Appendix G Weighting Technical Summary

Introduction

Data from the experimental statistics phase were weighted to take account of non-response, bias, and to improve representativeness. Three weights were produced for analysis: one for step 1 fieldwork, one for step 2 fieldwork and a combined (steps 1 and 2 fieldwork). These are described in Figure 15, Experimental phase weights.

Figure 15: Experimental phase weights

Description of weight	Name of weight	Number of cases weighted
Weight for step 1 fieldwork	step1_final_wt	5275
Weight for step 2 fieldwork	step2_final_wt	3549
Step 1 and 2 weight	step1and2_final_wt	8824 (3549 plus 5275)

There was no disproportionate sampling for either fieldwork step. To produce the step 1 fieldwork weight (step1_final_wt) and step 2 fieldwork weight (step2_final_wt) the same two processes were carried out:

- a logistic regression model for non-response within a household (run for households with more than one adult eligible to take part in the survey), then

- a calibration to population estimates

A household non-response model was considered and tested on data from the 2021 pilot survey. However, it did not meaningfully improve the weights and so was not done for the experimental phase.

To produce the combined weight for step 1 and step 2 (step1and2_final_wt), the separate step 1 and step 2 weights were simply stacked then checked for bias and outliers.

Running a logistic regression model for non-response

Addresses issued for step 1 fieldwork were randomly allocated to one of two selection types ('conditions') which specified the maximum number of adults (aged 18 years and over) to be selected from each address to complete the survey. In condition 1 addresses (14610 issued), up to two adults were selected; in condition 2 addresses (7343 issued), up to four adults were selected.

As a first step in the weighting process, a number of household and area level variables were tested for association with one or two fully productive adults per household. As only 91 condition 2 households yielded three or four completions, the model was simplified to a logistic regression of either 'one completion' or 'two or more' completions, including:

- Government Office Region (GOR)

- tenure

- socio-economic classification (area only)

- social grade (area only)

- education (area only)

- employment (area only)

- ethnicity (area only)

- car ownership (area only)

- urban-rural classification

- number of adults

- number of children

- household income

- household type

- age

- output area classification

- population density

- Index of Multiple Deprivation.

- which questionnaire approach the household was assigned to (step 2 sample only)

The final non-response model for step 1 included: GOR, household income, household tenure, and whether the household contained someone of retirement age (employment). The final response model for step 2 included: GOR, household income, household type, and output area classification. The predicted probabilities from the models were used to create non-response weights for households with more than one eligible adult. Both weights were checked for outliers and left untrimmed. Weights for responding households with only one eligible adult were set to one.

Calibration to population estimates

The non-response weights were then calibrated to estimates of the eligible population. Calibration weighting adjusts the weights so that characteristics of the weighted achieved sample match population estimates, reducing bias. Population figures for calibration were taken from, the ONS mid-year population estimates (Estimates of the population for the UK, England, Wales, Scotland and Northern Ireland - Office for National Statistics (ons.gov.uk)),[1] the Office for National Statistics (ONS) Index of Multiple Deprivation (IMD) estimates (Populations by Index of Multiple Deprivation (IMD) decile, England and Wales, 2020 - Office for National Statistics and Population Estimates by Scottish Index of Multiple Deprivation (SIMD) | National Records of Scotland (nrscotland.gov.uk)) and the most recent Labour Force Survey (for step 1: UK Data Service › Study[2] and for step 2 UK Data Service › Study).[3]

Calibration stages used the same variables for the step 1 and step 2 weights: age, sex, GOR, IMD, tenure, ethnicity, and education. For the step 1 weight, dropping tenure and

---

[1] Mid-2020 estimates for step 1 weighting and mid-2021 for step 2 weighting (more up-to-date estimates were available at the time the step 2 weighting was carried out).

[2] Quarterly Labour Force Survey, April - June 2022.

[3] Quarterly Labour Force Survey, July - September 2022.

education from the calibration was tested, but it did not meaningful improve the weights, and so was not used.

After calibration, the step 1 and step 2 weights were checked for outliers but left untrimmed. To produce the combined weight for step 1 and step 2 (step1and2_final_wt), the separate step 1 and step 2 weights were simply stacked then checked for bias and outliers, and it was found no changes were needed. Figure16 provides a summary of the design effects and other statistics of the weights.

Figure16: Summary of the design effects and other statistics of the weights

Weight	Number	Design Effect (2d.p.)	Effective sample size (0d.p.)	Efficiency (0d.p.)
step1_final_wt	5275	1.33	3955	75%
step2_final_wt	3549	1.29	2750	77%
step1and2_final_wt	8824	1.32	6703	76%