methods, data, analyses | 2026, pp. 1-30 DOI: 10.25521/mda.584

Survey Mode Change by Necessity: Evaluating Survey Dataset Mode and Frequency Change During a Global Pandemic

Nicole D. James[1] and Jamie C. Moore[1]

[1] *Institute for Social and Economic Research, University of Essex*

Abstract

The COVID-19 pandemic significantly affected how surveys were administered. One change was suspending face-to-face (F2F) interviewing, the primary mode for many surveys, and shifting to mainly web-based designs. Another was more frequent data collection. The former mode shift had already been occurring in surveys pre-pandemic, but at a slower pace and not to the same extent. Hence, these design changes were made with limited information on their impact on dataset performance. This paper investigates the extent to which information on three components of performance (response rates, non-response bias risks, and variation in sample subgroup non-response) in pre-pandemic surveys predicted similar outcomes in pandemic surveys using data from two *Understanding Society: The UK Household Longitudinal Study* (UKHLS) surveys. The UKHLS main survey is an annual panel survey with a mixed mode (including F2F) design. The UKHLS COVID-19 Study was a high-frequency (mostly monthly or bimonthly) pandemic-era survey in which main survey participants were invited to complete web questionnaires. Both longitudinal (respondents to the current wave and all prior waves) and cross-sectional (all respondents) datasets are studied. All three performance components differ between the two surveys, implying that main survey information on them was of limited predictive value in the COVID-19 Study. In particular, using a bias minimization strategy based on main survey information in the COVID-19 Study would have inflated costs and potentially negatively impacted dataset performance. The implications of these findings, both for surveys similar to the COVID-19 Study and surveys more generally, are then discussed.

Keywords: survey modes, representativeness, coefficients of variation, dataset performance, COVID-19



© The Author(s) 2026. This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 License. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

General population panel surveys aim to collect data from a sample that reflects a population over time. However, a challenge is that response rates are declining (de Leeuw & de Heer, 2002; Luiten et al., 2020). Non-response is important because it reduces dataset size and survey estimate precision (Groves et al., 2009). In addition, if it differs across sample subgroups defined by characteristics correlated with survey variables, survey variable estimates can deviate from population values (non-response biases), causing invalid inference.

There are various ways in which surveys can seek to minimize non-response biases. Bias prevention measures, such as offering incentives to under-represented subgroups, are used before or during fieldwork to reduce differential non-response (e.g., Groves, 2006; Groves et al., 2009; Peytchev et al., 2010). Bias adjustment measures, such as supplying non-response weights (e.g., Valliant et al., 2018; see Little & Rubin, 2020, for other methods), are used after data collection to reduce remaining biases. It should also be noted that bias prevention measure success improves bias adjustment effectiveness (e.g., Moore et al., 2024; Schouten et al., 2016).

Survey Mode and Non-Response Biases

The mode by which interviews are conducted can be an important determinant of non-response biases. Interviews may be face-to-face (F2F) or by telephone, paper, or web. Previously, F2F has mostly been used in general population surveys in the UK. It has higher response rates than other modes but also higher costs. Recently, less costly web mode has become increasingly used (Baker et al., 2010; Couper et al., 2007; Lugtig et al., 2014; Schonlau & Couper, 2017; Schonlau et al., 2009). However, response rates tend to be lower and non-response biases higher than with F2F (Cornesse & Bosnjak, 2018; Daikeler et al., 2020; Nicolaas et al., 2014). Hence, many surveys implement sequential mixed-mode designs in which web non-respondents are followed up F2F or by telephone to increase response and reduce biases (e.g., Bianchi et al., 2017; Burton & Jäckle, 2020; Cornesse & Bosnjak, 2018; Couper et al., 2007; Nicolaas et al., 2014).

The COVID-19 Pandemic and Non-Response Biases

The impact of mode on non-response biases was especially important during the COVID-19 pandemic. Many countries enacted restrictions to slow virus spread that significantly impacted surveys. F2F mode was often not possible, with

Direct correspondence to
Nicole D. James, Institute for Social and Economic Research, University of Essex, Wivenhoe Park, Colchester, CO4 3SQ, United Kingdom
E-mail: njames@essex.ac.uk

mostly web or telephone used instead. This caused difficulties regarding predicting biases and differential subgroup non-response that any bias minimization efforts would seek to address. As noted in the Survey mode and non-response biases section, similar mode changes were occurring before the pandemic, but at a slower pace and not to the extent now necessary, so it was not known if pre-pandemic information was of use for such a task.

In addition, there was a need for timely information on pandemic impacts. This was often not possible with existing surveys, so new, more frequently fielded surveys, often with eligible samples drawn from existing surveys, were initiated (e.g., Blom et al., 2020; Brown et al., 2021; Burton et al., 2020; Gummer et al., 2020). These surveys were subject to questions regarding the utility of pre-pandemic information from parent surveys for predicting non-response biases and differential subgroup non-response in the same way that existing surveys were, with the addition that their increased frequencies might also impact performance.

Quantifying Non-Response Biases

Quantifying non-response biases is often difficult because non-respondent information does not exist, and population values are not available (Hand, 2018). Hence, in early work, the response rate was often used as an indirect indicator of risks of biases, with high rates presumed to indicate lower risks (Schouten et al., 2011, 2012). However, empirical work has tended to find low correlations between response rates and biases (Groves, 2006; Groves & Peytcheva, 2008; Schouten et al., 2009). Therefore, alternative indicators have been sought.

One set of indicators that are useful in this regard are representativeness indicators (Schouten et al., 2012; see also the Evaluation Approach section). These evaluate how well respondents reflect the eligible sample by quantifying variation in sample member response propensities estimated using auxiliary information available for all members. Low propensity variation suggests that respondents are a random subset of the sample with respect to auxiliary covariates (dataset representativeness) and that bias risks are low. In addition, partial indicators exist that decompose propensity variation (i.e., dataset non-representativeness) into that correlated with auxiliary covariates and categories. In the case of covariate level indicators, this enables identification of covariates for inclusion in models underlying bias adjustments (Moore et al., 2021). In the case of category level indicators, it enables identification of under-represented sample subgroups that should be targeted by bias prevention measures (Schouten et al., 2012). Statistical inference concerning indicators is also possible (for alternative indicators, see Little & Vartivarian, 2005; Nagelkerke, 1991; Särndal, 2011; Wagner, 2010).

Research Questions

Little work exists on how mode and frequency changes in pandemic surveys impacted response rates, non-response bias risks, and differential sample subgroup non-response or whether impacts differed from those in the pre-pandemic surveys, that is, whether knowledge of the latter was useful to predict the former. Concerning differential subgroup non-response, this includes both subgroup identity (so that they can be targeted by bias prevention measures) and the identity of defining auxiliary covariates (so that they can be incorporated in bias adjustment measures); together, we term these correlates of response propensity variation and, combined with response rates and bias risks, components of dataset performance. These questions are of broad relevance. The pandemic was unprecedented in recent times, but similar could happen again, with comparable effects on surveys. In addition, even without such events, web may completely replace F2F in future surveys, so information on impacts on the aforementioned quantities when F2F is not used is of value. This paper addresses these knowledge gaps by evaluating and comparing components of dataset performance in *Understanding Society: The UK Household Longitudinal Study* (UKHLS) (Institute for Social and Economic Research, 2025) and the UKHLS COVID-19 Study, a higher-frequency, pandemic-era survey with an eligible sample drawn from participants of the former survey (Institute for Social and Economic Research, 2021). The UKHLS used F2F, web, and telephone modes. The COVID-19 Study was restricted to the latter two modes. Three research questions are investigated:

RQ1: Do response rates in the COVID-19 Study differ from those in the main survey?

RQ2: Do non-response bias risks in the COVID-19 Study differ from those in the main survey?

RQ3: Do the auxiliary covariate correlates of response propensity variation in the COVID-19 Study differ from those in the main survey?

Representativeness indicators are used to evaluate bias risks and correlates of response propensity variation. Both longitudinal (only respondents to the wave in question and all previous waves) and cross-sectional (all respondents to the wave) datasets are considered, as both are used by researchers. Eight waves from each survey are studied. The paper proceeds as follows. In Data and Methods, the two surveys, the evaluation datasets, representativeness indicators, and the evaluation approach are described. The main findings are reported in the Results section. Finally, the implications of these findings for survey design are examined in the Discussion.

Data and Methods

Data

UKHLS Main Survey

Understanding Society: The UK Household Longitudinal Study (UKHLS), hereafter termed "the main survey," is a panel survey with a target population of adults aged 16+ in households in the UK. The eligible sample is constructed from probability samples, with non-response carefully modeled (Lynn, 2009; Lynn & Kaminska, 2010). Interviews are conducted annually. The first person in a sample household to begin the questionnaire is asked to complete a household grid identifying all household members and a household questionnaire including items about the household. Then, all adults in the household are asked to complete an individual questionnaire. The survey began in 2009, with F2F interviews. At wave 3, telephone mode was introduced, and at wave 7, web mode. By the end of 2025, 15 waves had been achieved. It currently has a sequential mixed-mode design: some participants are allocated to web (a percentage that has grown over time to around 80%) and others to F2F, with follow-up in other modes. The current incentive protocol for adult sample members is offered on an individual basis and is mostly dependent on whether the sample member has participated in the previous wave. Those who participated in the last wave or turned 16 in the current wave receive an unconditional £20 gift voucher, whereas those who did not participate in the last wave or who are new entrants receive a £20 conditional gift voucher on completion of the survey. Moreover, those in web-first households that complete the survey within five weeks receive an additional £10 gift voucher. See Institute for Social and Economic Research (2025) for more details.

UKHLS COVID-19 Study

In April 2020 the UKHLS team fielded the UKHLS COVID-19 Study (hereafter termed "the COVID-19 Study"), a primarily web-based survey with a target population of adults in the UK that focused on the impacts of the pandemic and government policy responses on individuals and families in the UK (Burton et al., 2020). The survey's eligible sample was those in HHs responding at main survey waves 8 or 9 who had not attrited, died, or emigrated before the study began. The survey consisted of an individual interview, with no collection of household information, that is, unlike the main survey, it was not a household panel survey. The COVID-19 Study was also fielded more frequently than the main survey, mostly at monthly or bimonthly intervals. Nine waves were fielded, the last in September 2021. At waves 1 and 6, some web non-respondents were followed up by telephone. These interviews are not considered here. Sample members received a conditional £2 incentive for each survey completed that could be exchanged for gift cards, electronic vouchers, or donated to charity. In the final

wave, sample members were offered a conditional £10 incentive for completing the survey. Serology testing for COVID-19 antibodies was offered at wave 8 with greater financial incentives at waves 8 and 9. See the Institute for Social and Economic Research (2021) for more details of the study.

Evaluation Datasets

Evaluation datasets are constrained by the need for auxiliary covariate information on both survey respondents and non-respondents so that representativeness indicators could be computed to address research questions 2 and 3. Hence, given that such information does not exist for main survey wave 1 non-respondents, the evaluation-eligible sample for the main survey is defined as all main survey wave 1 respondents with complete wave 1 information for the eight auxiliary covariates included in the indicator response propensity models (= 46,885 individuals; see Table 1 for the auxiliary covariates and Table 2 for a summary of evaluation-eligible sample selection for the two surveys). For the cross-sectional datasets, the first respondents (all respondents to the wave) considered are from wave 2. Then, given the same evaluation-eligible sample, respondents from each wave up to and including wave 9, the last pre-pandemic wave, are considered. For the longitudinal datasets, the first respondents (respondents to the wave in question and all previous waves) considered are the wave 3 respondents. Then, as with the cross-sectional datasets, respondents are considered for each wave up to wave 9.

Table 1

Auxiliary Covariates Used in Evaluations and Their Categories

Auxiliary Covariate	Categories
Sex (sex_dv)	1) Male; 2) Female
Ethnicity (ethn_dv)	1) Not ethnic minority; 2) Ethnic minority
Age (age_dv)	1) 16-19; 2) 20-24; 3) 25-34; 4) 35-44; 4) 45-54; 5) 55-64; 6) 65+
Employment Status (jbstat)	1) Employed, 2) in education or training; 3) Unemployed; 4) Retired; 5) Other
Household Composition (hhtype_dv)	1) 1 adult, no children; 2) 1 adult, children; 3) Couple, no children; 4) Couple, children; 5) 2+ adults, no children; 6) 2+ adults, children
Marital Status (marstat)	1) Single; 2) Married; 3) Separated/divorced; 4) Widowed
Highest Qualification (hiqual_dv)	1) Degree or equiv.; 2) A-Level or equiv.; 3) GCSE or equiv.; 4) Other; 5) None
Tenure (tenure_dv)	1) Owned; 2) Rented; 3) Other

Note. UKHLS variable names in brackets.

For comparability, a similar approach is used for the COVID-19 Study evaluation datasets even though main survey wave 9 information exists for many survey-eligible sample members. The evaluation-eligible sample is defined as study wave 1 respondents with complete information for the eight auxiliary covariates included in the indicator response propensity models (= 14,777 individuals). For the cross-sectional datasets, the first respondents considered are from wave 2; then, given the same evaluation-eligible sample, waves 3 to 9 respondents are considered. For the longitudinal datasets, the first respondents considered are from wave 3; then, given the same evaluation-eligible sample, waves 4 to 9 respondents are considered. The UKHLS datasets are available from the UK Data Service (University of Essex & Institute for Social and Economic Research, 2021, 2023).

Table 2

Overview of the UKHLS Main Survey and COVID-19 Study Samples and Evaluation Datasets

	UKHLS Main Survey	UKHLS COVID-19 Survey
Survey type	Household panel	Individual panel
Target population	UK adult (aged 16+) population	UK adult (aged 16+) population
Survey frequency	Annual Wave 1: Jan 2009–Mar 2011 Waves 2-9: Jan 2010–May 2019	Waves 1-4 monthly, waves 5-8 bimonthly, wave 9 6 months later Wave 1: Apr 2020 Waves 2-9: May 2020–Sep 2021
Survey interview mode	Web, face to face and telephone	Web and telephone (latter not considered in this paper)
Survey eligible sample	Adults living in sampled HHs	Adults living in HHs in which someone responded to main survey waves 8 or 9
Evaluation eligible sample	Wave 1 respondents with complete auxiliary covariate information (design and non-response weighted) = 46,885 individuals	Wave 1 respondents with complete auxiliary covariate information (design and non-response-weighted) = 14,777 individuals
Evaluated waves	2-9 (cross-sectional) 3-9 (longitudinal)	2-9 (cross-sectional) 3-9 (longitudinal)

Note. The table presents the samples and datasets evaluated in this paper.

Evaluation Approach

RQ1: Differences in Response Rates

To evaluate differences in response rates between the two surveys, for each evaluation dataset described in the Evaluation datasets section, the response rate is computed as (number of respondents / evaluation-eligible sample size) * 100.

RQ2: Differences in Non-Response Bias Risks

To evaluate differences in non-response bias risks between the two surveys, representativeness indicators are computed. These indicators quantify variation in response propensities estimated given the parameter estimates from a logistic regression model including auxiliary covariates available for all sample members (Schouten et al., 2012). Weights can be applied in regression models to map sample members to the study population. If response is correlated with the auxiliary covariates and such covariates correlated with survey variables, low variation implies high respondent representativeness compared to the (weighted) eligible sample and low non-response bias risk (see Nishimura et al., 2017; Schouten et al., 2016 for empirical support). Specifically, Coefficients of Variation of the response propensities (CVs) are used. Overall CVs quantify overall dataset representativeness. The overall CV is, for sample size n and auxiliary covariate set $\boldsymbol{x}$ producing the propensity vector p_x

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

where $\hat{p}_i$ is the response propensity of subject i and $\hat{\bar{p}}$ average response propensity. The less propensities differ, the smaller the overall CV and the greater overall respondent representativeness. There are two reasons why CVs are used rather than R-indicators (see Quantifying non-response biases section) in evaluations. First, they quantify comparative representativeness more accurately when, as occurs in this paper (see Results), the response rates of compared datasets differ (see Moore et al., 2018; Schouten et al., 2009 for details). Second, overall CVs provide a measure of non-response risk on a scale meaningful to survey designers: assuming that non-response is correlated with the auxiliary covariates and that such covariates are correlated with survey variables, they approximate the maximal absolute standardized non-response bias of estimated sample-based survey variable means (see Schouten et al., 2009). Sampling bias adjustments to CVs to account for bias caused by estimating response propensities and CV standard errors that can be converted into 95% confidence intervals (CIs) to enable statistical inference are also available (for details, see de Heij et al., 2015; Schouten et al., 2011; Shlomo et al., 2012).

Sample bias-adjusted overall CVs and their 95% CIs are computed for each evaluation dataset. When CVs are used for comparative purposes, auxiliary covariates should be the same (Schouten et al., 2009). Hence, eight auxiliary covariates collected from wave 1 respondents in both surveys are used to estimate response propensities (see Table 1 for covariates). These are all known correlates of response propensities in the UKHLS or similar surveys (e.g., Lugtig et al., 2014; Lynn & Borkowska, 2018; Moore et al., 2018; Uhrig, 2008). The evaluation-eligible samples are also weighted. The wave 1 non-response weights supplied with the two surveys are included in the regression models used to estimate response propensities. For the main survey, the weight is the survey design weight adjusted for wave 1 non-response and post-stratified to 2011 UK population estimates. For the COVID-19 Study, it is the main survey wave 9 cross-sectional non-response weight (the design weight adjusted for waves 1-9 non-response and post-stratified to 2019 UK population estimates) adjusted for COVID-19 Study wave 1 non-response (Institute for Social and Economic Research 2021, 2025). These weights correct for unequal selection probabilities into the surveys and map the evaluation-eligible samples to the UK population at different time points: June 2011 for the main survey and June 2019 for the COVID-19 Study. Empirical work has shown that main survey weighted survey estimates perform well in approximating relevant population benchmarks (Benzeval et al., 2020; Borkowska, 2019) and that COVID-19 Study wave 1 weighted estimates approximate similar main survey wave 9 weighted estimates (Moore et al., 2024). Hence, overall CVs quantify non-response bias risks compared to the UK population at these time points.

There was no alternative to using weights mapping evaluation samples to the UK population at different times given the disparity in time periods covered by the two sets of datasets. It should also be noted that although the weights account for the sociodemographic aspects of selection into the evaluation-eligible samples for the two surveys, most of the COVID-19 sample will have previously responded to the main survey and may be positively disposed towards the survey. However, even given this, the evaluations still address the question of whether pre-pandemic main survey non-response bias risks are useful to predict those in the COVID-19 Study. Note as well that potentially correlated responses among members of the same household are not accounted for in evaluations. Information on household membership is only available for the main survey, which, given the need for the same response propensity model to be used (see the previous paragraph), causes issues when comparing surveys. In addition, methods to incorporate such information in CV computation do not exist and developing them is beyond the scope of this paper. The R code of de Heij et al. (2015) is used to compute indicators.

RQ3: Differences in the Correlates of Response Propensity Variation

To evaluate differences in the correlates of response propensity variation between surveys, partial covariate and covariate category CVs are computed. These indicators, which are fully described in the Appendix, quantify response propensity variation associated with propensity model auxiliary covariates and their categories. Two forms exist. Unconditional CVs (CV_us) quantify univariate associations with propensity variation, with zero values indicating no association. For covariate CV_us, larger positive values indicate greater non-representativeness associated with the covariate. For category CV_us, positive values indicate over-representation and negative values under-representation, with values of greater magnitude indicating greater associated non-representativeness. Conditional CVs (CV_cs) quantify multivariate associations, that is, propensity variation associated with the covariate after the impacts of other covariates have been accounted for. For both covariate and category CV_cs, larger positive values indicate greater associated, solely attributable, non-representativeness. A large CV_u coupled with a small CV_c equivalent implies an impact that is also correlated with other auxiliary covariates.

Given this, partial CVs are useful when designing bias minimization measures. CV_us identify covariates to include in bias adjustment models and under-represented subgroups to target with bias prevention measures. CV_cs can ensure measures are efficient, that is, that covariates with correlated impacts are not (both) included in bias adjustment models (model overfitting causes precision loss: see Little & Vartivarian, 2005), and under-represented groups with correlated impacts are not (both) targeted by bias prevention measures (Schouten et al., 2012). For each evaluation dataset, sample bias-adjusted partial CVs (for details, see de Heij et al., 2015; Shlomo & Schouten, 2013) are computed for the auxiliary covariates included in response propensity models and their categories (see Table 1), and also their standard errors (see de Heij et al., 2015; Shlomo et al., 2013) for conversion into 95% CIs for statistical inference. The R code of de Heij et al. (2015) is used to compute CVs. The code for all evaluations in the paper is available from the authors.

Results

Table 3 summarizes all evaluation results. It should be noted that while the results of the partial conditional CV analyses are described below, relevant tables and figures are presented in the Appendix.

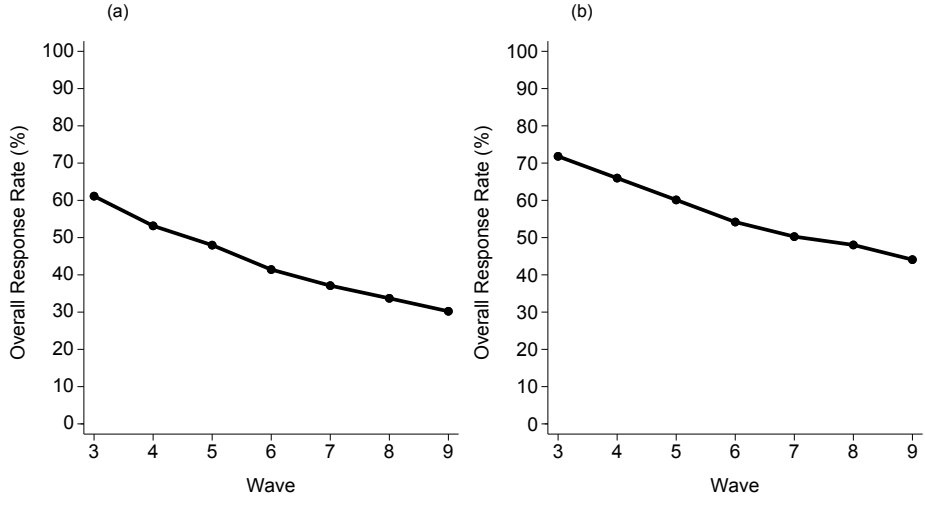
Longitudinal Datasets

RQ1: Differences in Response Rates

Response rates are reported in Figure 1 (see Appendix Table A1 for tabulation). They are 11 to 14 percentage points higher in the COVID-19 Study, likely due to many of the study-eligible sample having previously responded to the main survey and being positively disposed toward the survey (see RQ2: Differences in non-response bias risks section). They also decrease over time (across waves), due to attrition. However, after wave 6 in the COVID-19 Study, the rate of decrease slows slightly, probably partly due to the serology testing for COVID-19 antibodies offered at wave 8 and the greater financial incentives offered at waves 8 and 9 (see UKHLS COVID-19 Study section) reducing attrition. Such testing and incentives are also likely to play a role in explaining patterns observed with CVs in later waves in the COVID-19 Study (see below).

Figure 1

Response Rates for the (a) UKHLS Main Survey; (b) COVID-19 Study Longitudinal Datasets



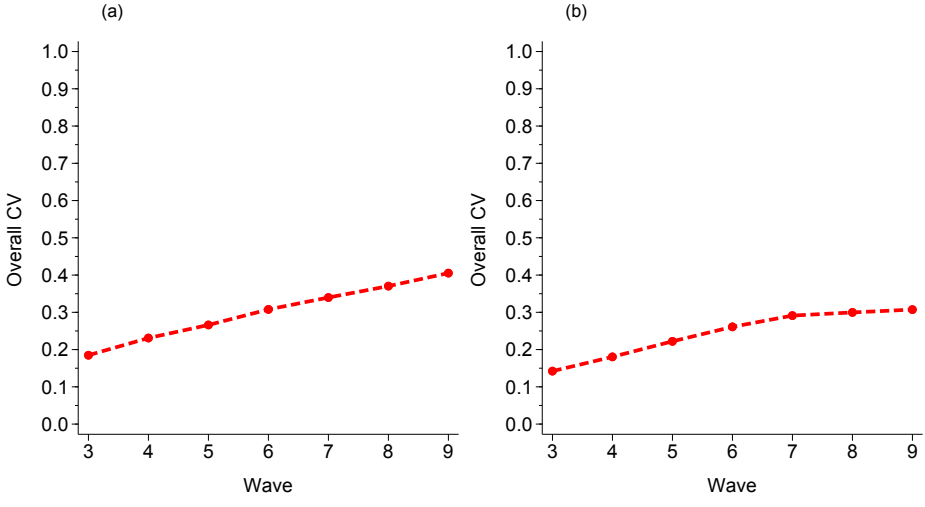
RQ2: Differences in Non-Response Bias Risks

Overall CVs are also reported in Figure 2 (see Table A2 for tabulation and 95% CIs). For all evaluated datasets, CV 95% CIs do not overlap zero, implying statistically significant respondent non-representativeness and non-response bias risks compared to eligible samples. CVs increase over time, implying increasing non-representativeness and bias risks, though at a slower rate after wave 6

in the COVID-19 Study. They are also always significantly smaller (i.e., CV 95% CIs do not overlap) in the COVID-19 Study: at wave 3, 0.14 versus 0.18 in the main survey and at wave 9, 0.31 versus 0.41. This pattern is likely a consequence of the changes in response rates over time in the two surveys and suggests a negative correlation between rates and respondent non-representativeness and bias risks.

Figure 2

Overall CVs for the (a) UKHLS Main Survey; (b) COVID-19 Study Longitudinal Datasets

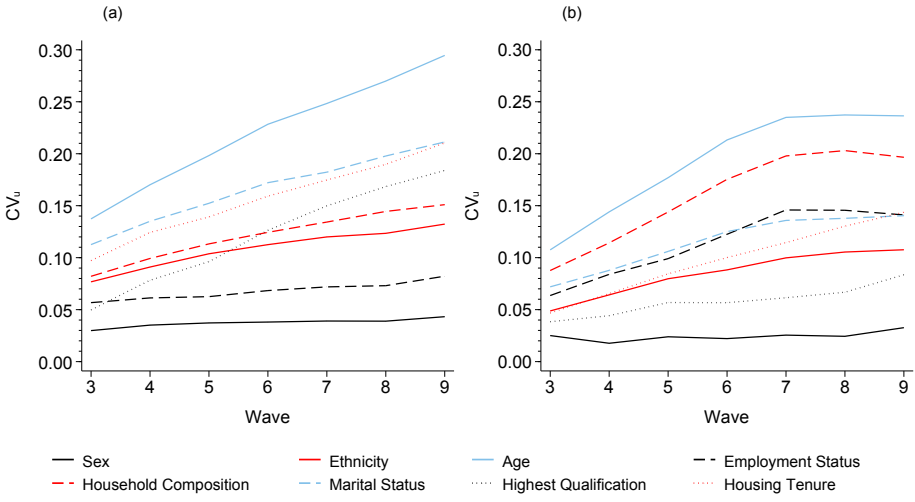


RQ3: Differences in the Correlates of Response Propensity Variation

Auxiliary covariate CV_us are reported in Figure 3 (see Appendix Table A3 for tabulation and 95% CIs). In both surveys, CV_us for all covariates differ significantly from zero, implying that respondents are non-representative with respect to the covariate and mostly increase over time. In the COVID-19 Study though, they may remain similar to wave 7. As with overall CVs, CV_us are smaller in the COVID-19 Study. Covariate CV_cs are reported in Appendix Figure A1 (see Appendix Table A4 for tabulations and 95% CIs). These quantify multivariate associations, with a significant CV_c and a similar CV_u implying an impact on representativeness independent of other auxiliary covariates (see RQ3: Differences in the correlates of response propensity variation section). They are smaller than CV_us but follow largely similar patterns, with all the covariates almost always significant, implying that they have independent impacts on representativeness.

Figure 3

Covariate CV_us for the: (a) UKHLS Main Survey; and (b) COVID-19 Study Longitudinal Datasets



It is not possible to report on category CVs for all auxiliary covariates here due to space constraints, so three important covariates are focused on: age, highest qualification, and tenure. Under-represented categories, the targets of bias prevention measures, are reported. With CV_us, zero values imply that respondents are representative with respect to the category, and significant (i.e., with 95% CIs that do not overlap zero) negative values imply under-representation. Category CV_us for the focal covariates are reported in Figures 4, 5, and 6, and tabulations of CV_us and their 95% CIs for all auxiliary covariate categories are in Appendix Table A5. CV_us are mostly smaller, and fewer under-represented categories occur in the COVID-19 Study. Concerning age in the main survey those aged 16-34 are significantly increasingly under-represented over time, and those aged 65+ are similarly under-represented from wave 7. In contrast, in the COVID-19 Study, those aged 16-44 are significantly under-represented. Under-representation increases over time, but only at earlier waves for those aged 25-44.

Concerning highest qualification, in the main survey, those with A-levels, GCSEs, or no qualifications are significantly under-represented, increasingly so over time, especially for those with no qualifications. Those with other qualifications are significantly increasingly under-represented from wave 7. In contrast, in the COVID-19 Study, those with other qualifications are significantly increasingly under-represented from wave 5, those with A-levels from wave 8, and those with No qualifications at wave 9. Concerning tenure, in the main survey, those in rented accommodation are significantly increasingly under-repre-

sented, as are (though much less so) those in other accommodation. In contrast, in the COVID-19 Study, only those in rented accommodation are significantly under-represented, increasingly so over time (note that possible explanations for some of these patterns are given in the Discussion). Category CV_cs for the focal covariates are reported in Appendix Figures A2, A3, and A4, with tabulations of CV_cs and their 95% CIs for all covariates presented in Appendix Table A6. They have the same interpretation as covariate CV_cs. Although they are smaller than CV_cs, otherwise they follow largely similar patterns. They are significant for all the above categories, implying that they have independent impacts on representativeness.

Figure 4

Covariate Category CV_us for the Covariate Age for the (a) UKHLS Main Survey; and (b) COVID-19 Study Longitudinal Datasets

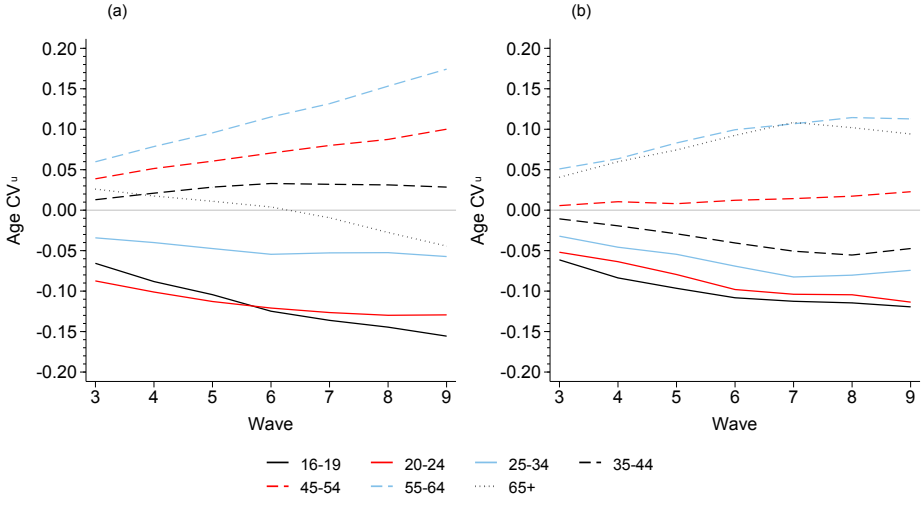


Figure 5

Covariate Category CV_us for the Covariate Highest Qualification for the (a) UKHLS Main Survey; and (b) COVID-19 Study Longitudinal Datasets

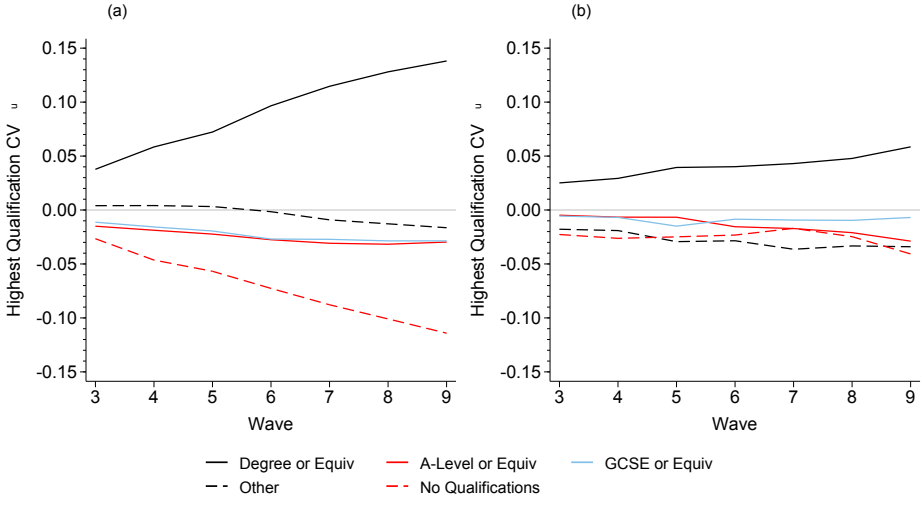
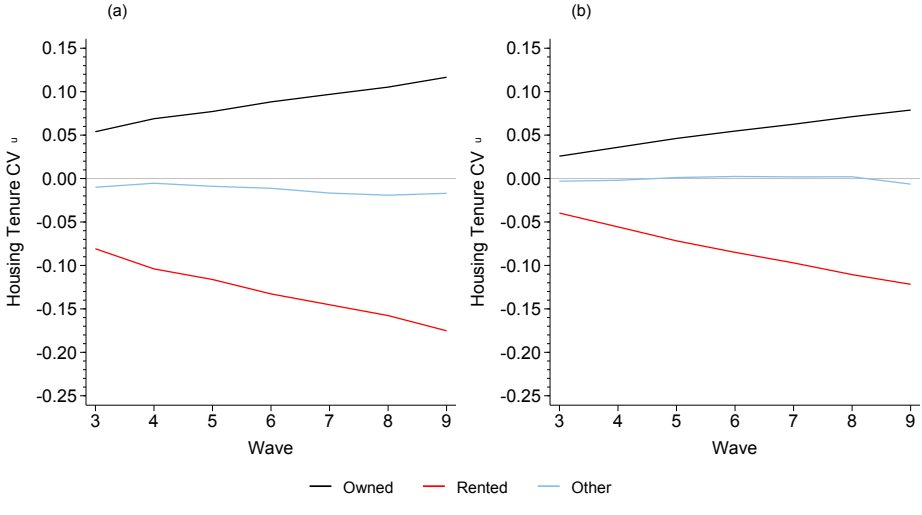


Figure 6

Covariate Category CV_us for the Covariate Tenure for the (a) UKHLS Main Survey; and (b) COVID-19 Study Longitudinal Datasets



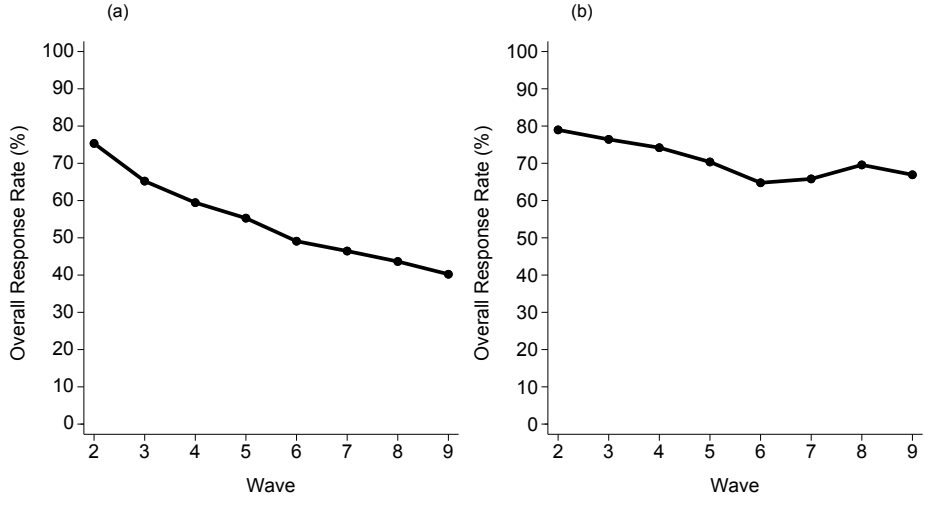
Cross-Sectional Datasets

RQ1: Differences in Response Rates

Response rates are reported in Figure 7 (see Appendix Table A7 for tabulation). Rates are higher than in longitudinal datasets, as expected given the less restrictive inclusion criteria (all respondents to a wave rather than only respondents to the wave and all previous waves). They are also higher in the COVID-19 Study, a difference that increases over time to 27 percentage points because main survey rates decrease over time, whereas COVID-19 Study rates decrease (at a slower rate) only until wave 6, then increase slightly. As with the longitudinal datasets, this pattern is likely due to study sample members being positively disposed toward the survey and the offering of serology testing at wave 8 and greater financial incentives at waves 8 and 9, which in addition here may have led to previous wave non-respondents re-engaging with the survey. Such testing and incentives also probably play a role in explaining patterns observed with CVs at later waves in the COVID-19 Study (see below).

Figure 7

Response Rates for the (a) UKHLS Main Survey; (b) COVID-19 Study Cross-Sectional Datasets



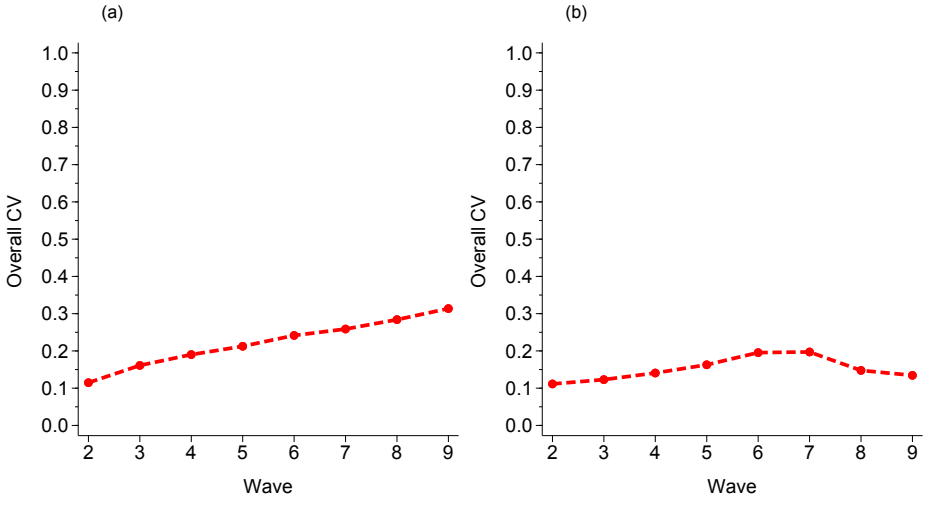
RQ2: Differences in Non-Response Bias Risks

Overall CVs are also reported in Figure 8 (see Appendix Table A8 for tabulation and 95% CIs). They are smaller than in longitudinal datasets, probably due to higher response rates, though they are always statistically significant. CVs in the

main survey increase over time, but in the COVID-19 Study, they increase (at a slower rate) only to wave 7, then decrease. After wave 2, they are significantly smaller in the COVID-19 Study, though differences are smaller (at wave 3, 0.12 versus 0.16 in the main survey; at wave 9, 0.13 versus 0.31) than in longitudinal datasets. As with the latter datasets, this pattern is likely a consequence of the changes in response rates over time in the two surveys and suggests a negative correlation between rates and respondent non-representativeness and bias risks.

Figure 8

Overall CVs for the (a) UKHLS Main Survey; (b) COVID-19 Study Cross-Sectional Datasets

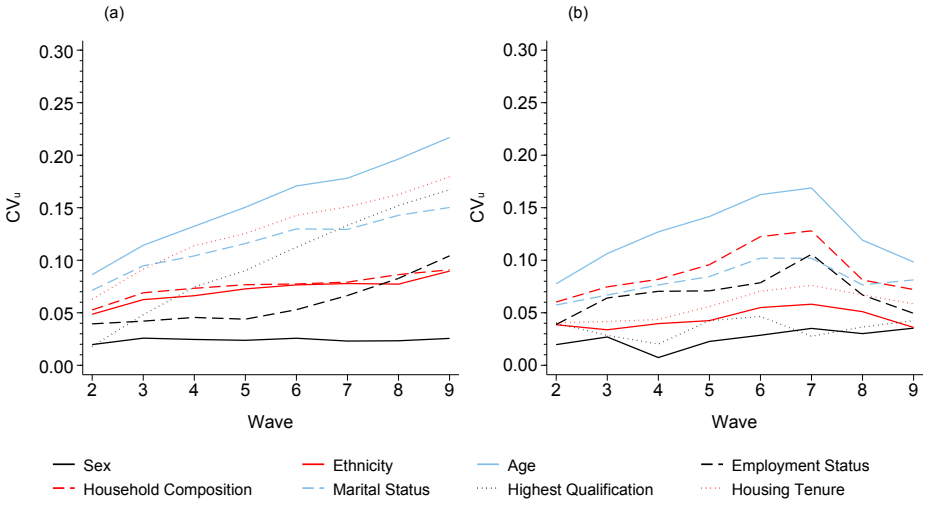


RQ3: Differences in the Correlates of Response Propensity Variation

Auxiliary covariate CV_us are reported in Figure 9 (see Appendix Table A9 for tabulation and 95% CIs). As with overall CVs, they tend to be smaller than in longitudinal datasets and, although all significant, smaller in the COVID-19 Study. They mostly increase over time, though in the COVID-19 Study they often decrease from wave 7. Covariate CV_cs are reported in Appendix Figure A5 (see Appendix Table A10 for tabulation and 95% CIs). As in longitudinal datasets, they are smaller than CV_us but otherwise follow largely similar patterns. They are almost always significant, implying that the covariates have independent impacts on representativeness.

Figure 9

Covariate CV_us for the: (a) UKHLS Main Survey; and (b) COVID-19 Study Cross-Sectional Datasets



Category CV_us for the focal covariates age, highest qualification, and tenure are reported in Figures 10, 11, and 12, with tabulations of CV_us and their 95% CIs for all auxiliary covariate categories presented in Appendix Table A11. They are mostly smaller than in longitudinal datasets and smaller in the COVID-19 Study, with fewer under-represented categories occurring. Concerning age, in the main survey, those aged 16-34 are significantly increasingly under-represented over time, and those aged 65+ are similarly under-represented from wave 7. In contrast, in the COVID-19 Study, those aged 16-24 are significantly increasingly under-represented up to waves 7-8, then under-representation decreases. Those aged 25-34 and 35-44 exhibit a similar pattern from waves 3 and 5, respectively, but are no longer significant at wave 9.

Concerning highest qualification, in the main survey, those with no qualifications are significantly increasingly under-represented. Those with A levels, GCSEs or other qualifications are often significantly slightly under-represented. In contrast, in the COVID-19 Study, only those with GCSEs are significantly under-represented, at wave 5. Concerning tenure, in the main survey, those in rented accommodation are significantly increasingly under-represented. Those in other accommodations are often significantly slightly under-represented. In contrast, in the COVID-19 Study, only those in rented accommodation are significantly under-represented, at a similar rate over time. Category CV_cs for the focal covariates are reported in Appendix Figures A6, A7, and A8, and tabulations of CV_cs and their 95% CIs for all covariates are in Appendix Table A12. As in the longitudinal datasets, they are smaller than CV_us but otherwise follow

similar patterns. CV_cs are significant for all the above categories, implying that they have independent impacts on representativeness.

Figure 10

Covariate Category CV_us for the Covariate Age for the (a) UKHLS Main Survey; and (b) COVID-19 Study Cross-Sectional Datasets

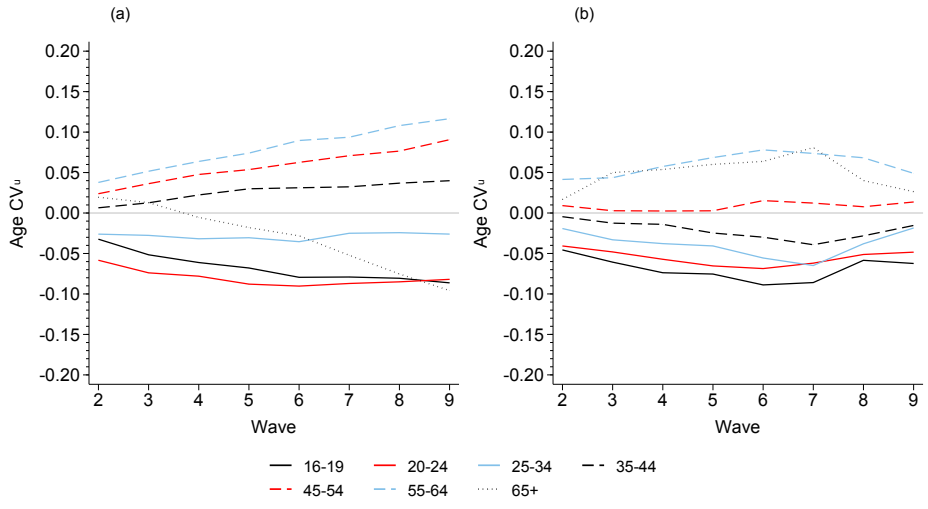


Figure 11

Covariate Category CV_us for the Covariate Highest Qualification for the (a) UKHLS Main Survey; and (b) COVID-19 Study Cross-Sectional Datasets

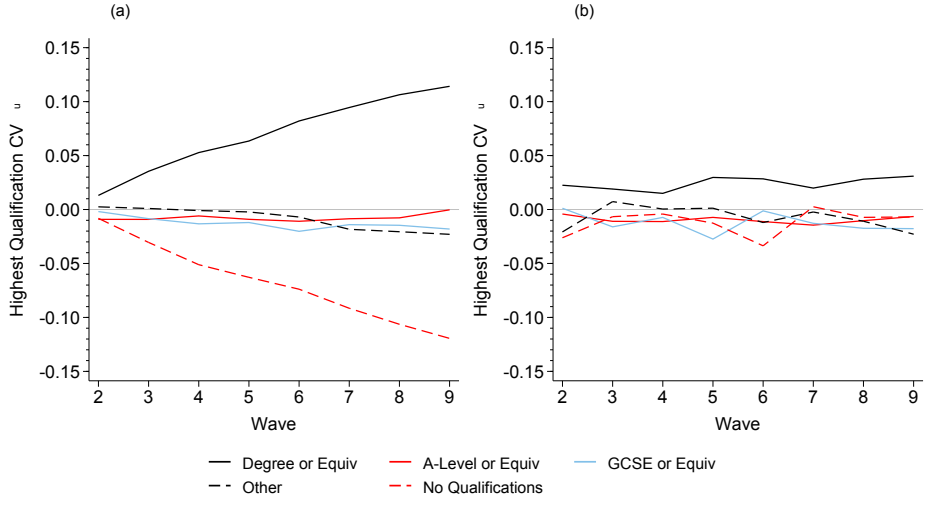


Figure 12

Covariate Category CV_us for the Covariate Tenure for the (a) UKHLS Main Survey; and (b) COVID-19 Study Cross-Sectional Datasets

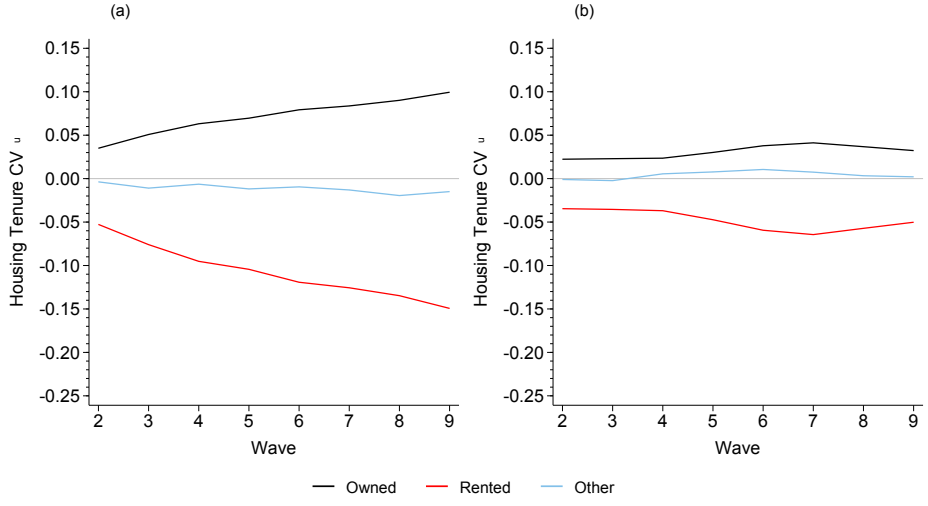


Table 3

Summary of the Results of the Evaluations

Research question	Longitudinal	Cross-sectional
RQ1: Do response rates differ between surveys?	Higher in COVID-19 Study. Decrease over time (waves), but rate slows from wave 7 in COVID-19 Study.	Higher in COVID-19 Study. Higher than in longitudinal datasets. Decrease over time but increases in wave 7 in COVID-19 Study.
RQ2: Do non-response bias risks differ between surveys?	Lower (representativeness greater) in COVID-19 Study. Increase over time, but rate slows after wave 7 in COVID-19 Study.	Lower (representativeness greater) in COVID-19 Study. Lower than in longitudinal datasets. Increase over time but decrease from wave 7 in COVID-19 Study.

Table 3 *continued*

Research question	Longitudinal	Cross-sectional
RQ3: Do the correlates of response propensity variation differ between surveys?	Covariates: In both surveys, all eight auxiliary covariates have significant independent impacts on representativeness at all waves. Impacts mostly smaller in COVID-19 Study. Impacts mostly increase over time, but rate often slows after wave 7 in COVID-19 Study. Categories (with significant independent impacts): Impacts mostly smaller in COVID-19 Study. Impacts mostly increase over time, less so at later waves in COVID-19 Study. Fewer significant impacts in COVID-19 Study. Impact rankings (largest to smallest at wave 3) in main survey: Rented accomm.; no qualifications; age 20-24; age 25-34; age 16-19; A-level qualifications; GCSE qualifications; other accomm. In COVID-19 Study (no highest qualification categories at wave 3): age 25-34; age 35-44; age 16-19; age 20-24; rented accomm.	Covariates: In both surveys, all eight auxiliary covariates have significant independent impacts on representativeness at most waves. Impacts mostly smaller in COVID-19 Study. Impacts mostly smaller than in longitudinal datasets. Impacts mostly increase over time, but rate often slows after wave 7 in COVID-19 Study. Categories (with significant independent impacts): Impacts mostly smaller in COVID-19 Study. Impacts mostly increase over time, trends vary after waves 6 to 7 in COVID-19 Study. Fewer significant impacts in COVID-19 Study. Impact rankings (largest to smallest at wave 2) in main survey: Rented accomm.; age-20-24; age 25-35; no qualifications; age 16-19; A-level qualifications. In COVID-19 Study (no highest qualification categories at wave 2): age 20-24; age 16-19; rented accomm.

Note. Concerning RQ3, the paper focuses on under-represented age, highest qualification, and tenure categories, with those with significant negative unconditional CVs and significant conditional CVs (i.e., with independent impacts on representativeness) at first waves studied mentioned here (see also RQ3: Differences in the correlates of response propensity variation section). The first longitudinal wave studied is wave 3. The first cross-sectional wave studied is wave 2. Auxiliary covariate impacts are not ranked because all are significant and independent and should be included in bias adjustment models.

Discussion

This paper evaluates whether components of dataset performance in the pre-COVID-19 pandemic UKHLS main survey were predictive of those in the UKHLS COVID-19 Study. The main survey is an annual panel survey that uses F2F, telephone, and web modes. The COVID-19 Study was a mostly web-based, higher-frequency survey fielded to collect information about pandemic impacts, with an eligible sample drawn from main survey participants. Even given such design differences, the only information available on likely response rates, non-response bias risks, and correlates of response propensity variation in the COVID-19 Study was from the main survey. The latter correlates are important because surveys often seek to minimize biases by using bias prevention measures to increase responses in under-represented sample subgroups (Groves, 2006; Groves et al., 2009; Peytchev et al., 2010) and/or post-data collection bias adjustment measures that incorporate auxiliary covariates defining them (for example, non-response weights: Valliant et al., 2018). Both longitudinal (only respondents to the wave and all previous waves) and cross-sectional (all respondents) datasets are evaluated. Eligible samples consist of wave 1 respondents, as no information existed on main survey wave 1 non-respondents. Waves 2 to 9 cross-sectional and waves 3 to 9 longitudinal datasets are studied.

Research findings are summarized in Table 3. Response rates are higher in the COVID-19 Study. This finding, which is contrary to most other work comparing web and mixed-mode surveys (see, for example, studies cited in Moore et al., 2025; Olson et al., 2021), is probably due to its sample including main survey respondents likely positively disposed towards surveys. Rates also generally decrease over time due to attrition, though less so in cross-sectional datasets due to the less restrictive inclusion criteria, meaning that they are higher than in longitudinal datasets. That said, in the COVID-19 Study, rates of decrease slow (longitudinal datasets) or response rates increase (cross-sectional datasets) at waves 8-9. This is likely due to COVID-19 antibody testing and increased financial incentives to complete the survey being offered at these waves, reducing attrition and leading non-respondents to re-engage with the survey.

To evaluate non-response bias risks, overall Coefficients of Variation of response propensities (CVs: see Schouten et al., 2012 and RQ2: Differences in non-response bias risks section) are computed. These quantify variation in response propensities estimated using a model with auxiliary covariates available for all sample members. Low variation implies representativeness compared to the sample with respect to the auxiliary covariates and low bias risks. Statistical inference is possible. Eight auxiliary covariates from wave 1 of the surveys are included in response propensity models. Sample members are weighted using non-response weights supplied with the surveys to adjust for sample selection and map them to the UK population. All datasets studied are non-representative of the samples. Representativeness is higher in the COVID-19 Study, a finding

that is contrary to that of other work comparing web and mixed-mode surveys (see Cornesse & Bosnjak, 2018, for a meta-analysis). It is also higher in cross-sectional datasets and mostly decreases over time, though in the COVID-19 Study decreases are slow (longitudinal datasets) or reverse (cross-sectional datasets) at waves 8 and 9. Given also these latter findings, it is probable that the observed between survey differences are a consequence of the response rate differences between them and negative correlations between rates and non-representativeness / bias risks.

To evaluate correlates of response propensity variation, partial covariate and covariate category CVs are computed (see Schouten et al., 2012; and RQ3: Differences in the correlates of response propensity variation section). These decompose propensity variation into that associated with auxiliary covariates / categories. They identify covariates with independent impacts on representativeness to include in bias adjustment models and similar under-represented subgroups to target with bias prevention measures. All auxiliary covariates almost always have such impacts in both surveys. Reflecting overall CVs, impacts are smaller in the COVID-19 Study and cross-sectional datasets and often increase over time, though in the COVID-19 Study often rates decline / impacts decrease at waves 8 and 9. This implies that all the covariates should be included in bias adjustment models for both surveys. Note, though, that the performance of such models at reducing remaining biases increases with unadjusted dataset representativeness (Moore et al., 2024; Schouten et al., 2016), so they would be expected to perform better with the more representative COVID-19 Study datasets.

To save space, category CVs for only the covariates age, highest qualification, and tenure are reported on in the text (those for the other covariates are presented in the Appendix). Again, impacts are mostly smaller in the COVID-19 Study and in cross-sectional datasets. There are fewer under-represented categories in the COVID-19 Study. In longitudinal datasets, in the main survey, those aged 16-34 and from wave 7, those aged 65+, are under-represented, whereas in the COVID-19 Study, those aged 16-44 are. In addition, in the main survey, those with A levels, GCSEs, other qualifications (from wave 7), and no qualifications are under-represented, whereas in the COVID-19 Study, only the first three categories are. Moreover, in the main survey, those in both rented and other accommodations are under-represented, whereas in the COVID-19 Study, only the former are. Age and tenure results are similar in cross-sectional datasets, though some categories are not always under-represented. Moreover, in the COVID-19 Study, only those with GCSEs are under-represented, at wave 5. The ranking of these categories in terms of impact at first evaluated waves is reported in Table 3.

At least one of the above differences between surveys is probably due to the survey designs. Under-representation of those aged 65+ only in the main survey likely reflects 9 years of sample member mortality, whereas COVID-19 Study duration was only 18 months. Other differences may be due to changes in living

conditions during the pandemic. That those aged 35-44 are only under-represented in the COVID-19 Study (typically, those in younger age groups are also under-represented in surveys other than the main survey: see, for example, Behr, 2005) could reflect school closures during parts of the pandemic. This would have reduced the time parents had to complete surveys due to needing to supervise children (a similar pattern also exists with HH composition: couple with children; see Appendix). That there is less under-representation of highest qualification categories in the COVID-19 Study (generally, those with low education levels are under-represented in surveys other than the main survey: see, for example, Maslovskaya & Lugtig, 2022) could reflect individuals furloughed from their jobs for parts of the pandemic having more time to complete the survey. Importantly, fewer under-represented categories in the COVID-19 Study imply that the use of bias prevention measures based on main survey findings would have led to greater costs than necessary and also, if successful, potentially led to otherwise representative categories becoming over-represented, increasing bias risks. Rankings of categories in terms of impact, which would be useful to prioritize bias prevention measures, are reported in Table 3.

There are three limitations to this research. First, actual non-response biases are not quantified. This is because, as is often the case (see Hand, 2018), relevant population benchmarks do not exist. Instead, weights are used to map sample members to the UK population. Benzeval et al. (2020) and Moore et al. (2024) provide evidence of the utility of these weights, though it should be noted that for the COVID-19 Study they only adjust for the socio-demographic aspects of sample selection: whether further biases are caused by many sample members previously responding to the main survey and being positively disposed towards surveys (see earlier) is unknown. Second, the weights map the two surveys to the UK population at different time points: March 2009–January 2011 with the main survey and June 2019 with the COVID-19 Study. Due to the time periods covered by the two sets of evaluated datasets, there was no alternative to this. That said, even given these limitations, the evaluations still address the research question that is the focus of the paper: whether information on main survey dataset performance is useful to predict similar results in the COVID-19 Study.

The third limitation concerns the use of auxiliary covariate information from wave 1. This presents issues in a longitudinal context because, although some covariates do not change over time (for example, ethnicity), others do (for example, age), with the consequence that at later waves some individuals will be wrongly categorized. There was no alternative to this approach because the CVs used in this paper require auxiliary covariate information for all sample members. This issue affects the main survey evaluations more than the COVID-19 evaluations due to the difference in time periods datasets cover (9 years vs. 18 months). Its likely impact on results is difficult to predict, though it should be noted that again the evaluations still address the question of whether main

survey information on dataset performance is useful to predict similar results in the COVID-19 Study.

Notwithstanding these limitations, the findings of the research in this paper imply that components of dataset performance in the main survey were of limited use in predicting similar in the COVID-19 Study. This is unsurprising given the changes in survey design and living conditions during the pandemic but has not previously been shown. This does not mean that the COVID-19 Study datasets were of poor quality; indeed, the findings suggest they had lower non-response bias risks than the main survey. Equally, the main survey has been shown to support high-quality inferences (Benzeval et al., 2020). Rather, it means that response rates, bias risks, and under-represented subgroups differ between surveys, with one important consequence being that using a bias minimization strategy based on main survey information in the COVID-19 Study would have led to unnecessarily inflated costs and different, potentially negative impacts on biases (see previous paragraphs). Fortunately, given this, in reality, custom bias prevention (telephone sampling of web non-respondents at several waves: see RQ1: Differences in response rates section) and adjustment (newly developed non-response weights) measures were used in the study and performed well in reducing biases (Moore et al., 2024).

These findings are important despite the end of the COVID-19 pandemic and the COVID-19 Study. Other pandemics may occur in the future and lead to demand for similar new surveys. In addition, the shift to less costly web mode was already occurring in the main survey, albeit at a slower pace and not to the same extent (Institute for Social and Economic Research 2021), and it is not inconceivable that one day it may completely replace F2F (though currently there are no plans to do so), so information on scenarios where F2F is not used is of value. Should either of these events occur, the findings imply that targeted research will be needed to properly predict impacts on components of dataset performance.

The findings also have relevance beyond the two UKHLS surveys. Other new surveys were fielded during the COVID-19 pandemic (e.g., Blom et al., 2020; Brown et al., 2021; Burton et al., 2020; Gummer et al., 2020). These were often also derived from existing surveys, had increased collection frequencies and mode differences, and were fielded with limited information on how changes would impact datasets. In addition, shifts towards increased web mode use outside of the pandemic are universal (Bianchi et al., 2017; Burton & Jäckle, 2020; Cornesse & Bosnjak, 2018; Couper et al., 2007; Institute for Social and Economic Research, 2025; Nicolaas et al., 2014). Hence, evidence that in such situations information from existing surveys is likely to be of limited use for prediction is of value. The findings here, though, are unlikely to be of use for making more specific inferences. The CVs used to evaluate bias risks depend on the auxiliary covariates used to model response propensities (Schouten et al., 2009, 2012), and

propensity correlates and / or signs of associations can differ between surveys (Behr et al., 2005; Luiten et al., 2020). Hence, as with the UKHLS surveys, targeted research will be needed to accurately predict the impacts of such design changes on datasets.

Acknowledgements

Understanding Society is an initiative funded by the Economic and Social Research Council and various government departments, with scientific leadership by the Institute for Social and Economic Research, University of Essex, and survey delivery by the National Centre for Social Research (NatCen) and Verian (formerly Kantar Public). The research data are distributed by the UK Data Service. The authors would also like to thank Prof. Peter Lynn for helpful comments.

Data and Code Availability Statement

Data are available to download from the UK Data Service:

University of Essex, & Institute for Social and Economic Research. (2021). *Understanding Society: COVID-19 Study, 2020-2021* (SN: 8644; Version 11th Edition) [data collection]. UK Data Service. https://doi.org/10.5255/UKDA-SN-8644-11

University of Essex, & Institute for Social and Economic Research. (2023). *Understanding Society: Waves 1-13, 2009-2022 and Harmonised BHPS: Waves 1-18, 1991-2009* (SN: 6614; Version 18th Edition) [data collection]. UK Data Service. https://doi.org/10.5255/UKDA-SN-6614-19

The Stata scripts are available as supplementary material accompanying this article.

The R code to compute CVs (RISQ 2.1) is available as supplementary material and from:

de Heij, V., Schouten, B., & Shlomo, N. (2015). *RISQ Manual 2.1: Tools in SAS and R for the computation of R-indicators, partial R-indicators and partial coefficients of variation*. Representativity Indicators for Survey Quality. https://hummedia.manchester.ac.uk/institutes/cmist/risq/RISQ-manual-v21.pdf

Conflict of Interest

The authors declare no conflicts of interest.

Ethics

The University of Essex Ethics Committee has approved all data collection on Understanding Society main study, COVID-19 surveys and innovation panel waves, including asking consent for all data linkages except to health records.

Funding

This research was supported by the University of Essex Social Sciences Doctoral Scholarship awarded to Nicole D. James.

References

Baker, R., Blumber, S. J., Brick, J. M., Couper, M. P., Courtright, M., Dennis, J. M., Dillman, D., Frankel, M. R., Garland, P., Groves, R. M., Kennedy, C., Krosnick, J., Lavrakas, P. J., Lee, S., Link, M., Piekarski, L., Rao, K., Thomas, R. K., & Zahs, D. (2010). AAPOR REPORT ON ONLINE PANELS. *Public Opinion Quarterly, 74*(4), 711-781.

Behr, A., Bellgardt, E., & Rendtel, U. (2005). Extent and Determinants of Panel Attrition in the European Community Household Panel. *European Sociological Review, 21*(5), 489-512. https://doi.org/10.1093/esr/jci037

Benzeval, M., Bollinger, C. R., Burton, J., Crossley, T. F., & Lynn, P. (2020). The Representativeness of Understanding Society [Working Paper]. *Understanding Society Working Paper Series, 2020-08*. https://www.understandingsociety.ac.uk/research/publications/publication-526039/

Bianchi, A., Biffignandi, S., & Lynn, P. (2017). Web-Face-to-Face Mixed-Mode Design in a Longitudinal Survey: Effects on Participation Rates, Sample Composition, and Costs. *Journal of Official Statistics, 33*(2), 385-408. https://doi.org/10.1515/jos-2017-0019

Blom, A. G., Cornesse, C., Friedel, S., Krieger, U., Fikel, M., Rettig, T., Wenz, A., Juhl, S., Lehrer, R., Möhring, K., Naumann, E., & Reifenscheid, M. (2020). High-Frequency and High-Quality Survey Data Collection: The Mannheim Corona Study. *Survey Research Methods, 14*(2), 171-178. https://doi.org/10.18148/srm/2020.v14i2.7735

Borkowska, M. (2019). Improving Population and Sub-Group Coverage: Who is missing and what can be done about it? *Understanding Society Working Paper Series, 2019-15*. https://www.understandingsociety.ac.uk/research/publications/publication-525964/

Brown, M., Goodman, A., Peters, A., Ploubidis, G. B., Sanchez, A., Silverwood, R., & Smith, K. (2021). COVID-19 Survey in Five National Longitudinal Studies: Waves 1, 2 and 3 User Guide (Version 3). UCL Centre for Longitudinal Studies and MRC Unit for Lifelong Health and Ageing. https://cls.ucl.ac.uk/wp-content/uploads/2017/02/UCL-Cohorts-COVID-19-Survey-user-guide.pdf

Burton, J., & Jäckle, A. (2020). Mode Effects [Working Paper]. *Understanding Society Working Paper Series*, *2020-05*. https://www.understandingsociety.ac.uk/research/publications/publication-526036/

Burton, J., Lynn, P., & Benzeval, M. (2020). How *Understanding Society*: The UK Household Longitudinal Study Adapted to the COVID-19 Pandemic. *Survey Research Methods*, *14*(2), 235-239. https://doi.org/10.18148/srm/2020.v14i2.7746

Cornesse, C., & Bosnjak, M. (2018). Is There an Association between Survey Characteristics and Representativeness? A Meta-Analysis. *Survey Research Methods*, *12*(1), 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. *Social Science Research*, *36*(1), 131-148. https://doi.org/10.1016/j.ssresearch.2005.10.002

Daikeler, J., Bošnjak, M., & Lozar Manfreda, K. (2020). Web Versus Other Survey Modes: An Updated and Extended Meta-Analysis Comparing Response Rates. *Journal of Survey Statistics and Methodology*, *8*(3), 513-539. https://doi.org/10.1093/jssam/smz008

de Heij, V., Schouten, B., & Shlomo, N. (2015). *RISQ Manual 2.1: Tools in SAS and R for the computation of R-indicators, partial R-indicators and partial coefficients of variation*. Representativity Indicators for Survey Quality. https://hummedia.manchester.ac.uk/institutes/cmist/risq/RISQ-manual-v21.pdf

de Leeuw, E., & de Heer, W. (2002). Trends in Household Survey Nonresponse: A Longitudinal and International Comparison. In R. Groves, D. Dillman, J. L. Eltinge, & R. J. Little (Eds.), *Survey Nonresponse* (pp. 41-54). Wiley.

Groves, R. M. (2006). Nonresponse Rates and Nonresponse Bias in Household Surveys. *Public Opinion Quarterly*, *70*(5), 646-675. https://doi.org/10.1093/poq/nfl033

Groves, R. M., Fowler, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2009). *Survey Methodology* (2nd ed.). Wiley.

Groves, R. M., & Peytcheva, E. (2008). The Impact of Nonresponse Rates on Nonresponse Bias: A Meta-Analysis. *Public Opinion Quarterly*, *72*(2), 167-189. https://doi.org/10.1093/poq/nfn011

Gummer, T., Schmiedeberg, C., Bujard, M., Christmann, P., Hank, K., Kunz, T., Lück, D., & Neyer, F. J. (2020). The Impact of COVID-19 on Fieldwork Efforts and Planning in Pairfam and FReDA-GGS. *Survey Research Methods*, *14*(2), 223-227. https://doi.org/10.18148/srm/2020.v14i2.7740

Hand, D. J. (2018). Statistical Challenges of Administrative and Transaction Data. *Journal of the Royal Statistical Society Series A: Statistics in Society*, *181*(3), 555-605. https://doi.org/10.1111/rssa.12315

Institute for Social and Economic Research. (2021). *Understanding Society COVID-19 User Guide*. University of Essex. https://www.understandingsociety.ac.uk/documentation/covid-19/user-guide/

Institute for Social and Economic Research. (2025). *Understanding Society: Waves 1-15, 2009-2024 and Harmonised BHPS: Waves 1-18, 1991-2009, User Guide*. University of Essex. https://www.understandingsociety.ac.uk/documentation/mainstage/user-guides/main-survey-user-guide/

Little, R. J. A., & Rubin, D. B. (2020). *Statistical Analysis with Missing Data* (3rd ed., Vol. 793). Wiley. https://doi.org/10.1002/9781119482260

Little, R. J. A., & Vartivarian, S. (2005). Does Weighting for Nonresponse Increase the Variance of Survey Means? *Survey Methodology*, *31*(2), 161-168.

Lugtig, P., Das, M., & Scherpenzeel, A. (2014). Nonresponse and Attrition in a Probability-Based Online Panel for the General Population. In M. B. Callegaro, Reg; Bethleham,

Jelke; Göritz, Anja S.; Krosnick, Jon A.; Lavrakas, Paul J. (Ed.), *Online Panel Research* (pp. 135-153). https://doi.org/10.1002/9781118763520.ch6

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*(3), 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*. https://www.understandingsociety.ac.uk/research/publications/publication-514007/

Lynn, P., & Borkowska, M. (2018). Some Indicators of Sample Representativeness and Attrition Bias for BHPS and Understanding Society. *Understanding Society Working Paper Series*, *2018-01*. https://www.understandingsociety.ac.uk/research/publications/publication-524851/

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

Maslovskaya, O., & Lugtig, P. (2022). Representativeness in Six Waves of Cross-National Online Survey (CRONOS) Panel. *Journal of the Royal Statistical Society Series A: Statistics in Society, 185*(3), 851–871. https://doi.org/10.1111/rssa.12801

Moore, J., 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 Paper Series*, *2024-04*. https://www.understandingsociety.ac.uk/research/publications/working-paper/understanding-society/2024-04/

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 [Working Paper]. *Understanding Society Working Paper Series*, *2025-02*. https://www.understandingsociety.ac.uk/research/publications/working-paper/understanding-society/2025-02/

Moore, J., Durrant, G., & Smith, P. (2018). Data Set Representativeness during Data Collection in Three UK Social Surveys: Generalizability and the effects of auxiliary covariate choice. *Journal of the Royal Statistical Society Series A: Statistics in Society, 181*(1), 229–248. https://doi.org/10.1111/rssa.12256

Moore, J., Durrant, G., & Smith, P. (2021). Do Coefficients of Variation of Response Propensities Approximate Non-Response Biases during Survey Data Collection? *Journal of the Royal Statistical Society: Series A (Statistics in Society), 184*(1), 301–323.

Nagelkerke, N. J. D. (1991). A Note on a General Definition of the Coefficient of Determination. *Biometrika, 78*(3), 691-692.

Nicolaas, G., Calderwood, L., Lynn, P., & Roberts, C. (2014). *Web Surveys for the General Population: How, why and when?* (National Centre for Research Methods Report, Issue. National Centre for Research Methods. http://eprints.ncrm.ac.uk/3309/

Nishimura, R., Wagner, J., & Elliott, M. (2016). Alternative Indicators for the Risk of Non-Response Bias: A Simulation Study. *International Statistical Review, 84*(1), 43-62. https://doi.org/10.1111/insr.12100

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., Sterrett, D., Su, J., Suzer-Gurtekin, Z. T., Turakhia, C., & Wagner, J. (2021). Transitions from Telephone Surveys to Self-Administered and Mixed-Mode Surveys: AAPOR Task Force Report. *Journal of Survey Statistics and Methodology, 9*(3), 381–411. https://doi.org/10.1093/jssam/smz062

Peytchev, A., Riley, S., Rosen, J., Murphy, J., & Lindblad, M. (2010). Reduction of Nonresponse Bias in Surveys through Case Prioritization. Survey Research Methods, 4(1), 21-29. https://doi.org/10.18148/srm/2010.v4i1.3037

Särndal, C.-E. (2011). The 2010 Morris Hansen Lecture Dealing with Survey Nonresponse in Data Collection, in Estimation. *Journal of Official Statistics, 27*(1), 1-21.

Schonlau, M., & Couper, M. P. (2017). Options for Conducting Web Surveys. *Statistical Science*, *32*(2), 279-292. https://doi.org/10.1214/16-STS597

Schonlau, M., van Soest, A., Kapteyn, A., & Couper, M. (2009). Selection Bias in Web Surveys and the Use of Propensity Scores. *Sociological Methods & Research*, *37*(3), 291-318. https://doi.org/10.1177/0049124108327128

Schouten, B., Bethlehem, J., Beullens, K., Kleven, Ø., Loosveldt, G., Luiten, A., Rutar, K., Shlomo, N., & Skinner, C. (2012). Evaluating, Comparing, Monitoring, and Improving Representativeness of Survey Response Through R-Indicators and Partial R-Indicators. *International Statistical Review, 80*(3), 382-399. https://doi.org/10.1111/j.1751-5823.2012.00189.x

Schouten, B., Cobben, F., & Bethlehem, J. (2009). Indicators for the Representativeness of Survey Response. *Survey Methodology, 35*(1), 101-113.

Schouten, B., Cobben, F., Lundquist, P., & Wagner, J. (2016). Does More Balanced Survey Response Imply Less Non-Response Bias? *Journal of the Royal Statistical Society Series A: Statistics in Society, 179*(3), 727-748. https://doi.org/10.1111/rssa.12152

Schouten, B., Shlomo, N., & Skinner, C. (2011). Indicators for Monitoring and Improving Representativeness of Response. *Journal of Official Statistics, 27*(2), 231-253.

Shlomo, N., & Schouten, B. (2013). *Theoretical Properties of Partial Indicators for Representative Response*. https://hummedia.manchester.ac.uk/institutes/cmist/risq/shlomo-schouten-2013.pdf

Shlomo, N., Schouten, B., & de Heij, V. (2013, 5–7 March). *Designing Adaptive Survey Designs using R-Indicators* New Techniques and Technologies for Statistics Conference, Brussels, Belgium. https://hummedia.manchester.ac.uk/institutes/cmist/risq/shlomo-schouten-heij-2013.pdf

Shlomo, N., Skinner, C., & Schouten, B. (2012). Estimation of an Indicator of the Representativeness of Survey Response. *Journal of Statistical Planning and Inference, 142*(1), 201-211. https://doi.org/10.1016/j.jspi.2011.07.008

Uhrig, S. C. N. (2008). The Nature and Causes of Attrition in the British Household Panel Survey. *ISER Working Paper Series*, *2008-05*. https://www.iser.essex.ac.uk/research/publications/working-papers/iser/2008-05

University of Essex, & Institute for Social and Economic Research. (2021). *Understanding Society: COVID-19 Study, 2020-2021* (SN: 8644; Version 11th Edition) [data collection]. UK Data Service. https://doi.org/10.5255/UKDA-SN-8644-11

University of Essex, & Institute for Social and Economic Research. (2023*). Understanding Society: Waves 1-13, 2009-2022 and Harmonised BHPS: Waves 1-18, 1991-2009* (SN: 6614; Version 18th Edition) [data collection]. UK Data Service. https://doi.org/10.5255/UKDA-SN-6614-19

Valliant, R., Dever, J. A., & Kreuter, F. (2018). *Practical Tools for Designing and Weighting Survey Samples* (2nd ed.). Springer International Publishing. https://doi.org/10.1007/978-3-319-93632-1

Wagner, J. (2010). The Fraction of Missing Information as a Tool for Monitoring the Quality of Survey Data. *Public Opinion Quarterly*, *74*(2), 223-243. https://doi.org/10.1093/poq/nfq007

Online First publication date: 2026-11-23