UCL Institute of Education

Next Steps

2019 Web Survey

User Guide (Version 1)

October 2025

CENTRE FOR LONGITUDINAL STUDIES

Economic and Social Research Council

Contact

Data queries: help@ukdataservice.ac.uk

Questions and feedback about this user guide: clsdata@ucl.ac.uk.

Authors

Matt Brown, Morag Henderson, Darina Peycheva, Sarab Rihal, Liam Wright.

How to cite this guide

Brown, M., Henderson, M., Peycheva, D., Rihal, S., Wright, L. (2025) *Next Steps 2019 Web Survey User Guide (Version 1)*. UCL Centre for Longitudinal Studies.

Data citation and acknowledgement

You should cite the data and acknowledge CLS following the guidance from cls.ucl.ac.uk/data-access-training/citing-our-data/

Centre for Longitudinal Studies

Centre for Longitudinal Studies (CLS)

UCL Social Research Institute

University College London

20 Bedford Way, London WC1H 0AL

www.cls.ucl.ac.uk

The UCL Centre for Longitudinal Studies (CLS) is an Economic and Social Research Council (ESRC) Resource Centre. It is home to a unique series of UK national cohort studies. It is part of the UCL Social Research Institute, based at the IOE, UCL Institute of Education.

This document is available in alternative formats. Please email the Centre for Longitudinal Studies at clsdata@ucl.ac.uk.

Contents

About Next Steps

Next Steps is a longitudinal cohort study, following a nationally representative group of nearly 16,000 people born in England in 1989-90. The study began when cohort members were 14 years old. With sweeps every year for the first seven years, it has captured rich information about their educational trajectories during adolescence.

Next Steps has since documented early adulthood experiences at Age 25 and Age 32. These adulthood sweeps have a wider scope, and include measures of health, wellbeing, family formation and labour market outcomes (among others), providing unparalleled insight into the many different aspects of this millennial generation's lives.

A vital source of evidence, Next Steps has had a major influence on national education policy and cast light on a wide range of important social issues, including the association between zero-hours contracts and mental health.

This new deposit includes the data collected in a short web survey conducted with Next Steps participants in 2019 when they were aged 29-30. Participants were asked about political attitudes, political engagement, voting in the 2017 election and the 2016 EU Referendum, personal achievements and societal challenges faced by their generation.

1. Introduction

There were seven initial sweeps of Next Steps, conducted annually by the Department for Education between 2004 and 2010. The Age 25 Survey (sweep 8 – 2015/16) and the Age 32 Survey (sweep 9 – 2022/23) were run by the Centre for Longitudinal Studies (CLS) at the UCL Institute of Education (IOE).

The Next Steps 2019 Web Survey took place between August and September 2019, in between the Age 25 and Age 32 Surveys. It was conducted by CLS. CLS conducts annual 'keeping-in-touch' exercises in which Next Steps participants are asked to confirm or update their contact details. The 2019 Web Survey was conducted as part of the 2019 keeping-in-touch exercise.

This user guide provides information about the data arising from the Next Steps 2019 Web Survey and accompanies the deposit of the data at the UK Data Service.

In addition to this user guide the data deposit includes:

Next Steps – 2019 Web Survey Questionnaire.

Data, questionnaires and user guides for all other sweeps are also available at UKDS. All datasets use a common ID – NSID.

2. The 2019 Web Survey

CLS conducts an annual ‘keeping-in-touch’ exercise in which Next Steps participants are sent an Update leaflet which provides news and findings from the study and are asked to confirm or update their contact details. In 2019, Next Steps participants were asked, for the first time, to update their contact details via an online survey.

CLS conducted an experiment in which three different approaches to requesting contact details updates were trialled. The invitation to the first group simply asked participants to login to the survey to check and update their contact details. The invitation to the second group told participants that they would additionally be asked to provide some feedback about their experience of taking part in Next Steps. The invitation to the third group told participants that they would be asked some additional questions about their lives. Participants were randomly allocated to these three groups. The aim of the experiment was to assess which approach would be most effective in encouraging completion of the survey.

The survey was conducted using Qualtrics – an online survey platform.

The data in this deposit are the answers to the substantive questions put to participants in the third group of the experiment.

3. Fieldwork

3.1 Issued sample

A total of 12,242 study members were invited to complete the survey. 3,622 were randomly allocated to the group to be asked substantive questions.

3.2 Fieldwork period

Fieldwork was conducted between 6th August 2019 and 13th September 2019.

3.3 Fieldwork stages

A sub-sample of 500 study members were randomly allocated to a 'soft-launch'. The soft-launch was conducted to ensure that the web survey was working as intended prior to invitations being sent to all participants. Soft-launch fieldwork began two weeks earlier than the remaining fieldwork.

3.4 Contact strategy

Participants were sent an invitation via post and email. Two email reminders and one postal reminder was sent to those who had not completed the web survey.

The letters and emails sent to the three groups only differed with regard to the description of the content of the survey. All participants were informed that completing the survey would take 5-10 minutes.

3.5 Incentives

Next Steps typically offers incentives for completion of surveys, but on this occasion no incentive was offered.

4. Response rate

4.1 Overall response

A total of 12,242 study members were invited to complete the survey and 2,646 did so, resulting in a response rate of 22%.

3,622 were invited to complete the version of the survey which included substantive questions and 727 did so, resulting in a response rate of 20% (there was no significant difference in response rate between any of the experimental groups). It is this group whose data is included in this deposit. Amongst the 727 participants there were 9 partial completions.

The response rate for the group allocated to the substantive questions is shown in Table 1 below:

Table 1. Overall response for 2019 Web Survey

Outcome	Frequency	Percent
Productive	727	20%
Fully productive	*718*	
Partially productive	*9*	
Unproductive	2,895	80%
Total issued sample	3,622	100%

5. Questionnaire

The substantive questions included in the 2019 Web Survey covered political attitudes and interest, voting in the 2017 General Election, voting in the 2016 EU Referendum, desired outcome from Brexit, personal achievements and aspirations, and societal challenges facing the Next Steps generation. A more detailed summary is provided in Table 2 below.

Table 2. Questionnaire content

Module title	Full questionnaire content
Politics	Political opinions: Agreement/Disagreement with the following statements: • *It is easier now for people like me to get on and improve things for themselves than it was for my parents.* • *Britain today is a place where hard work is rewarded.* • *Politicians are mainly in politics for their own benefit and not for the benefit of the community.* Political interest: • Level of interest in politics. Voting: • Whether voted in 2017 General Election • Party voted for in 2017 General Election • Whether voted in 2016 EU Referendum • How voted in 2016 EU Referendum • Desired outcome from 2016 EU Referendum
Achievements and aspirations	Whether has achieved each of the below – or anticipates achieving each of the below by the age of 40: • Owning home, Having a good car, Earning a lot of money, Having a worthwhile job, Having children, Having a partner/being married, Being famous/having made name for oneself, Having a personal achievement in sport, the arts, travel etc.
Societal challenges	Main societal challenges facing Next Steps generation (top 3 from below): • Getting on the housing ladder, Finding a stable and secure job, Climate change and global warming, Having enough money to live comfortably, Having enough money in retirement, Political change and instability, Concern about war or terrorism, Leaving the EU, Social media and digitisation of everyday lives, Accessing good healthcare

6. Research Data

6.1 Licensing and data access

The Next Steps 2019 Web Survey data have been processed by CLS and supplied to the UK Data Service (UKDS). All users of the data need to be registered with the UKDS. Details of how to do this are available on the UK Data Service website. The 2019 Web Survey dataset is available from the UKDS as safeguarded data, which can be downloaded from the UKDS once the End User License (EUL) access conditions have been accepted by the user.

6.2 Datasets and data structure

The survey data can be found in the dataset **NS_2019_Web_Survey**.This dataset is structured in a flat (wide) format. i.e., where one record exists for each cohort member.

The weight variable W2019FINWT 'Weight: 2019 web survey final weight' is included in the dataset **NS9_2022_Longitudinal File**.

6.3 Data documentation

In addition to this User Guide, the following documentation accompanies the data deposited at the UKDS:

Table 3. Data documents

Name of the document	Content summary
Next Steps – 2019 Web Survey Questionnaire	This document provides the questions asked in the 2019 Web Survey

6.4 Identifiers

All datasets are primarily identified with the same research identifier (NSID) used for all Next Steps cohort data available at the UKDS.

6.5 Variable description

6.5.1 Variable order

The order in which variables appear in the datasets broadly follows the order of the questionnaire.

6.5.2 Variable names

Variable names in the dataset NS_2019_Web_Survey are prefixed with 'W2019' denoting the year of the web survey. Variable names for substantive questions have been chosen to reflect the content of the question and are documented in the online questionnaire. Table 4 below includes the dataset variable name and a description of the variable in the order of response where applicable.

Table 4. Variable names and descriptions

Variable Name	Variable Description
NSID	Cohort member identifier
W2019INTMTH	Date of interview: month
W2019INTYEAR	Date of interview: year
W2019OUTCOME	Whether interview fully or partially productive
W2019STAGE	Survey stage at which interview data collected
W2019SELFIMPROVE	Easier now to improve things for ourselves than it was for our parents
W2019HARDWORK	Britain today is a place where hard work is rewarded.
W2019POLITICIANS	Politicians mainly in politics for personal benefit and not for benefit of community
W2019POLITICS	Degree of interest in politics
W2019VOTEDGE2017	Whether voted in last General Election (June 2017)
W2019PARTYGE2017	Party voted for in last General Election (June 2017)
W2019VOTEDEU2016	Whether voted in the European Referendum (2016)
W2019CHOICEEU2016	Voting choice in the European Referendum (2016)
W2019BREXIT01	Rank of Brexit outcome - Remain in the EU
W2019BREXIT02	Rank of Brexit outcome - Leave with no withdrawal agreement
W2019BREXIT03	Rank of Brexit outcome - Leave with softer Brexit deal (Customs Union and/or Single Market)
W2019BREXIT04	Rank of Brexit outcome - Theresa May's withdrawal agreement and deal
W2019BREXIT05	Rank of Brexit outcome - Leave with different/new deal negotiated by Boris Johnson
W2019ACHIEVED01	Have achieved - Owning your own home
W2019ACHIEVED02	Have achieved - Having a good car
W2019ACHIEVED03	Have achieved - Earning a lot of money

Variable Name	Variable Description
W2019ACHIEVED04	Have achieved - Owning your own home
W2019ACHIEVED05	Have achieved - Having a good car
W2019ACHIEVED06	Have achieved - Earning a lot of money
W2019ACHIEVED07	Have achieved - Having a worthwhile job
W2019ACHIEVED08	Have achieved - Having children
W2019ACHIEVED09	Have achieved - Having a partner or being married
W2019ACHIEVE4001	Will have achieved by age 40 - Owning your own home
W2019ACHIEVE4002	Will have achieved by age 40 - Having a good car
W2019ACHIEVE4003	Will have achieved by age 40 - Earning a lot of money
W2019ACHIEVE4004	Will have achieved by age 40 - Having a worthwhile job
W2019ACHIEVE4005	Will have achieved by age 40 - Having children
W2019ACHIEVE4006	Will have achieved by age 40 - Having a partner or being married
W2019ACHIEVE4007	Will have achieved by age 40 - Being famous or making a name for yourself
W2019ACHIEVE4008	Will have achieved by age 40 - Having a personal achievement in sport/arts/ travel
W2019ACHIEVE4009	Will have achieved by age 40 - None of these
W2019CHALLENGE01	Main societal challenges facing your generation - Getting on the housing ladder
W2019CHALLENGE02	Main societal challenges facing your generation - Finding a stable and secure job
W2019CHALLENGE03	Main societal challenges facing your generation - Climate change and global warming
W2019CHALLENGE04	Main societal challenges facing your generation - Having enough money to live comfortably
W2019CHALLENGE05	Main societal challenges facing your generation - Having enough money in retirement

Variable Name	Variable Description
W2019CHALLENGE06	Main societal challenges facing your generation - Political change and instability
W2019CHALLENGE07	Main societal challenges facing your generation - Concern about war or terrorism
W2019CHALLENGE08	Main societal challenges facing your generation - Leaving the EU
W2019CHALLENGE09	Main societal challenges facing your generation - Social media and digitisation of everyday lives
W2019CHALLENGE10	Main societal challenges facing your generation - Accessing good healthcare
W2019CHALLENGE11	Main societal challenges facing your generation - None of these

6.5.3 Variable labels

Variable labels are based on the wording that can be found in the questionnaire documentation. Where necessary, labels have been modified to ensure they are comprehensible and accurate.

For multi-coded variables a suffix has been used to identify each response option. 01, 02, 03… have been used to denote the 1st, 2nd, 3rd, ...options respectively.

6.5.4 Value labels

The value labels for valid responses are based on the question responses used in the Computer-Assisted Interviewing (CAI) program as documented in the questionnaire documentation. Value labels have been individually reviewed and amended, where necessary.

6.6 Missing values

Where applicable, non-responses have been recoded to -9 = ‘Not answered’ where a response was expected. Otherwise -1 = ‘Not applicable’ has been applied to responses to questions with routing conditions and partially productive responses.

Additionally, the DATA_AVAILABILITY variable (included in the dataset **NS9_2022_Longitudinal File**) flags cohort members for whom no data is available from any sweep of the study, following requests to delete their survey data.

7. Response patterns and weights

7.1 Response patterns

The issued sample for first sweep of Next Steps was approximately 21,000 young people. A total of 15,770 households were interviewed in that initial wave, representing 74 per cent of the target sample, with both young people and their parents in scope to be interviewed. At Sweep 4, 352 ethnic boost interviews were added (Black Caribbean and Black African pupils, selected from the original [non-responding] school sample), taking the total number of cohort members who had taken part in the study up to 16,122.

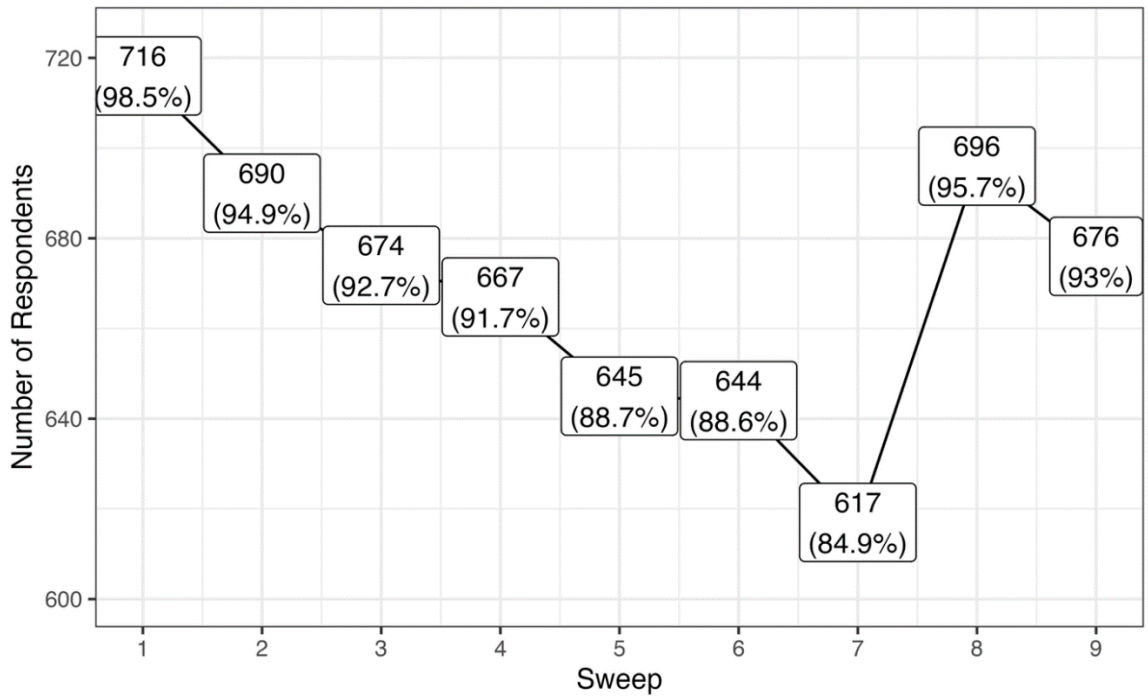


Figure 1: Proportion of web survey (substantive version) respondents responding to a given mainstage sweep

Note: Although the ethnic minority boost sample was added in Sweep 4, respondents to this sweep has been included in the denominator, which is why the proportion of web survey respondents who took part in Sweep 1 is <100%.

As discussed, 727 individuals responded to the version of the web-survey which included substantive questions, 4.5% of the total Next Steps sample. Figure 1 shows the proportion of these respondents who participated in a given mainstage sweep. Note, 576 (79.2%) web-respondents participated in each sweep up to Sweep 8. 543 (74.7%) also participated at Sweep 9.

7.2 Predicting response and weights

7.2.1 Introduction

The low proportion of total Next Steps participants who completed the version of the web survey which includes substantive questions reflects both survey design (only a subset of cohort members were invited to complete this version of the survey) and non-response to the study.

Non-response is common in longitudinal surveys. Missing values mean less efficient estimates because of the reduced size of the analysis sample but also introduce the potential for bias since respondents are often systematically different from non-respondents. To support researchers in producing robust analysis, we have developed comprehensive advice on how to deal with missing data that can be accessed on the CLS website here. The approaches we recommend to researchers capitalise on the rich data cohort members provided over the years before their nonresponse. These approaches include well known methods such as Multiple Imputation (MI), Inverse Probability Weighting (IPW), and Full Information Maximum Likelihood (FIML).

To correct for non-response in the Next Steps 2019 Web Survey, non-response weights are provided with the data, so that IPW analysis can be undertaken, either in isolation or in combination with MI or FIML. This section of the User Guide describes the derivation and implementation of these non-response weights. The weights were created and documented by Liam Wright and Richard Silverwood and closely followed the procedures used to derive non-response weights for CLS' COVID-19 Surveys. You can access the user guide on the CLS website here.

7.2.2 Target population and response

For the purposes of weighting, we defined the target population as any individual in Year 9 in England in 2003/04, reflecting the population from which the Next Steps sample was drawn. The response rate, based on this target population was 4.5% (n = 727 / 16,122)

7.2.3 Derivation of non-response weights

The derivation of the non-response rate proceeded as follows:

1. Within the sample corresponding to the target population, model survey response conditional on a set of covariates using logistic regression. The selection of covariates was informed from results of the CLS Missing Data Strategy which can be accessed on the CLS website here.
2. Among respondents to the 2019 Web Survey (substantive version), calculate the probability of response implied by each model.
3. Calculate the non-response weight as the inverse of the probability of response.
4. Examine the distribution of the non-response weight to decide whether truncation may be desirable, applying truncation, if so.
5. Given Next Steps used a complex sampling design, multiply the web survey non-response weights with initial survey weights to create a 'combined weight' to recover the initial target population.[1]
6. Finally, calibrate the combined non-response weight so that it summed to the number of web survey respondents (n = 727).

The variables included in the response model in are listed in Table 5. We used the same set of variables as used to derive weights for the Next Steps in CLS's COVID-19 Surveys, with the exception that we also included variables for (a) whether the participant was a boost respondent and (b) number of COVID-19 sweeps a

[1] For cohort members entering in Sweep 1, this was the non-response weight supplied with data for that sweep. For cohort members entering in Sweep 4 (i.e., the boost sample), this was the non-response weight supplied at that sweep. Both of these weights were designed to make the (weighted) Next Steps sample at those sweeps representative of the target population from which cohort members were drawn.

participant responded to (continuous; 0-3). Additionally, a small number of variables were excluded in models to derive the present weights due to data unavailability. The code used to clean the data and create the weights can be viewed here.

Table 5. Variables included in response model

Variable	Description
Sex	Categories: Male; Female
Ethnic Group	Categories: White; Mixed; Indian; Pakistani; Bangladeshi; Black Caribbean; Black African; Other
Parental Socio-Economic Class @ Sweep 1	Categories: Higher Managerial and professional occupations; Lower managerial and professional occupations; Intermediate occupations; Small employers and own account workers; Lower supervisory and technical occupations; Semi-routine occupations; Routine occupations; Never worked/long term unemployed
GHQ Caseness @ Sweep 2	Mean: 1.69; Min: 0; Max: 12
GHQ Caseness @ Sweep 8	Mean: 2.3; Min: 0; Max: 12
Voted in 2010 General Election	Categories: Yes; No
Highest NVQ @ Sweep 8	Categories: NVQ Level 1; NVQ Level 2; NVQ Level 3; NVQ Level 4; NVQ Level 5; None/Other
Economic Activity @ Sweep 8	Categories: Employed; Inactive; Unemployed; Education/Training; Other
Partnered @ Sweep 8	Categories: No; Yes (Non-Cohabiting); Yes (Cohabiting)
BMI @ Sweep 8	Mean: 25.23; Min: 13.27; Max: 67.15
Self-Rated Health @ Sweep 8	Categories: Excellent; Very good; Good; Fair; Poor
Smoking Status @ Sweep 8	Categories: Never; Former; Occasional; Daily
Social Support @ Sweep 8	Categories: Not at all; A little; Somewhat; A great deal
Log Income @ Sweep 8	Mean: 5.69; Min: 2.87; Max: 7.34
Boost Respondent	Categories: Main; Boost
# of Mainstage Sweeps Participated In	Mean: 6.04; Min: 1; Max: 9
# of COVID-19 Sweeps Participated In	Mean: 0.61; Min: 0; Max: 3

Missing values in the above variables were handled using multiple imputation (MI). The imputation models included these variables plus web survey response. Forty imputed datasets were produced and response models were fitted in each imputed dataset separately. Pooled model results are reported in Appendix 1. From these models, the probability of survey response was predicted for each respondent, with the non-response weight calculated as the inverse of the response probability. These

weights were then averaged across imputed models to get a single weight per individual.

We experimented with different levels of weight truncation. However, given there was a low overall response rate and there were very large differences in response probabilities between respondents, we opted against truncation for this data release. As a results, using the weights decreases the precision of estimates substantially (see next section). Data users should consider whether to use truncation in their own analyses using these weights.

7.2.4 Weight effectiveness

To examine the effectiveness of the derived non-response weight in restoring sample representativeness we compared sample descriptives in the full sample and in the (weighted and unweighted) sample of 2019 web survey respondents. We did this for variables that were used in the response model as well as two 'unseen' variables (risk taking behaviour and attitude to school, both measured at Sweep 1) which were not used in response models and are therefore not balanced mechanistically. A good performing weight is indicated by having a similar sample mean to the full sample. The results for unseen variables are presented in Figure 2 (results from 'seen' variables are presented in Appendix 1). We observed that there is considerable bias when using the unweighted web survey respondents but this is often reduced, though not completely, when applying the non-response weights. Further, given the high degree of variability in the weights, precision when using the weights (as indicated by the width of confidence intervals) is much weaker.

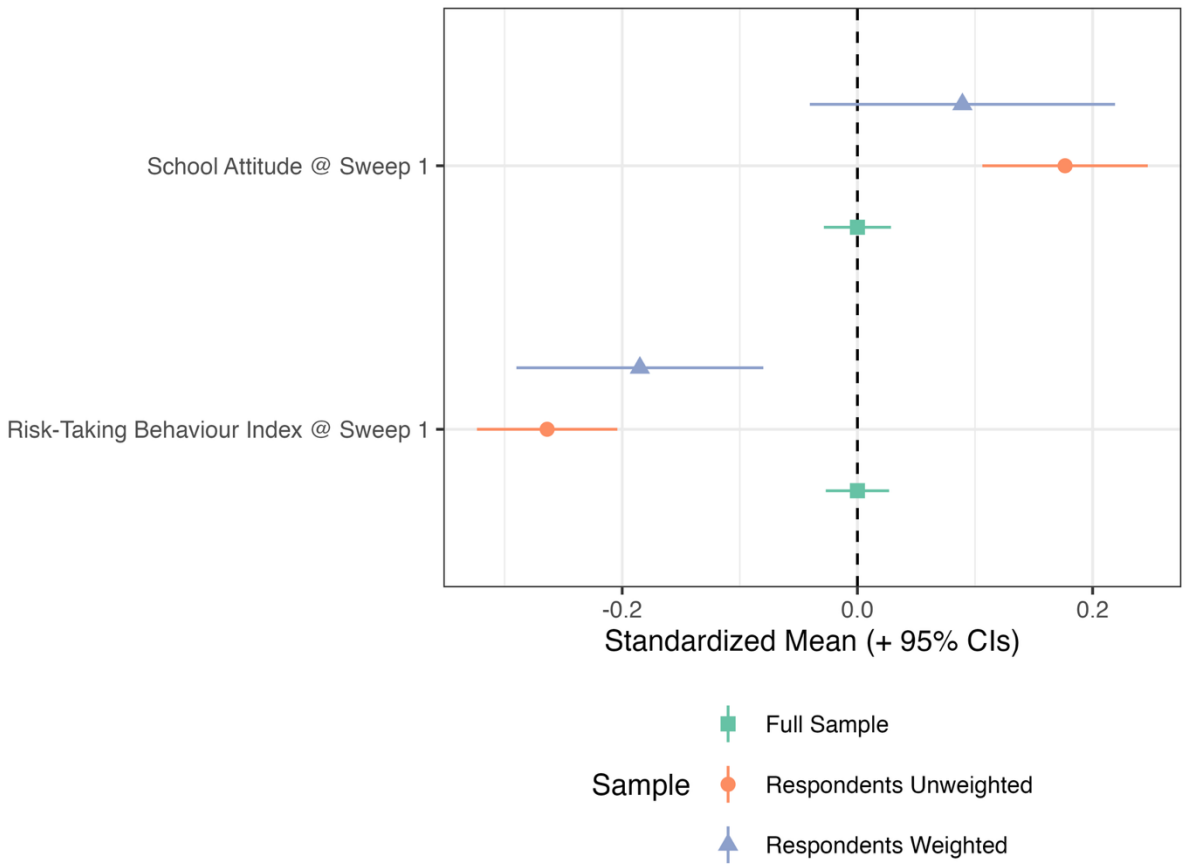


Figure 2: Means for two variables that were not used to create non-response weights in the target population and in weighted and unweighted 2019 Web Survey responding sample. Variables standardised (mean = 0, SD = 1) using the target population mean and SD

7.2.5 Implementation of non-response weights

Next Steps used a complex sampling design to recruit cohort members: cohort members were recruited from stratified samples of schools (approximately 30 students per school; see the LSYPE User Guide to Datasets: Wave 1 to Wave 7 for more detail). Analyses of Next Steps data should account for this complex sampling design by specifying primary sampling unit (PSU; SAMPPSU) and stratum (SAMPSTRATUM). In Stata, this can be achieved by first using `svyset` to specify the survey design, and then conducting analyses using the `svy` prefix, e.g.:

```
svyset SAMPPSU [pweight=W2019FINWT], strata(SAMPSTRATUM)
svy: proportion W2019VOTEDGE2017, citype(agresti)
```

In R, the `survey` package can be used to specify complex survey design, e.g.:

```
library(survey)
ns_svy <- svydesign(id = ~ SAMPPSU, strata = ~ SAMPSTRATUM, weights
= ~ W2019FINWT, data = ns_ws2019)
svytable(~ W2019VOTEDGE2017, ns_svy)
```

See relevant Stata help files (StataCorp 2023) and `survey` package documentation (Lumley 2011) for more information on using survey data in Stata and R. (Users with experience of the `tidyverse` may also want to use the `srvyr` package, which provides similar functionality to `survey` but within a `tidy` framework.)

Given the low numbers participating at the 2019 Web Survey, there may only be one PSU in a stratum, in which case users may consider allocating these observations to the modal stratum. Also, for some commands, users may find that certain functionalities have not been adapted for use with complex survey data. In this case, users may consider using survey weights without declaring complex survey design and noting this in the write up of their analyses. Users can also check whether accounting for complex survey design makes much difference in their particular analysis, by running analyses declaring and not declaring the complex design; as participants have now long left secondary school, and there are few observations per stratum, observations at the 2019 Web Survey may be independent (or at least less dependent) at the PSU level.

8. References

Brown, M., Fitzsimons, E., 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 4)*. London: 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

Department for Education. (2011). *LSYPE User Guide to the Datasets: Wave 1 to Wave 7*. https://doc.ukdataservice.ac.uk/doc/5545/mrdoc/pdf/lsype_user_guide_wave_1_to_wave_7.pdf

Lumley, T. (2011). *Complex surveys: a guide to analysis using R*. John Wiley & Sons.

StataCorp. 2023. *Stata 18 Survey Data Reference Manual*. College Station, TX: Stata Press.

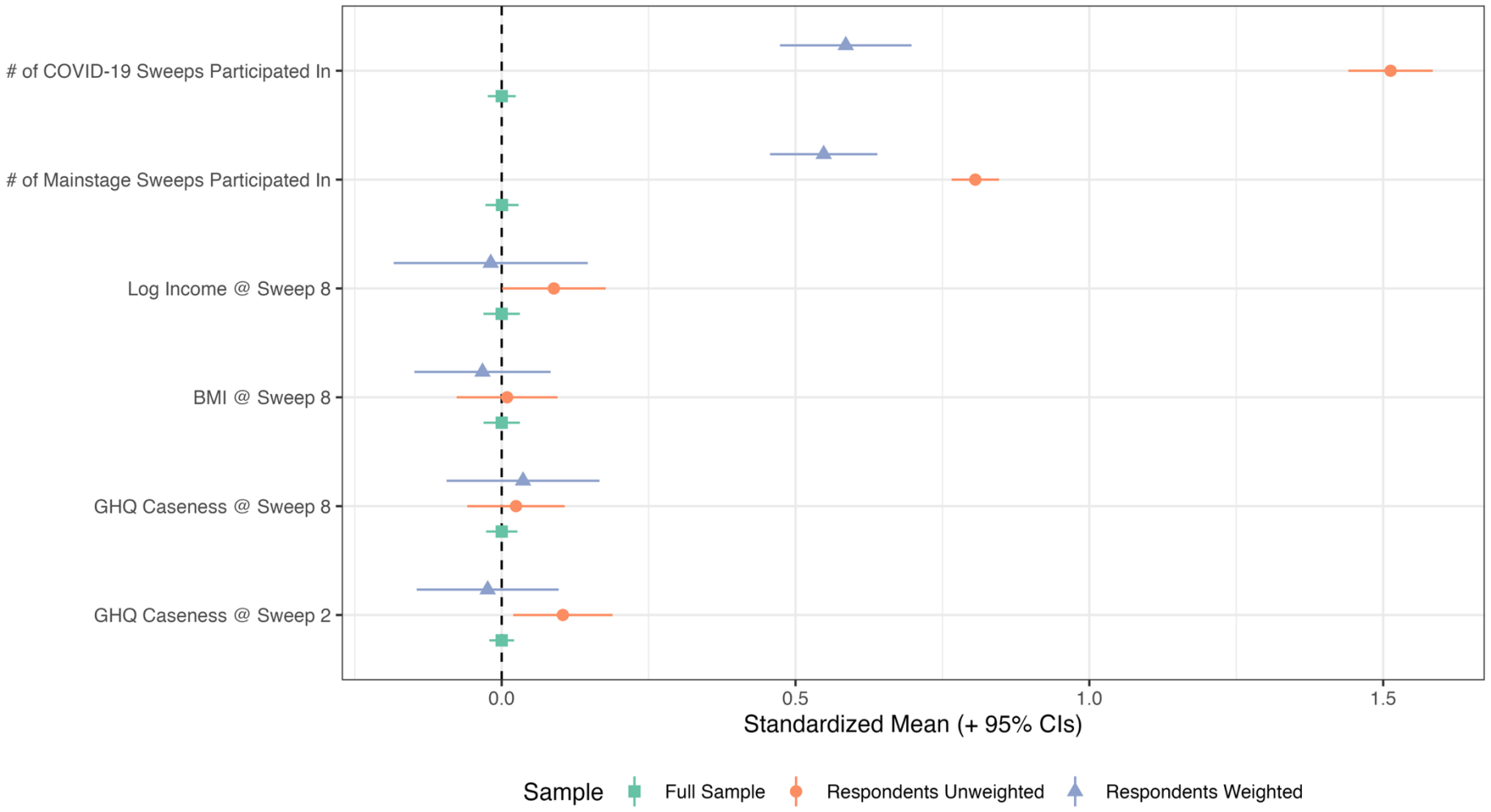
Appendix 1: Response model results

Appendix Table 1: Results from 2019 Web Survey Response Model

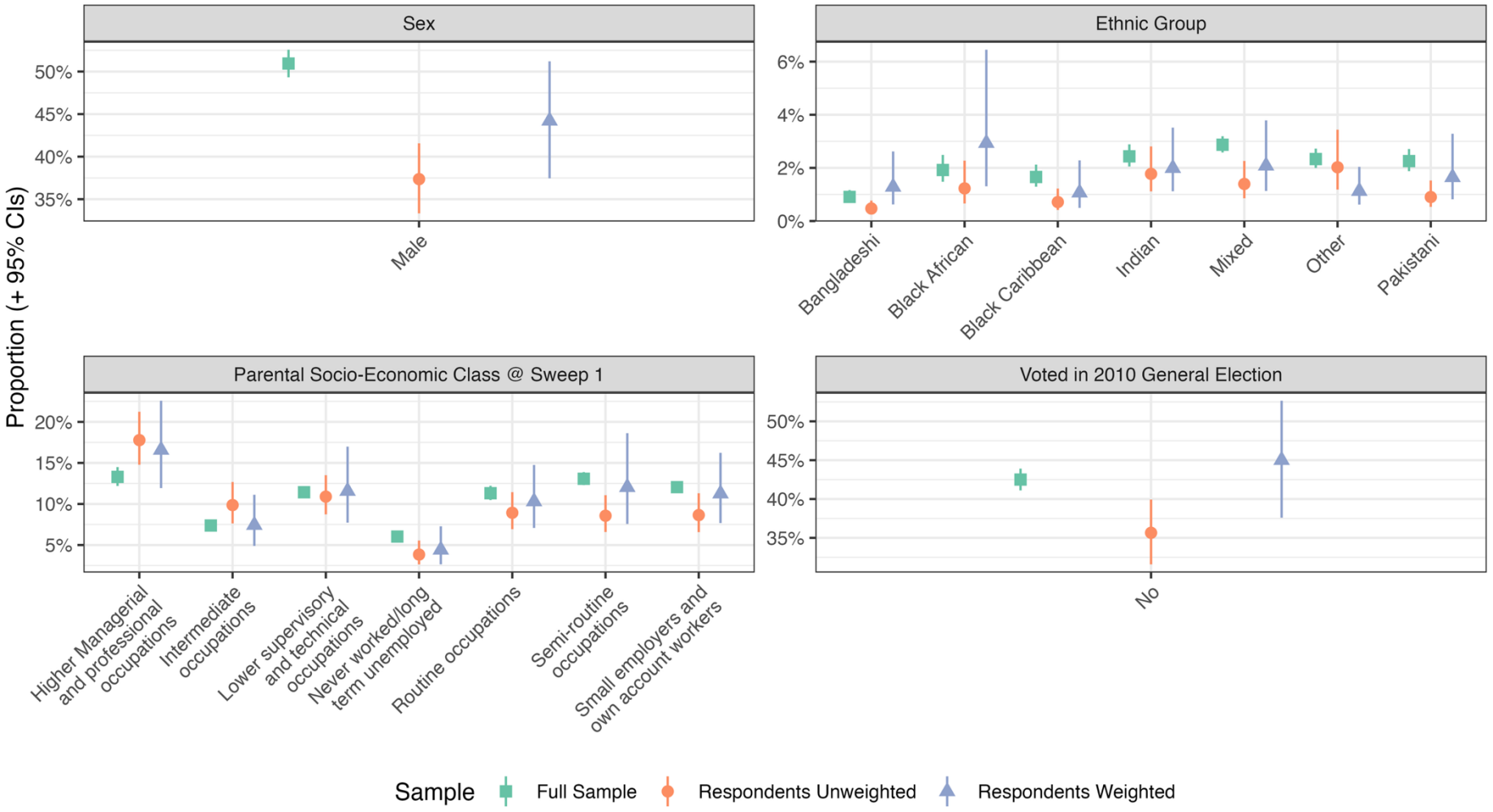
	Variable	OR (+ 95% CIs)
Sex	Female	1.01 (0.84, 1.2)
Ethnic Group	Mixed	0.73 (0.48, 1.12)
	Indian	0.64 (0.43, 0.94)
	Pakistani	0.56 (0.33, 0.93)
	Bangladeshi	0.74 (0.45, 1.22)
	Black Caribbean	0.81 (0.47, 1.39)
	Black African	1.19 (0.71, 1.99)
	Other	1.32 (0.81, 2.13)
Parental Socio-Economic Class @ Sweep 1	Lower managerial and professional occupations	1.24 (0.95, 1.6)
	Intermediate occupations	1.39 (0.98, 1.96)
	Small employers and own account workers	0.93 (0.67, 1.31)
	Lower supervisory and technical occupations	1.4 (1.01, 1.96)
	Semi-routine occupations	1.04 (0.73, 1.46)
	Routine occupations	1.31 (0.92, 1.86)
	Never worked/long term unemployed	1.6 (1.04, 2.45)
GHQ Caseness @ Sweep 2		1.02 (0.98, 1.05)
GHQ Caseness @ Sweep 8		1.03 (0.99, 1.06)

	Variable	OR (+ 95% CIs)
Voted in 2010 General Election	No	1.05 (0.86, 1.28)
Highest NVQ @ Sweep 8	NVQ Level 2	1.79 (1.12, 2.86)
	NVQ Level 3	1.87 (1.16, 3.01)
	NVQ Level 4	2.05 (1.28, 3.29)
	NVQ Level 5	2.3 (1.41, 3.74)
	None/Other	1.77 (0.94, 3.34)
Economic Activity @ Sweep 8	Other	0.76 (0.35, 1.64)
	Inactive	0.94 (0.61, 1.44)
	Unemployed	0.75 (0.46, 1.21)
	Education/Training	0.61 (0.4, 0.94)
Partnered @ Sweep 8	Yes (Non-Cohabiting)	1.08 (0.87, 1.35)
	Yes (Cohabiting)	1 (0.79, 1.25)
BMI @ Sweep 8		1 (0.99, 1.02)
Self-Rated Health @ Sweep 8	Very good	1.09 (0.89, 1.35)
	Good	0.95 (0.73, 1.22)
	Fair	0.53 (0.35, 0.82)
	Poor	0.99 (0.53, 1.82)
Smoking Status @ Sweep 8	Former	1.07 (0.84, 1.35)
	Occasional	1.12 (0.85, 1.47)
	Daily	0.78 (0.58, 1.05)
Social Support @ Sweep 8	A little	0.91 (0.46, 1.81)
	Somewhat	0.94 (0.48, 1.82)
	A great deal	1.05 (0.55, 2.02)

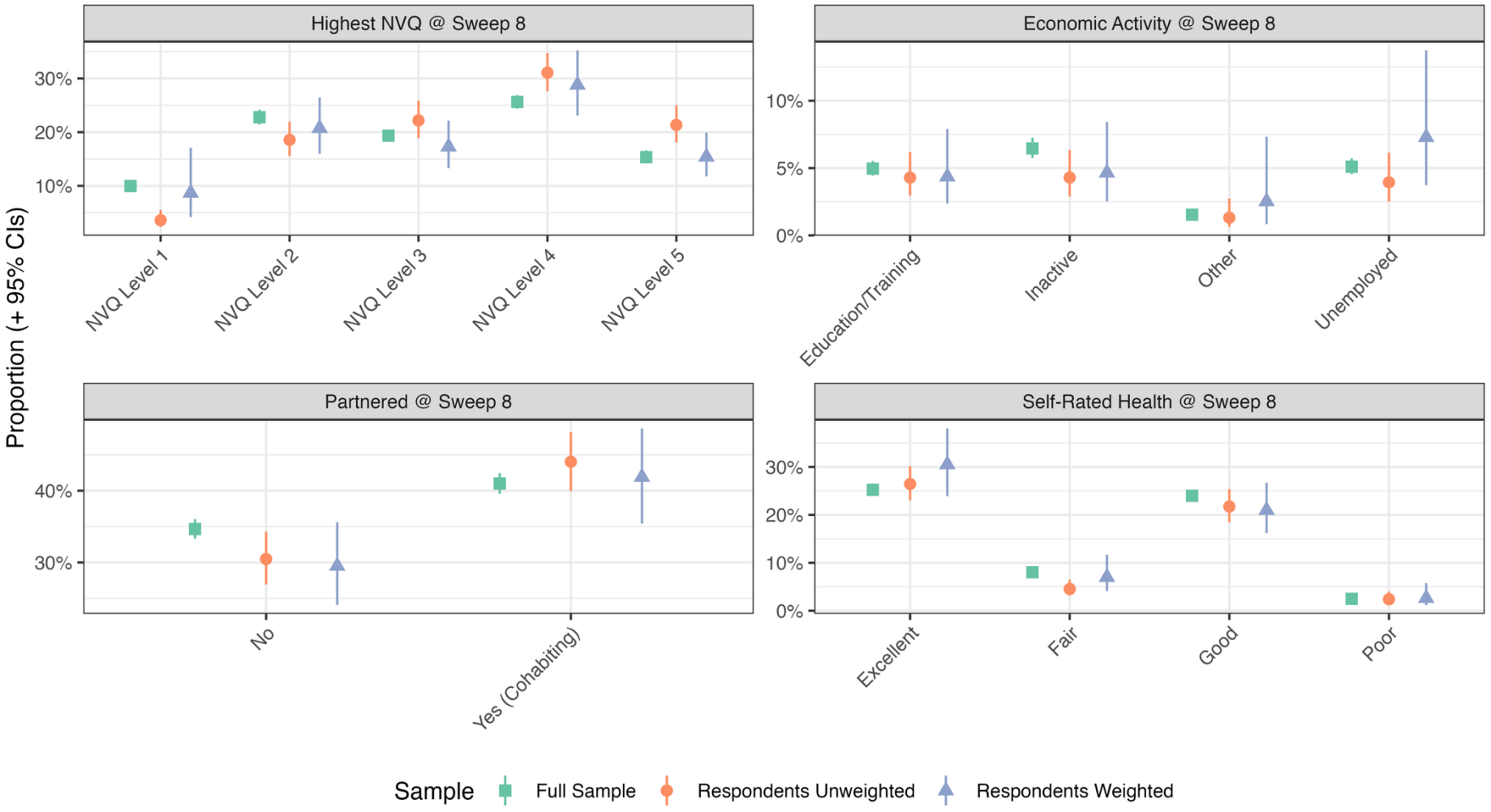
	Variable	OR (+ 95% CIs)
Log Income @ Sweep 8		0.94 (0.83, 1.06)
Boost Respondent	Boost	1.31 (0.63, 2.75)
# of Mainstage Sweeps Participated In		1.23 (1.15, 1.31)
# of COVID-19 Sweeps Participated In		2.85 (2.61, 3.1)
Intercept		0 (0, 0.01)



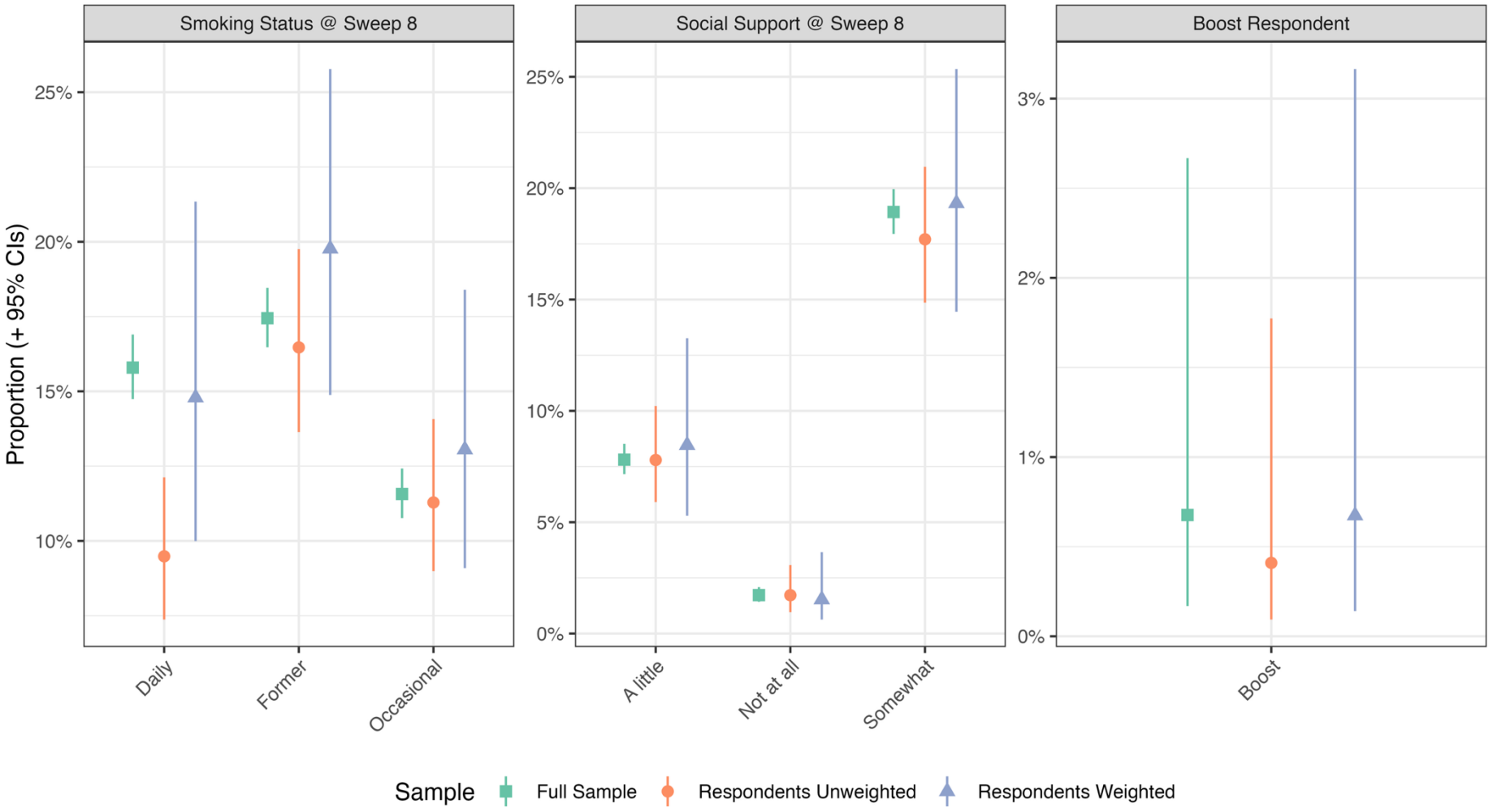
Appendix Figure 1: Means for continuous variables appearing in the response model – full sample and (weighted and unweighted) 2019 Web Survey respondents



Appendix Figure 2: Proportions for categorical variables appearing in the response model – full sample and (weighted and unweighted) 2019 Web Survey respondents (See also next two figures)



Appendix Figure 3: Proportions for categorical variables appearing in the response model – full sample and (weighted and unweighted) 2019 Web Survey respondents (See also next and previous figures)



Appendix Figure 4: Proportions for categorical variables appearing in the response model – full sample and (weighted and unweighted) 2019 Web Survey respondents (See also previous two figures)