

SUPPLEMENTARY MATERIAL

How to mitigate selection bias in COVID-19 surveys: evidence from five national cohorts

Contents

Methods S1: Sensitivity analyses for social class.....	2
Methods S2: Sensitivity analyses for childhood cognitive ability.	3
Table S1: Coding of all variables used for the derivation of weights and in the multiple imputation models.....	4
Fig. S1: Restoring representativeness for parental social class in COVID wave 2.....	9
Fig. S2: Restoring representativeness for parental social class in COVID Wave 3.	10
Fig. S3: Restoring representativeness for childhood cognitive ability in COVID wave 2.	11
Fig. S4: Restoring representativeness for childhood cognitive ability in COVID wave 3.	12
Fig. S5: Sensitivity analyses for parental social class.	13
Fig. S6: Sensitivity analyses for childhood cognitive ability.	14

Methods S1: Sensitivity analyses for social class.

We conducted sensitivity analyses for Next Steps and MCS which included both the original non-response weights as described in the COVID-19 Survey User Guide, as well as the non-response weights which we created specifically for the use of demonstrating how sample representativeness can be restored for parental social class. The original non-response weights for Next Steps and MCS included the exact same variable for parental social class which was used for the demonstration and thus had an unfair advantage relative to the multiple imputation models. However, as we can see in Fig. S5, the estimates based on the original non-response weights (blue bars) and the newly created non-response weights (orange bars) did not differ substantially. These were the newly created non-response weights used in the main paper.

Methods S2: Sensitivity analyses for childhood cognitive ability.

We conducted sensitivity analyses for NSHD, NCDS and MCS which included both the original non-response weights as described in the COVID-19 Survey User Guide, as well as the non-response weights which we created specifically for the use of demonstrating how sample representativeness can be restored for childhood cognitive ability. The original non-response weights for NSHD, NCDS and MCS included the exact same variable for childhood cognitive ability which was used for the demonstration and thus had an unfair advantage relative to the multiple imputation models. As we can see in Fig. S6, the estimates based on the original non-response weights (blue bars) and the newly created non-response weights (orange bars) did differ, showing that estimates are not as good when no measure of childhood cognitive ability was included in the derivation of weights. We furthermore created estimates based on newly created non-response weights (purple bars) which did not include the exact same childhood cognitive ability variable as in the demonstration but did include measures of childhood cognitive ability at other time points (the same measures which were included in the equivalent multiple imputation models). With the inclusion of other cognitive measures in the derivation of weights, we were again able to restore representativeness on childhood cognitive ability to a large extent. These were the newly created non-response weights (purple bars) used in the main paper.

Similarly, we ran sensitivity analyses which included estimates based on the original multiple imputation models (which included additional measures of childhood cognitive ability from time points other than the one used in the demonstration), as well as estimates based on multiple imputation models which only included the same variables that were used in the derivation of the original non-response weights from the COVID-19 survey user guide (no additional measures of cognitive ability from other time points). As expected, the estimates from the new multiple imputation models (pink bars) did not perform as well as the imputation models which allowed for inclusion of additional measures of cognitive ability from other time points.

In conclusion, when no measures of cognitive ability were included in the derivation of weights (orange bars) or multiple imputation models (pink bars), we were not equally successful in restoring sample representativeness for childhood cognitive ability as compared to when they were included (purple and green bars). For the results in the main paper, we use non-response weights and multiple imputation models that do include these additional measures, although not the exact same childhood cognitive ability variables as were used in the demonstration itself (to avoid an unfair advantage).

Table S1: Coding of all variables used for the derivation of weights and in the multiple imputation models.

NSHD	NCDS	BCS	Next Steps	MCS
Sex				
Birth 0 = Male / 1 = Female	Birth 0 = Male / 1 = Female	Birth 0 = Male / 1 = Female	Age 14 0 = Male / 1 = Female	9 months 0 = Male / 1 = Female
Ethnicity				
-	-	-	Age 14 1 = White / 2 = Indian, Pakistani, Bangladeshi / 3 = Black Caribbean, Black African / 4 = Mixed, Other	9 months Age 3 1 = White / 2 = Indian, Pakistani, Bangladeshi, Other Asian / 3 = Black Caribbean, Black African, Other Black / 4 = Mixed, Other
Parental social class				
Age 4 1 = Professional, intermediate / 2 = Skilled / 3 = Partly-, unskilled Age 11 1 = Professional, intermediate / 2 = Skilled / 3 = Partly-, unskilled	Birth 1 = Professional, managerial / 2 = Intermediate / 3 = Partly-, unskilled Age 11 1 = Professional, managerial / 2 = Intermediate / 3 = Partly-, unskilled	Birth 1 = Professional, managerial / 2 = Intermediate / 3 = Partly-, unskilled Age 10 1 = Professional, intermediate / 2 = Skilled / 3 = Partly-, unskilled	Age 14 1 = Managerial / 2 = Intermediate / 3 = Routine, semi-routine / 4 = Never worked	9 months 1 = Managerial / 2 = Intermediate / 3 = Routine, semi-routine Age 11 1 = Managerial / 2 = Intermediate / 3 = Routine, semi-routine
Number of rooms at home/persons per room				
Birth Household crowding	Birth Persons per room	Birth Number of rooms at accommodation	-	9 months Number of rooms at home
Cognitive ability				
Age 8	Age 7	Age 10	-	Age 5

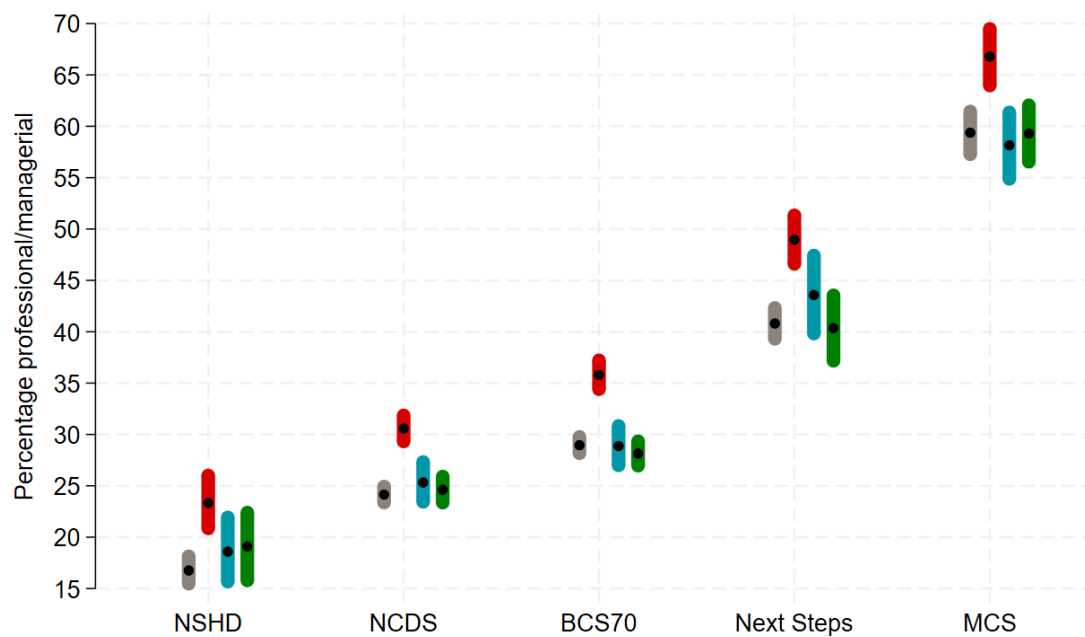
Standardised score (based on Reading Comprehension, Word Reading, Vocabulary and Picture Intelligence) Age 11 Standardised score (based on General Ability Test, Word Reading, Vocabulary and Arithmetic Test)	Standardised score (based on Southgate Group Reading Test, Copying Designs Test, Human Figure Drawing, Problem Arithmetic Test) Age 11 Total scores of General Ability Test, Reading Comprehension Test, Mathematics Test, Copying Design Tests	Standardised score (based on British Ability Scales, Edinburgh Reading Test, Friendly Maths Test, Spelling Dictation Task, Pictorial Language Comprehension Test) Age 5 Standardised score (based on English Picture Vocabulary Test, Human Figure Drawing Test, Copying Designs Test)		Standardised score based on British Ability Scales (BAS II Naming Vocabulary, BAS II Pattern Construction, BAS II Picture Similarities) Age 7 Standardised score for BAS II Word Reading Standardised score for NFER Progress in Math
Early life mental health				
Age 13 & 15 Externalising score 0 = absent / 1 = mild / 2 = severe Internalising score 0 = absent / 1 = mild / 2 = severe	Age 16 Externalising symptoms standardised score; Internalising symptoms standardised score	Age 16 Malaise sum score	Age 15 GHQ sum score	Age 11 SDQ Total Difficulties Score
Voting				
Age 26 Will you vote in the next General Election? 0 = No / 1 = Yes	Age 42 Did you vote in the last General Election? 0 = No / 1 = Yes	Age 42 Voted in General Election 2010 0 = Didn't vote / 1 = Voted	Age 20 Whether voted in 2010 General Election 0 = Yes / 1 = No	NA
Membership in organisations				
Age 43 0 = none / 1 = 1 / 2 = 2+	Age 42 Membership in organisations 0 = no / 1 = yes Membership in unions	Age 42 Member in organisations 0 = No organisations / 1 = 1 organisation / 2 = 2+ organisations	Age 26 Meetings for local groups/voluntary organisations 0 = Yes / 1 = No	Age 14 Youth clubs / Scouts / Etc 0 = At least once a month / 1 = Less than once a month

	0 = no / 1 = yes			
Internet access prior to web survey				
Age 69 How often do you take part in online social networks? 0 = Never / 1 = Not never	Age 50 Whether access to internet for reasons other than work? 0 = yes / 1 = no	Age 46 Time spent on internet 1 = None, little / 2 = Medium / 3 = Lots	Age 26 Time spent on social networking website 0 = None / 1 = Little / 2 = Lot	Age 14 Use internet at home 1 = Little, none / 2 = Medium / 3 = Lots
Consent for biomarkers				
Age 60-64 0 = neither, one / 1 = both	Age 44 0 = yes / 1 = no	Age 46 0 = No to one, both / 1 = Yes to both	-	-
Consent for linkages				
Age 60-64 0 = refused / 1 = yes	-	-	Age 26 0 = None / 1 = Some / 2 = All	-
Educational qualifications				
Age 26 0 = None attempted / 1 = Up to GCE 'O' Level / 2 = GCE 'A' Level / 3 = First or higher degree	Age 42 0 = None / 1 = NQV Level 1-3 / 2 = NVQ Level 4-5	Age 42 0 = None / 1 = NQV Level 1-3 / 2 = NVQ Level 4-5	Age 26 0 = None / 1 = NQV Level 1-3 / 2 = NVQ Level 4-5	9 months 0 = None / 1 = NQV Level 1-3 / 2 = NVQ Level 4-5
Economic activity				
Age 60-64 1 = Still in main occupation / 2 = Retired but still earning / 3 = Fully retired, Unemployed, Housewife	Age 50 0 = Currently employed / 1 = Not currently employed	Age 46 0 = Currently employed / 1 = Not currently employed	Age 26 0 = Currently employed / 1 = Not currently employed	Age 14 0 = Currently employed / 1 = Not currently employed
Partnership status				
Age 69 1 = Single & never married / 2 = Married / 3 = Separated, divorced, widowed	Age 50 1 = Single & never married / 2 = Married, civil partner / 3 = Separated, divorced, widowed	Age 46 1 = Never married, in CP / 2 = Married, CP / 3 = Separated, divorced, widowed	Age 26 0 = None / 1 = Spouse, civil partner / 2 = Cohabiting partner	Age 14 (Parental partnership status) 1 = None / 2 = Spouse, civil partner / 3 = Separated, divorced, widowed

Psychological distress				
Age 69 GHQ sum score	Age 50 Malaise sum score	Age 46 Malaise sum score	Age 26 GHQ sum score	Age 14 SMFQ sum score
BMI				
Age 69	Age 50	Age 46	Age 26	Age 11
Self-rated health				
Age 69 1 = Excellent, very good / 2 = Good / 3 = Fair, poor	Age 50 1 = Excellent, very good / 2 = Good / 3 = Fair, poor	Age 46 1 = Excellent, very good / 2 = Good / 3 = Fair, poor	Age 26 1 = Excellent, very good / 2 = Good / 3 = Fair, poor	Age 14 1 = Excellent, very good / 2 = Good / 3 = Fair, poor
Smoking status				
Age 69 1 = Current Smoker / 2 = Ex-smoker / 3 = Never smoked	Age 50 1 = Never / 2 = Former / 3 = Current	Age 46 1 = Never / 2 = Former / 3 = Current	Age 26 1 = Never / 2 = Former / 3 = Current	Age 14 0 = Never smoked / 1 = Current, former, tried
Maternal mental health				
-	-	-	-	9 months Malaise sum score
Social capital/social support				
Age 69 Frequency of meeting friends or relatives 1 = Never, almost never / 2 = Fairly frequently / 3 = Very frequently	Age 50 Frequency of meeting friends or relatives 1 = Never / 2 = Fairly frequently / 3 = Very frequently Has people to listen to 1 = A little, not at all / 2 = Somewhat / 3 = A great deal Most people can be trusted 1 = Most people can be trusted / 2 = Can't be too careful / 3 = Other/depends	Age 46 Meeting family and friends 1 = Never, rarely / 2 = Fairly frequently / 3 = Very frequently People around would be willing to listen 1 = A little, not at all / 2 = Somewhat / 3 = A great deal	Age 26 Meeting family and friends 1 = Never, rarely / 2 = Fairly frequently / 3 = Very frequently People around would be willing to listen 1 = A little, not at all / 2 = Somewhat / 3 = A great deal Trust scale (continuous, 0 - 10)	Age 14 Family, friends who help me feel safe, happy 0 = Very true / 1 = partly true, not true at all Someone I trust 0 = Very true / 1 = partly true, not true at all No one I feel close to 0 = Very, partly true / 1 = not true at all

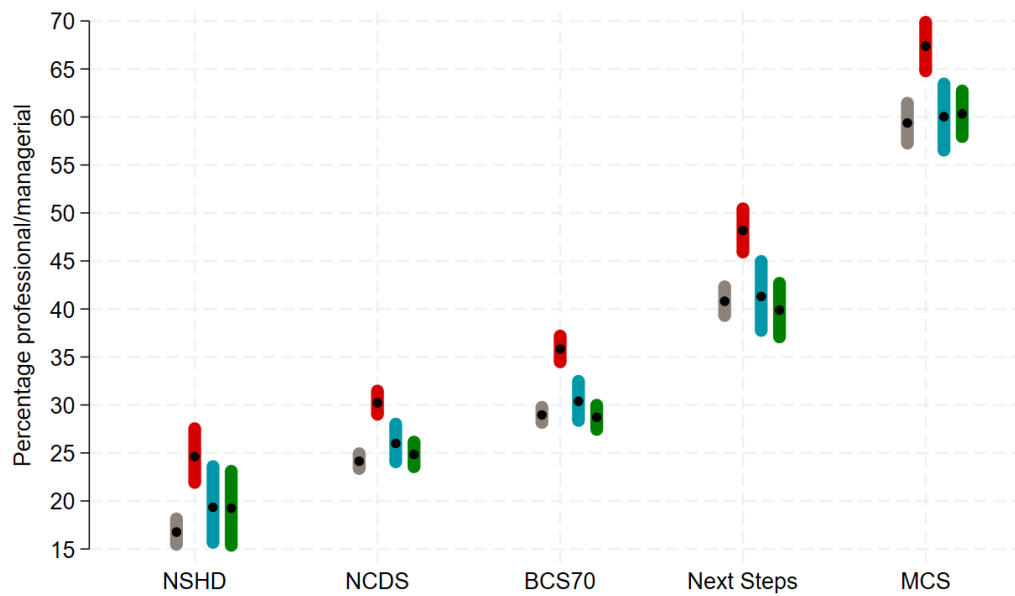
Income				
Age 69 Quantiles of total income	Age 55 Quantiles of total income	Age 42 Quantiles of total income	Age 26 Quantiles of total income	Age 14 Quantiles of total income
Number of non-responses across all previous sweeps				
Birth – age 69	Birth – age 55	Birth – age 42	Age 14 – age 26	9 months – age 14
Response to COVID-19 Wave 1 survey ^a				
Age 74	Age 62	Age 50	Age 30	Age 19
Response to COVID-19 Wave 2 survey ^e				
-	Age 62	Age 50	Age 30	Age 19

Fig. S1: Restoring representativeness for parental social class in COVID wave 2.



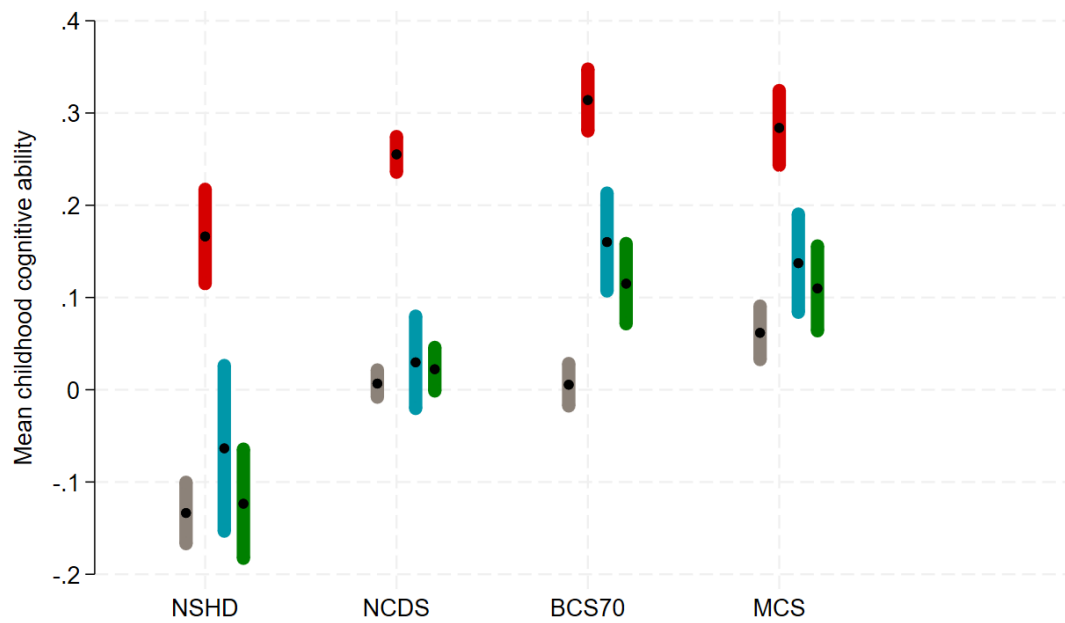
Percentage of highest social class (professional/managerial) in each cohort under different estimation approaches to account for non-response in the COVID-19 Wave 2 survey. Grey: using observed baseline data from the whole cohort; red: using observed baseline data from COVID-19 Wave 2 survey respondents only – unweighted (NCDS and BCS70) or using design weight only (NSHD, Next Steps and MCS); blue: using observed baseline data from COVID-19 Wave 2 survey respondents only – weighted using non-response weights (in addition to design weights as appropriate); green: using multiple imputation (plus design weights where appropriate). NSHD: National Survey of Health and Development; NCDS: 1958 National Child Development Study; BCS70: 1970 British Cohort Study; MCS: Millennium Cohort Study.

Fig. S2: Restoring representativeness for parental social class in COVID Wave 3.



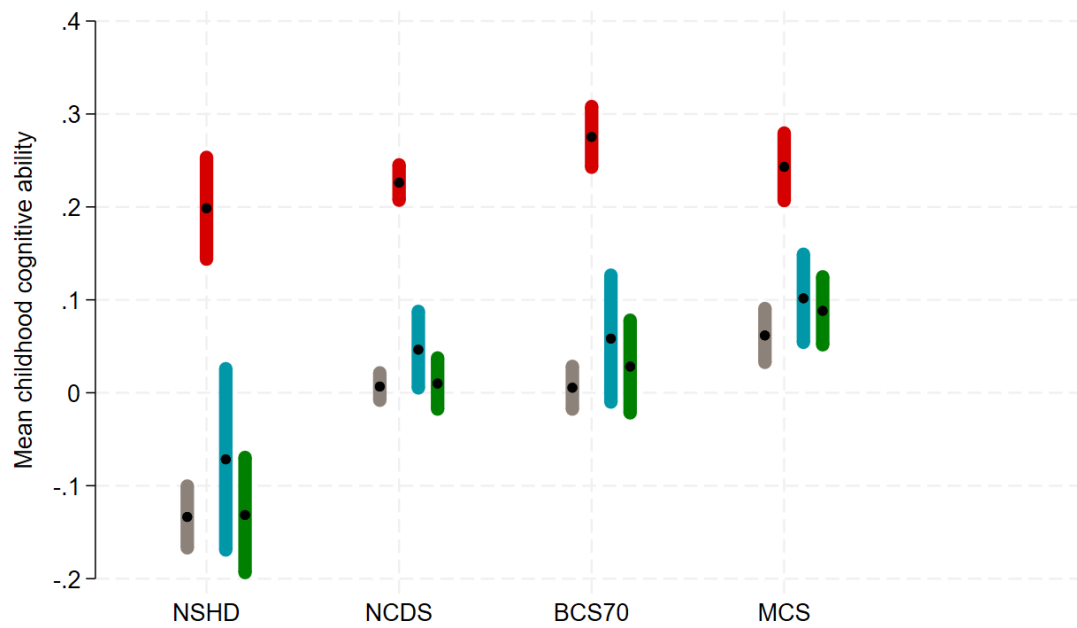
Percentage of highest social class (professional/managerial) in each cohort under different estimation approaches to account for non-response in the COVID-19 Wave 3 survey. Grey: using observed baseline data from the whole cohort; red: using observed baseline data from COVID-19 Wave 3 survey respondents only – unweighted (NCDS and BCS70) or using design weight only (NSHD, Next Steps and MCS); blue: using observed baseline data from COVID-19 Wave 3 survey respondents only – weighted using non-response weights (in addition to design weights as appropriate); green: using multiple imputation (plus design weights where appropriate). NSHD: National Survey of Health and Development; NCDS: 1958 National Child Development Study; BCS70: 1970 British Cohort Study; MCS: Millennium Cohort Study.

Fig. S3: Restoring representativeness for childhood cognitive ability in COVID wave 2.



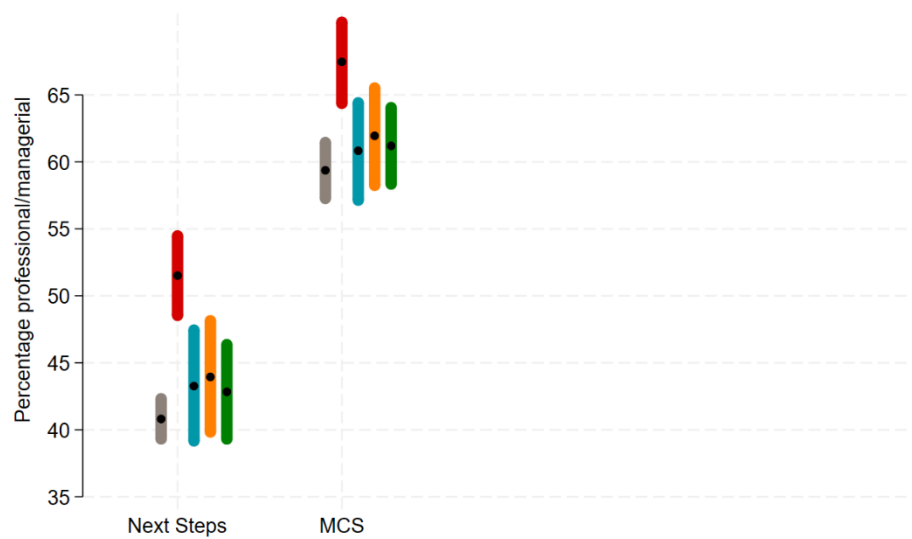
Mean of childhood cognitive ability in each cohort under different estimation approaches to account for non-response in the COVID-19 Wave 2 survey. Grey: using observed baseline data from the whole cohort; red: using observed baseline data from COVID-19 Wave 2 survey respondents only – unweighted (NCDS and BCS70) or using design weight only (NSHD, Next Steps and MCS); blue: using observed baseline data from COVID-19 Wave 2 survey respondents only – weighted using non-response weights; green: using multiple imputation. NSHD: National Survey of Health and Development; NCDS: 1958 National Child Development Study; BCS70: 1970 British Cohort Study; MCS: Millennium Cohort Study. Design weights were used in the estimation of means when available (NSHD, MCS) which explains why the mean of the standardised score is not always exactly 0.

Fig. S4: Restoring representativeness for childhood cognitive ability in COVID wave 3.



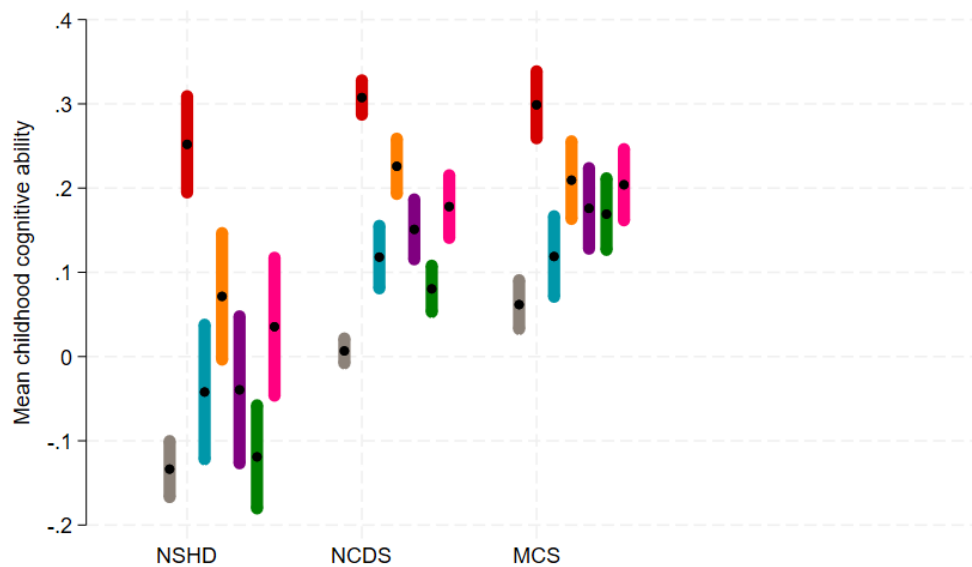
Mean of childhood cognitive ability in each cohort under different estimation approaches to account for non-response in the COVID-19 Wave 3 survey. Grey: using observed baseline data from the whole cohort; red: using observed baseline data from COVID-19 Wave 3 survey respondents only – unweighted (NCDS and BCS70) or using design weight only (NSHD, Next Steps and MCS); blue: using observed baseline data from COVID-19 Wave 3 survey respondents only – weighted using non-response weights; green: using multiple imputation. NSHD: National Survey of Health and Development; NCDS: 1958 National Child Development Study; BCS70: 1970 British Cohort Study; MCS: Millennium Cohort Study. Design weights were used in the estimation of means when available (NSHD, MCS) which explains why the mean of the standardised score is not always exactly 0.

Fig. S5: Sensitivity analyses for parental social class.



Percentage of highest social class (professional/managerial) in each cohort under different estimation approaches to account for non-response in the COVID-19 Wave 1 survey. Grey: using observed baseline data from the whole cohort; red: using observed baseline data from COVID-19 Wave 1 survey respondents only – unweighted (NCDS and BCS70) or using design weight only (NSHD, Next Steps and MCS); blue: using observed baseline data from COVID-19 Wave 1 survey respondents only – weighted using the original non-response weights from the COVID web survey user guide; orange: weighted using new non-response weights that do not include the same parental social class variable as in our example; green: using multiple imputation. MCS: Millennium Cohort Study.

Fig. S6: Sensitivity analyses for childhood cognitive ability.



Mean of childhood cognitive ability in each cohort under different estimation approaches to account for non-response in the COVID-19 Wave 1 survey. Grey: using observed baseline data from the whole cohort; red: using observed baseline data from COVID-19 Wave 1 survey respondents only – unweighted (NCDS and BCS70) or using design weight only (NSHD, Next Steps and MCS); blue: using observed baseline data from COVID-19 Wave 1 survey respondents only – weighted using the original non-response weights from the COVID-19 survey user guide; orange: weighted using new non-response weights that do not include the same cognitive ability variable used in our example; purple: weighted using new non-response weights that use a cognitive ability measure from a different wave than the one in our example (same as in multiple imputation model); green: using multiple imputation; pink: using multiple imputation based on the same variables used for the derivation of the original non-response weights (which means not including cognitive ability measures from other time points). NCDS: 1958 National Child Development Study; MCS: Millennium Cohort Study. Design weights were used in the estimation of means when available (MCS) which explains why the mean of the standardised score is not always exactly 0.