003)		
Region: West	0.164	0.002	0.001	0.154	0.011	0.009	0.172	-0.002	-0.001	0.062	-0.005	0.002	0.155	-0.043*	-0.030*
	(0.002)			(0.004)			(0.007)			(0.003)			(0.004)		
Household Location: Urban	0.756	0.001	-0.001	0.808	-0.013	-0.011	0.701	0.041	0.010	0.922	0.013	0.003	0.774	0.011	0.002
	(0.002)			(0.004)			(0.008)			(0.003)			(0.005)		
Behind with bills: No	0.932	0.000	-0.000	0.886	-0.008	-0.003	0.992	-0.002	-0.001	0.863	0.002	0.002	0.903	0.006	-0.002
	(0.001)			(0.003)			(0.002)			(0.004)			(0.004)		
Behind Council Tax: Yes	0.068	-0.002	-0.000	0.114	0.003	0.004	0.008	-0.005	-0.002	0.127	-0.027	-0.005	0.099	-0.026*	-0.007
	(0.001)			(0.003)			(0.002)			(0.004)			(0.004)		
Household income: 1st quintile	0.239	-0.000	-0.004	0.187	-0.034*	-0.017	0.492	0.067*	0.013	0.221	-0.010	-0.008			
	(0.002)			(0.004)			(0.009)			(0.005)					
Household income: 2nd quintile	0.204	0.000	0.001	0.188	0.001	0.003	0.256	-0.007	-0.001	0.185	-0.002	0.006			
	(0.002)			(0.004)			(0.008)			(0.005)					
Household income: 3rd quintile	0.186	-0.000	0.000	0.188	0.007	-0.001	0.134	-0.027*	-0.006	0.178	-0.004	-0.012			
	(0.002)			(0.004)			(0.006)			(0.004)					
Household income: 4th quintile	0.179	-0.000	0.000	0.196	0.006	0.001	0.061	-0.010	-0.002	0.185	0.008	0.003			
	(0.002)			(0.004)			(0.004)			(0.005)					
Household income: 5th quintile	0.192	0.000	0.002	0.241	0.019*	0.013	0.057	-0.023*	-0.003	0.229	0.008	0.010			
	(0.002)			(0.005)			(0.004)			(0.005)					

www.surveyfutures.net