LSE Enterprise

CHILDREN'S WELLBEING AND DEVELOPMENT OUTCOMES FOR AGES 5, 7 AND 11, AND THEIR PREDICTORS
FINAL REPORT

December 2017

AUTHORS

(Listed Alphabetically)

Mireia Borrell-Porta (Post-Doctoral Researcher, Department of Social Policy, Oxford University)

Kerris Cooper (Doctoral Researcher, Department of Social Policy, LSE)

Joan Costa Font (Associate Professor of Political Economy and Social Policy, LSE)

Chiara Orsini (Post-Doctoral Researcher and Visiting Researcher, Social Policy, LSE)

Berkay Ozcan* (Assistant Professor of Social Policy, LSE)

Lucinda Platt (Professor of Social Policy and Sociology, LSE)

*Team leader.

London School of Economics (Department of Social Policy and LSE Enterprise)

TABLE OF CONTENTS

1. Introduction

In this report, we aim to identify the factors present during pregnancy, infancy, and the early years of life that are associated with a wide range of outcomes for children at ages 5, 7, and 11 by conducting a comprehensive analysis of the UK Millennium Cohort Study (MCS), a large-scale multidisciplinary child cohort study. This study builds on an earlier project (the PREview study) that used the MCS to identify factors during pregnancy, shortly after birth, and at age 3 that were associated with child outcomes at age 5 (Hobcraft and Kiernan, 2010).[1] Compared to the earlier project, we cover a broader window of children's lives, as we take into account variables measured between birth and age 11; and we analyse a broader set of outcomes. These outcomes, which cover four domains of children's lives (health, cognitive, behavioural, and social), are explained in greater detail in Section 2.3. These extensions imply that we can potentially identify areas for early intervention up to the end of primary school and across the range of young children´s experience.

Specifically, this project focuses on the following research questions:

1. What child and family factors predict child outcomes in the domains of health; behaviour; cognitive and educational development; and peer relations at ages 5, 7, and 11?
2. What is the relative importance of these different predictors in shaping child outcomes?
3. Are any of these factors protective or markers of 'resilience'? That is, are they associated with more positive outcomes (or less negative outcomes) for those from less advantaged backgrounds?

This report contains the main findings of the study. The study was informed by a separate *Literature Review*; and the report is accompanied by a full set of appendices that include further details and additional analyses.

---

[1] *Hobcraft, J. N., & Kiernan, K. E. (2010). Predictive factors from age 3 and infancy for poor child outcomes at age 5 relating to children's development, behaviour and health: evidence from the Millennium Cohort Study.* University of York, York.

2. Millennium Cohort Study

2.1. Data

The Millennium Cohort Study (MCS) is a nationally representative, multi-topic survey of over 19,000 babies born between September 2000 and January 2002 who are followed over time. The sample population for the study was drawn from live births in the UK using Child Benefit Records. The sample was selected from a random sample of electoral wards, and was disproportionately stratified to ensure adequate representation of all four UK countries, deprived areas, and, in England, areas with high concentrations of Black and Asian families (Plewis 2007).[2] In this report we focus on those living in England only, as outlined further below under Sample.

The children and their families have been surveyed at ages 9 months, 3 years, 5 years, 7 years, and 11 years. The most recent survey took place when the cohort children were around 14 years old. Data have been collected for each survey from the child's main carer (typically the mother), as well from the main carer's co-resident partner (typically the father). From age 3 onwards, direct measures of height and weight and cognitive assessments were carried out with the cohort children. For each survey from age 7 onwards, the cohort children have filled in a self-completion questionnaire. Each cohort child's class teacher was also asked to complete a questionnaire as part of the surveys when the child was age 7 and age 11. In addition, permission to link both the children's and the parents' records to a variety of administrative sources has been sought on a number of occasions. Enhancement studies collecting bio-measures have also been carried out. For further information, see Connelly and Platt (2014).[3] Analysis of the MCS can provide extensive insights into children´s social, cognitive, educational and emotional development and their health up to the end of primary school and how this relates to their family context at different ages, as well as their earlier experiences. It is also possible to relate the findings to the broader educational, social and policy context in which these children of the new millennium are growing up.

The complex sample design with oversampling of children in certain circumstances (those living in deprived areas, ethnic minority children) enables us to provide robust analysis of groups of particular interest. By adjusting our estimates for the complex survey design and applying weights for initial non-response and subsequent attrition, we are nevertheless able to provide nationally representative information about children growing up in England.

2.2. Sample

The following sample selection and exclusion criteria have been applied:

— The sample includes only singleton children living in *England*.

Because we focus on England, our results can inform the practices and policies of Public Health England. The sample only includes children who are living in England in all survey

---

[2] Plewis, I. (ed). (2007) *Millennium Cohort Study First Survey: Technical Report on Sampling (4th Edition)*. London: Centre for Longitudinal Studies, Institute of Education.

[3] Connelly, R. and Platt, L. (2014) *Cohort Profile: UK Millennium Cohort Study (MCS)*. International Journal of Epidemiology, *43(6), 1719-1725.*

sweeps. This ensures that a comparable sample is produced across sweeps and outcomes. The small numbers of multiple births (twins and triplets) in the MCS are excluded as predictors (such as birth weight) and outcomes of these cases are likely to require different interpretations.

— The analysis focuses on children's outcomes measured at ages 5, 7 and 11.

However, predictors may be measured at any age. These predictors may include measures that form outcome measures at later sweeps. For example, behaviour at age 3 may be used to predict cognitive outcomes at age 5.

We use the relevant sample for the outcome and predictor set under consideration. The precise nature of the sample for any given analysis will depend on the instruments used to collect the data, responses to the specific questions, and how the data from different instruments are combined in the analysis. For example, the parental self-completion questionnaire, which contains information on behavioural outcomes, has a lower response rate than the main interviewer-administered questionnaire. The child physical and cognitive assessments have high response rates of over 90 per cent, but they do not correspond precisely to the responses to the main questionnaire. The self-completion questionnaires filled in by the children at age 7 had slightly lower rates of completion than the assessments and the child self-completion questionnaires filled in by the children at age 11 (see Hansen 2014[4] for detailed information on the response rates across the survey instruments).

In sum, the results we present in each section are estimated on samples that differ across outcome measures. For each outcome, we aim to achieve the maximum sample size, and include all respondents for whom we have relevant information on the predictor and outcome measures, regardless of their responses at other sweeps. For example, when we analyse age 7 outcomes with reference to predictors collected at sweep 1 (fixed and pre-natal measures) and at age 5 (the sweep preceding the outcome), we do not exclude the respondents who had missing observations at ages 3 or 11.

Each outcome is weighted using the survey weight (combining the design and the non-response weights) that applies to the time the outcome was measured. For example, the age 5 country-specific weights are used for the age 5 outcomes, and the age 7 country-specific weights are used for the age 7 outcomes.

For analysis of change in outcomes over time, we additionally use a balanced panel of all children continuously observed at all survey sweeps. In this case, we apply the country-specific weights at age 11 for this restricted sample.

2.3. Outcome Measures at age 5, 7, 11

Outcome variables for children across the age range were identified from the review of the previous PREview study as well as through prior agreement between PHE and LSE. We provide various measures of outcomes falling in four main domains, as explained below.

Health Outcomes

---

[4] *Hansen, K. (ed.) (2014) Millennium Cohort Study: A Guide to the Data Sets (ninth edition): First, Second, Third, Fourth and Fifth Surveys. London: Centre for Longitudinal Studies, Institute of Education.*

BMI, Overweight and Obesity

For measurement of weight / overweight as a marker of children's physical health and development outcomes, we use:

a) BMI (included as a derived variable in the data), a continuous measure of the relationship between height and weight, which we adjust for age (in months) and child sex.

b) Overweight / obesity combined, as a discrete outcome. There are derived variables for these measures in the MCS dataset and these variables have been constructed using the International Obesity Task Force (IOTF) reference curves for the UK in 1990.

We do not separate overweight from obesity because a) they are often combined in the literature [*see* the *Literature Review*] b) the numbers for obesity are relatively small for analysis c) childhood obesity is a relative rather than a clinical measure.

Self-Reported Child Health (SRCH)

Self-Reported Child Health (SRCH) is a subjective health measure evaluated by the child's parents. It is measured in four categories, from very good to poor. We treat this as a continuous measure and model it linearly. There is some debate in the literature about treating reported health categories as linear, thus we carried out additional analysis to check whether ordinal specifications produced different results.

Cognitive Outcomes

We use measures of three types of cognitive outcome as collected at ages 5, 7 and 11, as detailed below. All cognitive measures are analysed as continuous variables. For further details on the cognitive measures see Hansen (2014) and Connelly (2013).[5] Rather than using the age-adjusted score in the MCS (which only adjust in bands of three months), we adjust scores by month of birth and school year in the multivariate regressions. This is because cognitive ability is likely to be influenced both by relative age (in months) and additionally by the school year the children are in, with those in a higher school year tending to perform better even for a given age.

The British Ability Scales II (BAS II)

The BAS II is comprised of a number of core sub-tests [e.g. BAS Naming Vocabulary, BAS Picture Similarity, BAS Pattern Construction and BAS Verbal Similarities] designed to measure both cognitive ability and educational achievement. BAS II measures are available at ages 5, 7 and 11, although different sub-tests are taken at different ages, as follows:

- Naming Vocabulary: age 5
- Pattern Construction: age 5, age 7
- Picture Similarities: age 5
- Word Reading: age 7

---

[5] Connelly, R. (2013) MCS data note 2013/1: Interpreting test scores. London: Centre for Longitudinal Studies. Available at: http://www.cls.ioe.ac.uk/page.aspx?&sitesectionid=1242&sitesectiontitle=Data+notes

- Verbal Similarities: age 11

Derived variables of scores constructed using the BAS scoring rules and subsequently standardised are supplied in the MCS datasets.

National Foundation for Education Research (NFER) Standard Progress in Maths

At age 7, cohort members completed an adapted, shorter version of the NFER standard Progress in Maths test. This test assesses a child's mathematical skills and knowledge by asking them to complete a series of calculations in a paper and pencil exercise. The test is read aloud to the child and covers topics such as numbers, shapes, measurement and data handling. In order to complete a maths test within the restricted interview time an MCS version of the original the Progress in Maths (PiM) test was developed, this version of the test required the cohort members to complete fewer items than the full PiM test. As with the BAS II tests we use the standardized score provided in the dataset.

Key Stage 2

Department for Education administrative data in the form of the National Pupil Database (NPD) can be linked to the MCS conditional on parental consent. This provides maths and reading scores from the last year of primary school, the end of Key Stage 2 (age 10/11).

Behavioural Outcomes

Strengths and Difficulties Questionnaire (SDQ)

The Strength and Difficulties Questionnaire (SDQ) has five sub-scales with five items each: 1) Emotional Symptoms;
2) Conduct Problems;
3) Hyperactivity/Inattention;
4) Peer Problems; and
5) Prosocial behaviour.

It also allows for a Total Difficulties score, which is the sum of the sub-scales 1-4. We measure behavioural problems derived from the Strengths and Difficulties Questionnaire (SDQ) in the following ways:

— Total difficulties (as a continuous measure)

This measure is constructed as the sum of the four problem domains:
Emotional Symptoms, Conduct Problems, Hyperactivity/Inattention, Peer Problems (also as continuous variables)

We also measure as distinct behavioural outcomes:

— Externalising problems ('Hyperactivity scale' + 'Conduct scale')
— Internalising problems ('Emotion symptoms scale' + 'Peer Problems')
— Hyperactivity/ Attention Deficit

Note that 'Peer Problems' is used as a measure of social outcomes.

Social Outcomes

Bullying

For Bullying, we use the self-reported child measures from age 7 and 11 that measure frequency of bullying victimisation. [i.e. 'how often do other children bully you?' with three available response options: *Never; Some of the time; and All of the time*]. We treat this measure as ordinal or as categorical. For example, at age 7, we contrast bullied "all the time" with the other responses.

Peer Problems

We use the peer problems subscale of the SDQ as a continuous measure of peer problems.

2.4. Selecting Predictors

Guiding Principles for Selecting Predictors

Predictor variables are identified from our literature review and from the review of the PREview report. We have followed the steps and principles described below:

1. The *Literature Review* guided us in scoping areas for further exploration not covered in the PREview research.

2. We have matched the concepts from the *Literature Review* with measures that proxy for them in the MCS at each age.

3. We have been sensitive to the age-specificity of some factors identified in the *Literature Review*. We have selected the relevant age-appropriate measures.

4. When the evidence in the literature is strong and well-established for a given predictor of any child outcome, for example "family income", we include it as a control variable in our multivariate analyses. In this way, the associations we present are independent of the effects of these clearly influential variables.

As a result, we have different sets of predictors for different types of outcomes. We present the list of predictors for each outcome in approximately the same way, according to stages in children's and families' life courses, as described below. We also distinguish constant and time-varying factors.

The predictor variables can be categorical (e.g. parental work status) or continuous (e.g. income). We discuss below how our modelling strategies respect the distributional features of both types of variables. However, transformations and standardisations are applied when necessary. We outline the implications of different ways of aggregating responses.

Following common practice, and as operationalised in the PREview research, predictor variables are observed at least a wave before the outcome variables to respect the temporal order of predictors and outcomes, unless there is a good reason to look at contemporaneous associations in certain specifications. The lists of predictors included for each outcome for each age are presented in Appendix 1.

Groupings of Predictors

As noted, we aimed to group the predictors according to the part of the child's life they refer to, which corresponds to the organisation of the *Literature Review*. Namely we specify pre-natal, early childhood, time-varying and maternal/environmental factors, as appropriate for the outcome (see also Appendix 1).

3. Methods

To answer question 1, we model the bivariate associations between each of the predictor and outcome variables. For continuous outcome variables we use linear regression and the impact of the predictor can be straightforwardly interpreted as the change in the outcome for 1 unit change in the predictor. For categorical outcomes, we follow the practice of the PREview report and present odds ratios; but we also estimate marginal effects to enable us to compare the influence of different predictor variables.

For question 2, we turn to multivariate analysis, where we include all those predictors that were statistically significantly associated with the outcome in the bivariate analysis concurrently. We then compare their relative importance taking into account both size of the effect and its statistical significance.

For question 3, we select potential measures of resilience or 'protective factors' that are statistically significant in the multivariate analysis and interact them with measures of disadvantaged family background, specifically income and parental education. We report any statistically significant interactions in the summary of findings for each outcome.

3.1. Bivariate Descriptions

We first inspect the relationship between each predictor variable and the relevant outcome variables separately (bivariate analysis), using appropriate bivariate regression models.

Health Outcomes

- BMI is a continuous variable and associations are estimated using bivariate linear regression.
- Obesity / overweight is a binary outcome variable and associations are estimated using bivariate logistic regression.
- SRCH is an ordered outcome variable and associations are estimated using bivariate linear regression. For robustness we also estimate ordered logistic regressions.

Cognitive Outcomes

- All the cognitive outcomes are continuous measures and associations are estimated using bivariate linear regression.

Behavioural Outcomes

- All the behavioural outcomes are continuous measures and associations are estimated using bivariate linear regression.

Social Outcomes

- Bullying is an ordered categorical variable and associations are estimated using ordered logistic regression.
- Bullying 'all the time' / 'most days' is a binary measures and associations are estimated using binary logistic regression.
- "Peer problems" is a continuous measure and associations are estimated using bivariate linear regression.

LSE Enterprise

3.2. Multivariate Models

Our selection of multivariate model is again dependent on the outcome measures.

- For continuous outcomes we use linear regression (OLS); for binary outcomes, we use a non-linear specification [Logistic regression]; and for ordinal outcomes we estimate ordered logit models [Ologit].
- We additionally estimated child fixed effects models, where possible, to take advantage of longitudinal nature of the dataset. These models link changes in the levels of predictors to changes in child outcomes, and therefore require repeat measures of the relevant outcomes (and predictors), as we discuss further below.

In all multivariate analyses, we report two specifications, one with unbalanced samples, and one with the balanced samples used in the fixed effects models.

3.2.1. Child Fixed-Effects Models

A common strategy for estimating causal effects of predictors on outcome measures is the application of fixed-effects models to longitudinal data where there are repeated observations of dependent and independent variables over time for a given individual. Child-specific fixed effects eliminate all time-constant differences among children and their family environment. The child-fixed effects model is equivalent to including a separate dummy variable for each child that captures all unobserved, time-constant differences among children. Alternatively, it is equivalent to factoring out children's average outcomes and values on predictor variables and focusing on the deviations from these averages, and how they vary together.

The main advantage of these models is that all unmeasured, constant factors are controlled for and the coefficients of the factors that are changing over time are not affected by these factors. Because comparison is now made changes over time not between children, these models are also called "within-child models" (See Mclanahan, Tach, Schneider 2013[6]). The main disadvantage of these models is that it is impossible to estimate the contribution of time constant factors on the outcomes, even though these may be of substantial interest (and substantive relevance).

In order to estimate child fixed-effects, the first requirement is that child outcomes need to be repeated. Thus, it is important to note here that:

- Not all child outcomes are repeated and comparable across ages.
- We can obtain coefficients for only a subset of predictors that vary over time.
- Since we combine all ages to estimate these models, the coefficients are always estimated on the balanced sample and reported to have an average effect over ages 5, 7 and 11.
- Lack of significant association may not always mean that the predictors are not associated with the outcome. There may be too few observations that change or too little change to be able to identify the association with sufficient statistical power.

Overall, fixed effects models provide evidence of causal associations where they are statistically significant, under the assumption that the multivariate fixed effects specifications include all possible time-varying measures (i.e. no unobserved time-varying measure is

---

[6] McLanahan, S., Tach, L., & Schneider, D. (2013). The causal effects of father absence. Annual review of sociology, 399, 399.

excluded from the specifications).They can only be taken as evidence that we cannot dismiss the possibility of no causal relationship, where there is an absence of a statistically significant relationship. We return to these points in our discussion of results.

3.2.2. Interpretation of Coefficients

Coefficients estimated using OLS in linear models are directly interpreted as "effect sizes", because they are equal to "marginal effects". We interpret these coefficients as one unit increase/decrease in the dependent variable with a one unit increase/ decrease in the predictor.

We report "average marginal effects (AMEs)" calculated from multivariate logistic regressions when we analyse binary outcome measures. We report "average marginal effects" (AMEs) to ease interpretation and make it consistent with that from linear (OLS) regressions. AMEs from logistic regressions can be interpreted as the percentage point change in the outcome implied by a one unit change in the predictor.

3.3. Description of Multivariate Models for Each Outcome

Health Outcomes

BMI, Overweight and Obesity

- BMI is a continuous variable and associations are estimated using multivariate linear regression (OLS).
- Obesity / overweight is a binary outcome variable and modelled using multivariate logistic regression.

These variables have consistent measures across ages 5, 7, 11, i.e. three observation points, which allowed us to run fixed effects specifications using all three survey sweeps.

These models are run separately for boys and girls as it is customary in the literature on child adiposity.

Self-Reported Child Health

SRCH is an ordered outcome variable. For these purposes we treat it as continuous and associations are estimated using multivariate linear regression. For robustness we also check these estimates against those deriving from an ordered logistic regression.

This variable has consistent measures across ages 5, 7, 11, i.e. three observation points, which allows us to estimate fixed effects specifications using all three survey sweeps.

Cognitive Outcomes

All the cognitive outcomes are continuous measures and are modelled using multivariate linear regression models.

Only one cognitive measure is available for more than one sweep (Pattern Construction at ages 5 and 7), so we estimate fixed effects estimations for just two observation points. Fixed effects models estimated only on two data points are equivalent to a first difference model,

LSE Enterprise

where the differences between two data points of all variables are taken and the regression is estimated on these differences.

Behavioural Outcomes

All the behavioural outcomes are continuous measures and are modelled using multivariate linear regression.

The SDQ is measured consistently across ages 5, 7, 11, i.e. three observation points, which allowed us to run fixed effects specifications using all three survey sweeps.

Social Outcomes

- Bullying is an ordered categorical variable and is modelled using multivariate ordered logistic regression and, concentrating on 'all of the time' versus other responses, binary logistic regression.
- "Peer problems" is a continuous measure and is modelled using multivariate linear regression.

Bullying is measured at ages 7 and 11, and we estimate fixed effects models just for these survey sweeps [i.e. first difference models]. Peer problems are measured at all three ages (5, 7 and 11). We estimate fixed effects models using all three data points.

LSE Enterprise

4. Results

4.1. Bivariate Descriptions

Tabulations of all bivariate associations are provided in full in the Appendix. Below, we summarise the bivariate associations between each outcome and the predictors. A positive association means it makes the outcome more likely, a negative association means that factor makes the outcome less likely. For example, in Table 1, below, smoking during pregnancy is associated with higher BMI and a greater risk of overweight for both boys and girls, while parental marriage cohabitation at birth is associated with lower BMI and a lower risk of overweight at age 5 for boys, compared to those whose parents were not cohabiting when they were born. Since these are the simple bivariate associations, the relationships may be confounded by other factors. This is addressed in the subsequent analyses of multivariate relationships.

Health Outcomes

Obesity and Body Mass Index (BMI)

Associations between predictors and BMI and Obesity were estimated separately for boys and girls. The table below summarizes all the associations. Sex differences in the associations are mentioned in parentheses, indicating that given association is only significant for that sex.

Table 1: Bivariate Associations: BMI and Overweight by Age

	BMI			Obesity/Overweight		
	Age 5	Age 7	Age 11	Age 5	Age 7	Age 11
Smoking during pregnancy Yes	+ve	+ve	+ve	+ve	+ve	+ve
Low birth weight Yes	Ns	Ns	Ns	Ns	Ns	Ns
Gestational age Yes	Ns	Ns	Ns	Ns	Ns	Ns
Married/Cohabiting birth Yes	-ve (boys)	-ve (girls)	-ve (girls)	-ve (boys)	-ve (girls)	Ns
Income Quintiles age 9m Highest quintile	(base)	(base)	(base)	(base)	(base)	(base)
Fourth quintile	+ve	+ve	+ve	+ve (girls)	+ve (girls)	+ve
Third quintile	+ve	+ve	+ve	+ve (girls)	+ve	+ve
Second quintile	+ve (girls)	+ve	+ve	+ve (boys)	+ve	+ve
Lowest quintile	+ve (girls)	+ve	+ve	+ve	+ve	+ve
Income Quintiles age 3y Highest quintile	(base)	(base)	(base)	(base)	(base)	(base)
Fourth quintile	+ve	+ve	+ve	+ve (girls)	+ve (girls)	+ve
Third quintile	+ve	+ve	+ve	+ve(girls)	+ve	+ve
Second quintile	+ve (girls)	+ve	+ve	+ve (boys)	+ve	+ve
Lowest quintile	+ve (girls)	+ve	+ve	+ve	+ve	+ve
Month of birth Coefficient	Ns	Ns	Ns	Ns	Ns	Ns

Education						
NVQ level 5	(base)	(base)	(base)	(base)	(base)	(base)
NVQ level 4	Ns	Ns	+ve (girls)	Ns	Ns	+ve (girls)
NVQ level 3	Ns	Ns	+ve (girls)	+ve (girls)	Ns	+ve (girls)
NVQ level 2	Ns	+ve	+ve (girls)	Ns	+ve	+ve (girls)
NVQ level 1	Ns	+ve	+ve (girls)	+ve (boys)	Ns	+ve (girls)
None of these	+ve (boys)	+ve	+ve (girls)	+ve (boys)	+ve	+ve (girls)
Family structure age 9m						
Two parents/carers	(base)	(base)	(base)	(base)	(base)	(base)
One parent/carer	+ve	+ve (boys)	+ve	+ve	+ve (boys)	+ve
Mother's age at child's birth						
Additional year of age	Ns	Ns	Ns	Ns	Ns	Ns
Family structure age 3						
Two parents/carers	(base)	(base)	(base)	(base)	(base)	(base)
One parent/carer	+ve	+ve (boys)	+ve	+ve	+ve (boys)	+ve
Maternal weight before pregnancy						
Coefficient	+ve	+ve	+ve	+ve	+ve	+ve
Breastfeeding Not breastfed (base)						
Breastfed 1 day-12 weeks	+ve (boys)	Ns	Ns	Ns	Ns	Ns
Breastfed 13 weeks to 6 months	Ns	-ve	-ve	-ve (boys)	-ve (boys)	-ve
Breastfeed more than 6 months	-ve (girls)	Ns	Ns	Ns	Ns	Ns
Maternal BMI at 9m						
Coefficient	+ve	+ve	+ve	+ve	+ve	+ve
Working during pregnancy						
Yes	Ns	Ns	-ve	Ns	Ns	-ve
Working in sweep 1						
Yes	Ns	Ns	Ns	Ns	Ns	Ns
Regular bed times						
Yes	-ve (boys)	-ve	Ns	-ve (boys)	-ve	Ns
Fresh fruits						
Yes	Ns	Ns	Ns	Ns	Ns	Ns
Child enjoys exercise						
Yes		Ns	-ve		Ns	-ve
Child eats breakfast						
Yes	Ns	Ns	Ns	Ns	Ns	Ns

Summary of Findings

Here we summarise the associations of predictors with BMI and overweight. Unless we specify otherwise all associations are found at all ages and for both girls and boys.

Maternal Predictors

- Maternal weight before pregnancy positively predicts BMI and obesity among boys and girls at, but birth weight did not.
- For example, a one kg increase in maternal weight increases BMI at age 5 by 1.4 among boys and 1.7 among girls, and the probability of overweight by 2 percentage points among boys and 3 percentage points among girls.

LSE Enterprise

- Maternal BMI at 9 months is a strong and significant predictor of BMI and overweight: an increase of one unit on maternal BMI is associated with an increase of coefficient being 0.24 for child's BMI and of 2 percentage points on the probability of overweight at all ages.
- Breastfeeding has mixed associations, depending both on the duration of breast-feeding and the age at which BMI / overweight is measured. Breastfeeding up to 6 months is associated with lower BMI and reduced probabilities of being overweight, even up to age 11. But being breastfed for longer or shorter periods has no impact on weight relative to not being breastfed.
- When BMI and overweight are evaluated at age 11 we find that smoking during pregnancy continues to have a high correlation with BMI and overweight, and the same applies to maternal weight before pregnancy.
- When overweight is evaluated at age 7 we find some different results. Whether the mother smoked during pregnancy increases BMI and obesity by a very large magnitude that exceeds the effect of maternal BMI in sweep 1 both among boys and girls.

B) Family and Environmental Predictors

- Family composition is associated with both BMI and overweight: we find that children whose parents were single in sweep 1 exhibit both a higher BMI (boys exhibit 0.28 and girls 0.21 units higher BMI) and obesity. Consistently cohabiting parents are associated with lower BMI and lower risk of overweight among boys at all ages and among girls at ages 5 and 11.
- The detrimental effect of single (cohabiting) parenthood at sweep 1 are only significant among girls. However, single parenthood at sweep 3 is associated with a substantial increase in child´s BMI and overweight.
- Lower income at sweep one increases girls' BMI linearly. Parental socio-economic position in sweep 1 and 3 suggest that a lower income quintile leads to larger BMI and a greater risk of overweight but the effect is non-linear.
- Having regular bed time reduces BMI and the risk of overweight for both boys and girls at age 7, but only for boys at age 5.
- Lower maternal education (NVQ) appears to increase BMI and overweight among girls.
- Although eating behaviour is not associated with higher BMI and overweight, exercise is associated with reduced BMI and risks of overweight at age 11.

Self-Reported Child Health

Table 2, below, reports the association of relevant factors with self-reported child health at ages 5, 7 and 11.

Table 2: Bivariate Associations between Predictors and Self-Reported Child Health

	Age 5	Age 7	Age 11
Smoking during pregnancy Yes	-ve	-ve	-ve
Low birth weight yes	-ve	-ve	-ve

LSE Enterprise

Drinking			
Never	(base)	(base)	(base)
Less than once a month	+ve	+ve	+ve
1-2 times a month	+ve	+ve	+ve
1-2 times a week	+ve	+ve	+ve
3-4 times a week	Ns	Ns	Ns
5-6 times a week	Ns	Ns	Ns
Every day	+ve	Ns	Ns
Fresh vegetables once a day			
Yes	-ve	-ve	-ve
Home help child			
Yes	+ve	+ve	+ve
Felt low or sad			
Yes	-ve	-ve	-ve
Mothers general health			
Fair	Ns	+ve	Ns
Good/V Good	+ve	+ve	+ve
Excellent	+ve	+ve	+ve
Mother long standing illness			
Yes	-ve	-ve	-ve
Mother's mental health - Kessler			
coefficient	-ve	-ve	-ve
Income Quintiles			
Highest quintile	(base)	(base)	(base)
Fourth quintile	-ve	-ve	-ve
Third quintile	-ve	-ve	-ve
Second quintile	-ve	-ve	-ve
Lowest quintile	-ve	-ve	-ve
Psychosomatic symptoms (stomach ache)			
Yes	-ve	-ve	-ve
Education			
NVQ level 5	+ve	+ve	+ve
NVQ level 4	+ve	+ve	+ve
NVQ level 3	+ve	+ve	+ve
NVQ level 2	+ve	+ve	+ve
NVQ level 1	Ns	Ns	Ns
None of these	Base	Base	Base
Overseas qual only	-ve	-ve	-ve
Child picked or bullied			
Yes	-ve	-ve	-ve
Mother's age at child's birth			
20 to 29	(base)	(base)	(base)
12 to 19	+ve	+ve	+ve
30 to 39	+ve	+ve	+ve
40 plus	+ve	+ve	Ns
Sex of the child			
Female	+ve	+ve	Ns
Number of siblings			
1	+ve	Ns	Ns
2	Ns	Ns	-Ns
3+	Ns	-ve	-ve
Ethnicity			
Mixed	-ve	Ns	-ve
Indian	-ve	-ve	-ve
Pakistani and Bangladeshi	-ve	-ve	-ve
Black or Black British	-ve	-ve	-ve
Other Ethnic group (inc. Chinese, Other)	-ve	-ve	Ns
Breastfeeding (age last)			

LSE Enterprise

1d-3months	Ns	Ns	Ns
13 weeks- 6months	+ve	+ve	+ve
More than 6 months	+ve	+ve	+ve
Month of birth			
September	Base	Base	Base
October	Ns	Ns	Ns
November			
December	-ve	-ve	-ve
January	-ve	Ns	Ns
February	Ns	Ns	Ns
March	Ns	Ns	Ns
April	+ve	Ns	Ns
May	Ns	+ve	Ns
June	Ns	Ns	Ns
July	Ns	Ns	Ns
August	Ns	Ns	Ns
Exercise with child (Ref: No Exercise)			
At least once a year		Ns	Ns
Every few months		+ve	+ve
At least once a month		+ve	+ve
Once or twice a week		+ve	+ve
Several times a week		+ve	+ve
Every day		+ve	+ve
Child exercise (days week) (Ref: None)			
1		+ve	+ve
2		+ve	+ve
3		+ve	+ve
4		+ve	+ve
5 or more		+ve	+ve
Portion of fruit per day			
1		Ns	+ve
2		+ve	+ve
3 or more		+ve	+ve

Summary of Findings

A) Maternal Predictors

- Smoking during pregnancy reduces child self-reported health and the association persists to age 11, though it is smaller at that point.
- We find that mother drinking moderately (up to 1-2 times per week) during pregnancy show an unexpected positive effect on self-reported health of the child. The coefficients were consistently significant and decline from 0.15 to 0.10 for drinking 1-2 times per week across the child age.
- Maternal age at birth is correlated with child health. Better health at all child ages is associated with mothers who were 30 to 39 at the child's birth (though the association declines from 0.23 at age 5 to 0.15 at age 11).
- Maternal education predicts higher self-reported child health, particularly for women with a NVQ level of 4 and 5 (the coefficient for the higher NVQ takes the value of 0.38 at age 5 declining to 0.33 at age 11). The largest coefficient size is observed for mother's mental health around 0.67 at age 5 and 0.61 at age 11.
- Maternal long standing illness reduces child health in a similar magnitude across the child's early childhood.

LSE Enterprise

- Worse maternal mental health (as measured by a higher score on the Kessler scale) is negatively correlated with child health at all ages, with effect sizes of a similar magnitude (around -0.032 points on a four-point scale).
- Similarly, post-natal depression reduces self-reported health at all ages by a magnitude of -0.23 at age 5 with a smaller effect of -0.15 at age 11.

B) Family and Environmental Effects

- Low birth weight is associated with worse child health, and the coefficient size increases to 0.3 at age 11.
- Psychosomatic symptoms are also linked to worse self-reported health.
- Bullying is consistently associated with worse self-reported health across all ages.
- Non-white ethnicity is negatively correlated with self-reported health, and the effect is larger for Indian and Pakistani children than for other ethnicities, consistent with existing literature. The effect size for Pakistani is -0.56 at age 5 and only declines a small amount to –0.4 at age 11.
- Having more than three siblings is associated with lower self-reported health at all ages.
- Lower income is consistently and linearly negatively correlated with self-reported health.
- Breastfeeding between 13 weeks to 6 months or more than 6 months' increases child health, but the effect is strongest among mothers who breastfed from 13 weeks to 6 months.
- Girls exhibit a health advantage compared to boys, although the effect disappears by age 11.
- Parental and child exercise are correlated with higher self-reported health (e.g., correlation coefficient 0.38 of child exercising three days a week), similarly fruit consumption increases self-reported health (e.g. three portions a day show coefficient of 0.38).

Cognitive Outcomes

Table 3 below shows bivariate associations between predictors and various cognitive outcomes by age. As we discussed, the type of cognitive test varies by age. Therefore we summarize the results from Table 3 by age rather than groups of predictors. Tables 4 and 5 reports the Key Stage 2 test results for English (and Reading) and Maths (and Arithmetic) at age 11. Those results are summarized separately, following the tables. The summary of findings focuses on predictors that are consistently significant across tests within a given age.

Table 3: Bivariate Associations between Predictors and Cognitive Outcomes by Age

	Age 5			Age 7			Age 11
	Naming	**Picture**	**Pattern**	**Word**	**Maths**	**Pattern**	**Verbal**
Low birth weight							
Yes	+ve	Ns	Ns	+ve	Ns	Ns	+ve
Working at sweep 1							
Yes	+ve	+ve	+ve	+ve	+ve	+ve	+ve
Child care at sweep 1							
coefficient	+ve	+ve	+ve	+ve	+ve	+ve	+ve
Income Quintiles sweep 1							
Highest quintile	(base)	(base)	(base)	(base)	(base)	(base)	(base)
Fourth quintile	-ve	-ve	-ve	-ve	-ve	-ve	-ve
Third quintile	-ve	-ve	-ve	-ve	-ve	-ve	-ve
Second quintile	-ve	-ve	-ve	-ve	-ve	-ve	-ve
Lowest quintile	-ve	-ve	-ve	-ve	-ve	-ve	-ve
Income Quintiles sweep 2							
Highest quintile	(base)	(base)	(base)	(base)	(base)	(base)	(base)
Fourth quintile	-ve	-ve	-ve	-ve	-ve	-ve	-ve
Third quintile	-ve	-ve	-ve	-ve	-ve	-ve	-ve
Second quintile	-ve	-ve	-ve	-ve	-ve	-ve	-ve
Lowest quintile	-ve	-ve	-ve	-ve	-ve	-ve	-ve
Married/Cohabiting at birth							
Not in work	+ve	+ve	+ve	+ve	+ve	+ve	+ve
Education							
NVQ level 5	(base)	(base)	(base)	(base)	(base)	(base)	(base)
NVQ level 4	-ve	-ve	-ve	-ve	-ve	-ve	-ve
NVQ level 3	-ve	-ve	-ve	-ve	-ve	-ve	-ve
NVQ level 2	-ve	-ve	-ve	-ve	-ve	-ve	-ve
NVQ level 1	-ve	-ve	-ve	-ve	-ve	-ve	-ve
None of these	-ve	-ve	-ve	-ve	-ve	-ve	-ve
Overseas qual only	-ve	-ve	-ve	-ve	-ve	-ve	-ve
Family structure sweep 1							
Two parents/carers	(base)	(base)	(base)	(base)	(base)	(base)	(base)
One parent/carer	-ve	-ve	-ve	-ve	-ve	-ve	-ve
Family structure sweep 2							
Two parents/carers	(base)	(base)	(base)	(base)	(base)	(base)	(base)
One parent/carer	-ve	-ve	-ve	-ve	-ve	-ve	-ve
Mother's age							
Coeff	+ve	+ve	+ve	+ve	+ve	+ve	+ve
Gender of the child							
Female	+ve	+ve	+ve	+ve	Ns	+ve	Ns
Bleeding and threat of miscarriage early in pregnancy							
Yes	Ns	Ns	Ns	Ns	Ns	Ns	Ns

LSE Enterprise

Raised blood pressure, eclampsia /preeclampsia							
Yes	Ns	Ns	Ns	Ns	Ns	Ns	Ns
Working during pregnancy							
Coefficient	+ve	+ve	+ve	+ve	+ve	+ve	+ve
Working mother							
Coefficient	+ve	+ve	+ve	+ve	+ve	+ve	+ve
Breastfeeding (age last)							
1d-3months	+ve	Ns	+ve	+ve	Ns	+ve	+ve
13 weeks- 6months	+ve	+ve	+ve	+ve	+ve	+ve	+ve
More than 6 months	+ve	+ve	+ve	+ve	+ve	+ve	+ve

Summary of Findings

A) Age 5

Cognitive outcomes at age 5 are measured by three tests from British Ability Scales II (or BAS II): Naming Vocabulary, Picture Similarity, and Pattern Construction.

- Birth weight increases scores on naming vocabulary, but working mother and mother working during pregnancy and the child being in child care in year 1 are associated with higher scores in all three tests.
- Single parenthood is associated with lower scores on all three tests, especially vocabulary and pattern construction, and the opposite applies to cohabiting parents.
- Higher income and lower deprivation in sweeps 1 and 3 are associated with higher scores on all three tests, but especially on naming vocabulary.
- Children breastfed between more than 13 weeks reveal higher scores, as do those with more highly educated mothers.

B) Age 7

Children's cognitive abilities at age 7 are tested with three tests: BAS Pattern Construction, BAS Word Reading, and NFER Progress in Maths.

- Birth weight is associated with better word reading skills but not the other tests.
- Children with working mothers, working during pregnancy and children that used child care in the first year of birth have higher scores.
- Single parenthood is associated with lower scores in the three tests; and lower income and less deprivation at sweeps 1 and 3 are associated with monotonically higher scores.
- Children breastfed more than 6 months exhibit better scores than the rest, and breastfeeding is a protective risk.
- Maternal age and NVQ increases all scores, as well as starting school, but in contract attending nursery does not.

C) Age 11

At age 11, the BAS II verbal similarity test was accompanied by neuropsychological tests covering memory and strategic thinking, and decision making and risk taking, rather than additional tests of cognitive skills.

- Gender is not associated with higher scores, but birth weight is; and using child care at sweep 1, and mothers being in employment and working during pregnancy are associated with increased verbal similarity scores.
- Having cohabiting parents also shows positive effects, and by contrast, single parenthood shows a negative effect of similar magnitude.
- We find better outcomes for children born between May and September, those with lower deprivation scores and those whose parents are in the highest quintiles of income in both sweep 1 and 4.
- Consistent with other outcomes, children who were breastfed exhibit higher verbal similarity.
- Lower maternal education (NVQ) is linked to lower verbal similarity scores; while school club attendance, having a working mother and an older mother improves verbal similarity test scores.

LSE Enterprise

Key Stage 2 Outcomes: English and Mathematics Tests Scores

We report bivariate associations between Predictors and English and Reading Test Scores of children at the end of KS2 (age 11).

Table 4: Bivariate Associations of Predictors and English (and Reading) Test Scores

VARIABLES	Reading Score	English Score
Sex of the child		
Female	+ve	+ve
Birth weight		
Measures	+ve	+ve
Month of birth		
January	+ve	+ve
February	+ve	+ve
March	+ve	+ve
April	+ve	+ve
May	+ve	+ve
June	+ve	+ve
July	+ve	+ve
August	+ve	+ve
September	+ve	+ve
October	+ve	+ve
November	+ve	+ve
December	+ve	+ve
Child Childcare In Sweep 1	+ve	+ve
Mother Worked During Pregnancy	+ve	+ve
Mother Worked In Sweep 1	+ve	+ve
Single Parent In Sweep 1	-ve	-ve
Married Or Cohabiting At Birth	+ve	+ve
Deprivation at Sweep 1		
Index Mult Deprivation 10-20% Sweep 1	Ns	Ns
Index Mult Deprivation 20-30% Sweep 1	+ve	+ve
Index Mult Deprivation 30-40% Sweep 1	+ve	+ve
Index Mult Deprivation 40-50% Sweep 1	+ve	+ve
Index Mult Deprivation 50-60% Sweep 1	+ve	+ve
Index Mult Deprivation 60-70% Sweep 1	+ve	+ve
Index Mult Deprivation 70-80% Sweep 1	+ve	+ve
Index Mult Deprivation 80-90% Sweep 1	+ve	+ve
Index Mult Deprivation Highest Decile Sweep 1	+ve	+ve
Parental income quantile at Sweep 1		
Second Highest Income Quantile Sweep 1	-ve	-ve
Third Highest Income Quantile Sweep 1	-ve	-ve
Fourth Highest Income Quantile Sweep 1	-ve	-ve
Fifth Highest Income Quantile Sweep 1	-ve	-ve
Breastfeeding period		
Breastfeed Between 1 And 3 Months	+ve	Ns
Breastfeed Between 13 Weeks And 6 Months	+ve	+ve
Breastfeed More Than 6 Months	+ve	+ve
Bleeding And Threat Of Miscarriage Early In Pregnancy	Ns	Ns
Raised Blood Pressure, Eclampsia/Preeclampsia Or Toxaemia	Ns	Ns
Other Illness During Pregnancy	Ns	Ns

Deprivation at Sweep 4		
Index Mult Deprivation 10-20% Sweep 4	+ve	+ve
Index Mult Deprivation 20-30% Sweep 4	+ve	+ve
Index Mult Deprivation 30-40% Sweep 4	+ve	+ve
Index Mult Deprivation 40-50% Sweep 4	+ve	+ve
Index Mult Deprivation 50-60% Sweep 4	+ve	+ve
Index Mult Deprivation 60-70% Sweep 4	+ve	+ve
Index Mult Deprivation 70-80% Sweep 4	+ve	+ve
Index Mult Deprivation 80-90% Sweep 4	+ve	+ve
Index Mult Deprivation Highest Decile Sweep 4	+ve	+ve
Parental income quantile at Sweep 1		
Second Highest Income Quantile Sweep 4	-ve	-ve
Third Highest Income Quantile Sweep 4	-ve	-ve
Fourth Highest Income Quantile Sweep 4	-ve	-ve
Fifth Highest Income Quantile Sweep 4	-ve	-ve
National Vocational Qualification		
National Vocational Qualification 4	-ve	-ve
National Vocational Qualification 3	-ve	-ve
National Vocational Qualification 2	-ve	-ve
National Vocational Qualification 1	-ve	-ve
One Parent Carer At Sweep 4	-ve	-ve
Working Mother	+ve	+ve
Mother's Age	+ve	+ve
Whether Child Attends After School Club	Ns	+ve

Summary of Findings

- Girls and breastfed children have better reading and English scores.
- Low birth weight is negatively correlated with reading and English scores; but having married or cohabiting parents, and having a working mother during pregnancy or in childhood correlates with higher English and reading scores.
- Children of single parents, those with lower maternal education, and parents living in more deprived neighbourhoods and with lower socio-economic status tend to have lower reading and English scores.

LSE Enterprise

Table 5: Bivariate Associations between Predictors and Maths (and Arithmetic) Test Scores

VARIABLES	Arithmetic Scores	Math total Scores
Female Child	-ve	-ve
Birth Weight	+ve	+ve
Birth Month		
January	+ve	+ve
February	+ve	+ve
March	+ve	+ve
April	+ve	+ve
May	+ve	+ve
June	+ve	+ve
July	+ve	+ve
August	+ve	+ve
September	+ve	+ve
October	+ve	+ve
November	+ve	+ve
December	+ve	+ve
Child Childcare In Sweep 1	+ve	+ve
Mother Worked During Pregnancy	+ve	+ve
Mother Worked In Sweep 1	+ve	+ve
Single Parent In Sweep 1	-ve	-ve
Married Or Cohabiting At Birth	+ve	+ve
Deprivation at sweep 1		
Index Mult Deprivation 10-20% Sweep 1	Ns	+ve
Index Mult Deprivation 20-30% Sweep 1	+ve	+ve
Index Mult Deprivation 30-40% Sweep 1	+ve	+ve
Index Mult Deprivation 40-50% Sweep 1	+ve	+ve
Index Mult Deprivation 50-60% Sweep 1	+ve	+ve
Index Mult Deprivation 60-70% Sweep 1	+ve	+ve
Index Mult Deprivation 70-80% Sweep 1	+ve	+ve
Index Mult Deprivation 80-90% Sweep 1	+ve	+ve
Index Mult Deprivation Highest Decile Sweep 1	+ve	+ve
Parental Income at Sweep 1		
Second Highest Income Quantile Sweep 1	-ve	-ve
Third Highest Income Quantile Sweep 1	-ve	-ve
Fourth Highest Income Quantile Sweep 1	-ve	-ve
Fifth Highest Income Quantile Sweep 1	-ve	-ve
Breastfeeding period		
Breastfeed Between 1 And 3 Months	Ns	Ns
Breastfeed Between 13 Weeks And 6 Months	+ve	+ve
Breastfeed More Than 6 Months	+ve	+ve
Bleeding And Threat Of Miscarriage Early In Pregnancy	Ns	Ns
Raised Blood Pressure, Eclampsia/Preeclampsia Or Toxaemia	Ns	Ns
Other Illness During Pregnancy	-ve	-ve

Deprivation at sweep 4		
Index Mult Deprivation 10-20% Sweep 4	Ns	+ve
Index Mult Deprivation 20-30% Sweep 4	+ve	+ve
Index Mult Deprivation 30-40% Sweep 4	+ve	+ve
Index Mult Deprivation 40-50% Sweep 4	+ve	+ve
Index Mult Deprivation 50-60% Sweep 4	+ve	+ve
Index Mult Deprivation 60-70% Sweep 4	+ve	+ve
Index Mult Deprivation 70-80% Sweep 4	+ve	+ve
Index Mult Deprivation 80-90% Sweep 4	+ve	+ve
Index Mult Deprivation Highest Decile Sweep 4	+ve	+ve
Parental Income at Sweep 1		
Second Highest Income Quantile Sweep 4	-ve	-ve
Third Highest Income Quantile Sweep 4	-ve	-ve
Fourth Highest Income Quantile Sweep 4	-ve	-ve
Fifth Highest Income Quantile Sweep 4	-ve	-ve
National Vocational Qualification		
National Vocational Qualification 4	-ve	-ve
National Vocational Qualification 3	-ve	-ve
National Vocational Qualification 2	-ve	-ve
National Vocational Qualification 1	-ve	-ve
None Of The Above	-ve	-ve
One Parent Carer At Sweep 4	-ve	-ve
Working Mother	+ve	+ve
Mothers Age	+ve	+ve
Whether Child Attends After School Club	Ns	Ns

Summary of Findings

- Boys and children breastfed for more than 3 months have better arithmetic and maths scores.
- Low birth weight is negatively correlated with arithmetic and maths scores, but married or cohabiting and working mothers either during pregnancy or during childhood are associated with higher arithmetic and maths scores.
- Children of single parents, those with lower maternal education, and parents living in more deprived neighbourhoods and with lower socio-economic status have lower arithmetic and maths scores.

LSE Enterprise

Behavioural Outcomes

Here below, we report bivariate associations between predictors and four behavioural outcomes based on the Strength and Difficulties Questionnaire (SDQ): Total Difficulties, Externalising Behaviour, Internalising Behaviour and Hyperactivity/ Attention Deficit. The predictors in these sections broadly related to, first, mother's characteristics and behaviour, and then about family environment.

SDQ – Total difficulties [Sum of Four Problem Subscales]

Average values are between 6 and 7 on a scale ranging from 0-40.

Table 6: Bivariate Associations between Predictors and SDQ – Total Difficulties Score

	Age 5	Age 7	Age 11
Smoked in pregnancy			
Yes	+ve	+ve	+ve
Drinking during pregnancy			
Never	(base)	(base)	(base)
Less than once a month	-ve	-ve	-ve
1-2 times a month	-ve	-ve	-ve
1-2 times a week	-ve	-ve	-ve
3-4 times a week	-ve	Ns	Ns
5-6 times a week	-ve	Ns	Ns
Every day	-ve	Ns	Ns
Low birth weight			
Yes	+ve	+ve	+ve
Gestational age preterm			
Yes	+ve	+ve	+ve
Felt low or sad			
Yes	+ve	+ve	+ve
Depression - long standing illness			
Depression	+ve	+ve	+ve
Other illnesses	+ve	+ve	+ve
Depression - as diagnosed by Dr.			
Yes	+ve	+ve	+ve
Mother's mental health – Kessler			
Coefficient	+ve	+ve	+ve
Income Quintiles			
Highest quintile	(base)	(base)	(base)
Fourth quintile	+ve	+ve	+ve
Third quintile	+ve	+ve	+ve
Second quintile	+ve	+ve	+ve
Lowest quintile	+ve	+ve	+ve
Whether in work or not			
Not in work	+ve	+ve	+ve

LSE Enterprise

Education			
NVQ level 5	(base)	(base)	(base)
NVQ level 4	+ve	+ve	+ve
NVQ level 3	+ve	+ve	+ve
NVQ level 2	+ve	+ve	+ve
NVQ level 1	+ve	+ve	+ve
None of these	+ve	+ve	+ve
Overseas qual only	+ve	+ve	+ve
Family structure			
Two parents/carers	(base)	(base)	(base)
One parent/carer	+ve	+ve	+ve
Mother's age at child's birth			
20 to 29	(base)	(base)	(base)
12 to 19	+ve	+ve	+ve
30 to 39	-ve	-ve	-ve
40 plus	-ve	-ve	Ns
Sex of the child			
Female	-ve	-ve	-ve
Number of siblings			
1	-ve	-ve	-ve
2	-ve	-ve	-ve
3+	Ns	Ns	Ns
Ethnicity			
Mixed	+ve	Ns	Ns
Indian	Ns	Ns	-ve
Pakistani and Bangladeshi	+ve	+ve	+ve
Black or Black British	+ve	Ns	Ns
Other Ethnic group (inc. Chinese, Other)	Ns	Ns	Ns
Relationship with partner - happiness			
Coefficient	+ve	+ve	+ve
Importance of stimulation of the child			
Strongly agree	(base)	(base)	(base)
Agree	+ve	+ve	+ve
Neither agree nor disagree	+ve	+ve	+ve
Disagree	+ve	Ns	Ns
Strongly disagree	Ns	Ns	Ns

Summary of Findings

A) Maternal Predictors

- Smoking during pregnancy is associated with greater child behavioural problems, namely an increase of around 2 points.
- Maternal drinking is associated with a small negative effect on hyperactivity problems (i.e. drinking more reduces behavioural problems) except if the mother drinks every day. But heavy drinking is rare.
- Low birth weight and early gestational age are associated with a significant increase in child Total Difficulties of around 1-2 points.
- All measures of mothers' mental health/stress are associated with a strong positive effect of depression on child behaviour at all ages (i.e. depression increases total difficulties).

- Feeling low or sad: increase of around 2 points.
- Depression (self-report): increase of around 2-4 points.
- Depression as diagnosed by doctor: increase between 1-2 points.
- Kessler scale (continuous from 0 to 24): each point on the Kessler scale is associated with an increase of around 0.5 points.

B) *Family and Environmental Predictors*

- There is large significant association between income and externalising behaviour, with lower income associated with more problems. The second and lowest quintile groups have the strongest association with behavioural problems, increasing total difficulties by 2-5 points.
- Having no education or only level 1 NVQ, compared to level 4 NVQ, is associated with an increase in problems of around 2-5 points. Effect sizes are large.
- Children in one parent/carer families experience around 2.5 points more of total difficulties.
- Compared to children born to mothers aged 20-29 years old, children of younger mothers experience significantly more behavioural difficulties. The opposite is true for children born to older mothers.
- Girls are significantly associated with having fewer behavioural problems than boys.
- Having 1 or 2 siblings is significantly associated with fewer difficulties than no siblings. Having three or more siblings is not different to having no siblings.
- At age 5, Pakistani / Bangladeshi and Black / Black British children have more behavioural problems than white children. At age 7, the association with total difficulties is only significant for Pakistani or Bangladeshi children. The effect size and significance is similar to age 5.
- At age 11, the relationship remains significant for Pakistani or Bangladeshi children, but the size of the association is much smaller.
- There is significant association between lower levels of relationship quality/happiness and more child behavioural problems, especially strong for very unhappy couples.
- At all ages we observe significantly more difficulties among children whose parents do not think stimulation is important.

LSE Enterprise

SDQ – Externalising Problems

Externalising problems are created from the sum of hyperactivity/ attention problems and conduct problems. Average values are between 3 and 5 on a scale from 0-20.

Table 7: Bivariate Associations between Predictors and SDQ – Externalising Problems

	Age 5	Age 7	Age 11
Smoking			
Yes	+ve	+ve	+ve
Drinking			
Never	(base)	(base)	(base)
Less than once a month	Ns	Ns	Ns
1-2 times a month	Ns	Ns	-ve
1-2 times a week	-ve	-ve	Ns
3-4 times a week	Ns	Ns	Ns
5-6 times a week	-ve	Ns	Ns
Every day	Ns	Ns	Ns
Low birth weight			
Yes	+ve	+ve	+ve
Gestational age			
Yes	+ve	+ve	+ve
Felt low or sad			
Yes	+ve	+ve	+ve
Depression - long standing illness			
Depression	+ve	+ve	+ve
Other illnesses	+ve	+ve	+ve
Depression - as diagnosed by Dr.			
Yes	+ve	Ns	Ns
Mother's mental health - Kessler			
Coefficient	+ve	Ns	Ns
Mother's mental health - Kessler - categorical			
1	+ve	+ve	+ve
2	+ve	+ve	+ve
3	+ve	+ve	+ve
4	+ve	+ve	+ve
5	+ve	+ve	+ve
6	+ve	+ve	+ve
7 – high	+ve	+ve	Ns
Income Quintiles			
Fourth quintile	+ve	+ve	+ve
Third quintile	+ve	+ve	+ve
Second quintile	+ve	+ve	+ve
Lowest quintile	+ve	+ve	+ve
Whether in work or not			
No	+ve	+ve	+ve

Education			
NVQ level 5	(base)	(base)	(base)
NVQ level 4	+ve	Ns	Ns
NVQ level 3	+ve	+ve	+ve
NVQ level 2	+ve	+ve	+ve
NVQ level 1	+ve	+ve	+ve
None of these	+ve	+ve	+ve
Overseas qual only	+ve	+ve	+ve
Family structure			
One parent/carer	+ve	+ve	+ve
Mother's age at child's birth			
20 to 29	(base)	(base)	(base)
12 to 19	+ve	+ve	+ve
30 to 39	-ve	-ve	-ve
40 plus	-ve	-ve	-ve
Gender of the child			
Female	-ve	-ve	-ve
Number of siblings			
0	(base)	(base)	(base)
1	-ve	-ve	-ve
2	-ve	-ve	-ve
3+			
Ethnicity			
White	(base)	(base)	(base)
Mixed	Ns	Ns	Ns
Indian	Ns	Ns	-ve
Pakistani and Bangladeshi	+ve	+ve	Ns
Black or Black British	Ns	Ns	Ns
Other Ethnic group (including Chinese, Other)	Ns	Ns	Ns
Relationship with partner - happiness			
coefficient	+ve	+ve	+ve
Importance of stimulation of the child			
Strongly agree	(base)	(base)	(base)
Agree	+ve	+ve	+ve
Neither agree nor disagree	+ve	+ve	+ve
Disagree	Ns	Ns	Ns
Strongly disagree	Ns	Ns	Ns
Neighbourhood effects [base: safe]			
Coefficient	+ve	+ve	+ve

Summary of Findings

A) Maternal Predictors

- Mother smoking is associated with greater child externalising problems of around 2 points.
- Drinking has a negative – albeit small - effect on externalising problems (i.e. drinking more reduces behavioural problems) except if the mother drinks every day. There are few categories for which the coefficients are significant. Heavy drinking is rare.
- Low birth weight and gestational age increases child externalising problems by around 1 points, as expected.
- All measures of depression indicate a large negative association of depression with child behaviour.

LSE Enterprise

- Feeling low or sad: increase of around 1 point.
- Depression (asked to them): increase of around 1-2 points.
- Depression as diagnosed by doctor: increase around 1 points.
- Mental health Kessler scale (continuous from 0 to 24): increase of around 0.2 points for every point on the scale.

B) *Family and Environmental Predictors*

- Overall: most important factors are income and education. There are significant and large negative associations of income with child behaviour.
- This is especially the case for the second and lower quintile, which compared to the highest level, increase child externalising problems by 2-3 points.
- Compared to highest level, lower education increases externalising problems by between 2-3 points.
- Mother not working is negatively associated with externalising problems.
- Children in one parent/carer families experience around 1.5 points more total difficulties.
- Compared to children born to mothers from 20-29 years old, children of younger mothers experience more behavioural difficulties, and the effects are persistent across time. The opposite is true for children born to older mothers.
- Girls have fewer behavioural problems than boys.
- Having 1 or 2 siblings is significantly associated with fewer difficulties than no siblings.
- At ages 5 and 7, being from another ethnic group than white is significantly associated with increased externalising problems only for Pakistani or Bangladeshi children.
- At age 11 this association fades and even reverses, but is no longer significant. At the same time, being Indian becomes associated with fewer externalising problems.
- An increase in externalising problems at ages 5 and 7 is found for children whose parents think stimulation is not important. At age 11, the coefficients are generally lower, as are significance levels.
- The neighbourhood not being safe increases externalising problems.

SDQ – Internalising problems outcome

Internalising problems is the sum of emotional problems and peer problems.

Average values are between 1 and 3.5 on a scale from 0-20.

Table 8: Bivariate Associations between Predictors and SDQ – Internalising problems

	Age 5	Age 7	Age 11
Smoking			
Yes	+ve	+ve	+ve
Drinking			
Never	(base)	(base)	(base)
Less than once a month	-ve	-ve	-ve
1-2 times a month	-ve	-ve	-ve
1-2 times a week	-ve	-ve	-ve
3-4 times a week	-ve	Ns	Ns
5-6 times a week	-ve	Ns	Ns
Every day	Ns	Ns	Ns
Low birth weight			
Yes	+ve	+ve	+ve
Gestational age pre-term			
Yes	+ve	+ve	+ve
Felt low or sad			
Yes	+ve	+ve	+ve
Depression - long standing illness			
Depression	+ve	+ve	+ve
Other illnesses	+ve	+ve	+ve
Depression - as diagnosed by Dr.			
Yes	+ve	+ve	+ve
Mother's mental health - Kessler			
coefficient	+ve	+ve	+ve
Mother's mental health - Kessler - categorical			
0 - low	(base)	(base)	(base)
1	+ve	+ve	+ve
2	+ve	+ve	+ve
3	+ve	+ve	+ve
4	+ve	+ve	+ve
5	+ve	+ve	+ve
6	+ve	+ve	+ve
7 - high	+ve	+ve	Ns
Income Quintiles			
Highest quintile	(base)	(base)	(base)
Fourth quintile	+ve	+ve	+ve
Third quintile	+ve	+ve	+ve
Second quintile	+ve	+ve	+ve
Lowest quintile	+ve	+ve	+ve

Whether in work or not			
No	+ve	+ve	+ve
Education			
NVQ level 5	(base)	(base)	(base)
NVQ level 4	+ve	+ve	Ns
NVQ level 3	+ve	+ve	+ve
NVQ level 2	+ve	+ve	+ve
NVQ level 1	+ve	+ve	+ve
None of these	+ve	+ve	+ve
Overseas qual only	+ve	+ve	+ve
Family structure			
One parent/carer	+ve	+ve	+ve
Mother's age at child's birth			
20 to 29	(base)	(base)	(base)
12 to 19	+ve	+ve	+ve
30 to 39	+ve	-ve	-ve
40 plus	Ns	Ns	+ve
Gender of the child			
Female	Ns	-ve	Ns
Number of siblings			
1	-ve	-ve	-ve
2	-ve	-ve	-ve
3+	Ns	Ns	Ns
Ethnicity			
White	(base)	(base)	(base)
Mixed	+ve	+ve	Ns
Indian	+ve	Ns	Ns
Pakistani and Bangladeshi	+ve	+ve	+ve
Black or Black British	+ve	Ns	Ns
Other Ethnic group (inc Chinese,Other)	Ns	Ns	Ns
Relationship with partner - happiness			
coefficient	+ve	+ve	+ve
Importance of stimulation of the child			
Strongly agree	(base)	(base)	(base)
Agree	+ve	+ve	+ve
Neither agree nor disagree	+ve	+ve	Ns
Disagree	+ve	Ns	Ns
Strongly disagree	Ns	Ns	-ve
Neighbourhood effects [base:safe]			
coefficient	+ve	+ve	+ve

Summary of Findings

A) Maternal Predictors

- If the mother smokes, child internalising problems increase around 0.5-1 points.

LSE Enterprise

- Drinking has a negative – albeit small - effect on internalising problems (i.e. drinking more reduces behavioural problems) except for if the mother drinks every day. At age 7, even when the mother drinks every day, the effect of drinking on behavioural problems is negative, but not significant. Few mothers report drinking a lot and so categories with heavy drinking are not significant.
- Low birth weight and pre-term gestational age increases child internalising problems by around 0.5 points.
- All measures of depression have strong and negative effects on child internalising problems.
    - Feeling low or sad: increase of around 1 point.
    - Depression (asked to them): increase of around 1-2.5 points.
    - Depression as diagnosed by doctor: increase between 0.5-1 points.
    - Mental health Kessler scale (continuous from 0 to 24): increase of 0.2 points for each point on the problem scale.

B) *Family and Environmental Predictors*

- There are significant effects of income, with lower income affecting negatively internalising behaviour. This is especially the case for the second and bottom quintile, which, compared to the highest level, increases child internalising problems by 1-2 points.
- Having no education or level 1 NVQ, compared to the highest level increases internalising behaviour scores by between 1 and 2 points.
- Mother not working is associated with increased internalising problems.
- Children in one parent/carer families experience around 1 point more of internalising problems.
- Compared to children born to mothers aged 20-29 years old, children of younger mothers' experience more difficulties, and the effects are stronger at ages 5 and 7 than at age 11. The opposite is true for children born to older mothers. The coefficients for mothers between 30-39 years-old are constant across child ages. Instead, for 40 plus mothers, coefficients are not significant.
- Girls are associated with fewer internalising problems than boys but only at age7.
- Having 1 or 2 siblings is significantly associated with fewer internalising problems than no siblings. Having three or more siblings is not significantly different to havingno siblings.
- At age 5, being from a minority ethnic group is associated with increased internalising problems; especially for Pakistani or Bangladeshi children. And at ages 7 and 11, the association is only significant for Pakistani and Bangladeshi children. At ages 5 and 7 the effects is around 1.5 points and at age 11 it is 0.6.
- Children of very unhappy couples are more likely to have internalising problems.
- Parental attitudes towards child stimulation have inconsistent associations with internalising problems. At age 5, we observe more internalising problems for children whose parents think stimulation is not relevant (an effect size of around 1 point for those who strongly disagree). At age 7, the relationship is similar but not statistically significant and at Age 11 the size of the effect as well as its significance is smaller. Those who strongly disagree with the statement seem to be associated with around 1 point less in internalising problems and significance level is 5%.
- The neighbourhood not being safe is associated with more internalising problems.

LSE Enterprise

SDQ – Hyperactivity/Attention Deficit Problems Outcome

Hyperactivity/attention deficit problems are measured in a 10-point scale, from lower to higher problems. Average values are around 3 on a scale of 0 -10.

Table 9: Bivariate Associations between Predictors and SDQ- Hyperactivity/Attention Deficit Problems

	Age 5	Age 7	Age11
Smoking			
yes	+ve	+ve	+ve
Drinking			
Never	(base)	(base)	(base)
Less than once a month	-ve	Ns	Ns
1-2 times a month	Ns	Ns	-ve
1-2 times a week	-ve	-ve	Ns
3-4 times a week	Ns	Ns	Ns
5-6 times a week	-ve	Ns	Ns
Every day	Ns	Ns	Ns
Low birth weight			
Yes	+ve	+ve	+ve
Gestational age preterm			
Yes	+ve	+ve	+ve
Felt low or sad			
Yes	+ve	+ve	+ve
Depression - long standing illness			
Depression	+ve	+ve	+ve
Other illnesses	+ve	+ve	+ve
Depression - as diagnosed by Dr.			
Yes	+ve	+ve	+ve
Mother's mental health - Kessler			
coefficient	+ve	+ve	+ve
Mother's mental health - Kessler - categorical			
0 - low	(base)	(base)	(base)
1	+ve	+ve	+ve
2	+ve	+ve	+ve
3	+ve	+ve	+ve
4	+ve	+ve	+ve
5	+ve	+ve	+ve
6	+ve	+ve	+ve
7 – high	+ve	Ns	Ns
Income Quintiles			
Highest quintile	(base)	(base)	(base)
Fourth quintile	+ve	+ve	+ve
Third quintile	+ve	+ve	+ve
Second quintile	+ve	+ve	+ve
Lowest quintile	+ve	+ve	+ve
Whether in work or not			
No	+ve	+ve	+ve

LSE Enterprise

Education	36		
NVQ level 5	(base)	(base)	(base)
NVQ level 4	Ns	Ns	Ns
NVQ level 3	+ve	+ve	+ve
NVQ level 2	+ve	+ve	+ve
NVQ level 1	+ve	+ve	+ve
None of these	+ve	+ve	+ve
Overseas qual only	+ve	+ve	+ve
Family structure			
Two parents/carers	(base)	(base)	(base)
One parent/carer	+ve	+ve	+ve
Mother's age at child's birth			
20 to 29	(base)	(base)	(base)
12 to 19	+ve	+ve	+ve
30 to 39	-ve	-ve	-ve
40 plus	-ve	-ve	Ns
Gender of the child			
Male	(base)	(base)	(base)
Female	-ve	-ve	-ve
Number of siblings			
0	(base)	(base)	(base)
1	-ve	-ve	-ve
2	-ve	-ve	-ve
3+	-ve	-ve	-ve
Ethnicity			
White	(base)	(base)	(base)
Mixed	Ns	Ns	Ns
Indian	Ns	Ns	-ve
Pakistani and Bangladeshi	+ve	+ve	Ns
Black or Black British	Ns	-ve	Ns
Other Ethnic group	Ns	Ns	Ns
Relationship with partner - happiness			
coefficient	+ve	+ve	+ve
Importance of stimulation of the child			
Strongly agree	(base)	(base)	(base)
Agree	+ve	+ve	+ve
Neither agree nor disagree	+ve	+ve	+ve
Disagree	Ns	Ns	Ns
Strongly disagree	Ns	Ns	Ns
Neighbourhood effects [base:safe]			
coefficient	+ve	+ve	+ve

Summary of Findings

A) Maternal Predictors

- If the mother smokes, child behavioural problems are around 1 point higher.
- Results for mother's drinking are the same as previous problem scales.

- Low birth weight and pre-term gestational age are associated with around 0.5 more points on the hyperactivity scale.
- As with the other problem scales, all measures of maternal depression indicate a strong negative effect on child hyperactivity.
  - Feeling low or sad: increase of around 0.6 point.
  - Depression (asked to them): increase of around 1 points.
  - Depression as diagnosed by doctor: increase around 0.5 points.
  - Mental health Kessler scale (continuous from 0 to 24): increase of around 0.1 points for each point on the scale.

B) *Family and Environmental Predictors*

- There is a large and significant effect of income, with lower income negatively associated with child behaviour.
- Compared to highest educational level child hyperactivity problems are higher by more than 1 point when their mother is educated to the lowest and second-lowest educational categories.
- Mother not working is associated with higher hyperactivity scores, and the size of association is large.
- Children in one parent/carer families experience around 1 point more of hyperactivity problems.
- Compared to children born to mothers aged 20-29 years old, children of younger mothers experience more behavioural difficulties, and the effects are persistent across ages. For those children born to mothers aged 40+, the effect is positive but fades away with the age of the child.
- Girls are associated with fewer hyperactivity problems than boys.
- Having siblings is significantly associated with lower hyperactivity scores than having no siblings.
- At ages 5 and 7, Pakistani or Bangladeshi children have higher rates problem scores.
- At age 11 this association fades and even reverses, but is not statistically significant.
- Children of very unhappy couples are more likely to exhibit higher levels of hyperactivity.
- Parents' agreement that child stimulation is important has a differing effect over ages. At ages 5 and 7, children whose parents think stimulation is not relevant have more problems, while at age 11 the coefficients are smaller in size. Note however, that there are only small numbers of parents who express disagreement with the statement.
- Living in a neighbourhood deemed not safe increases hyperactivity problems.

Social Outcomes

Bullying outcomes

We use child self-reported experience of bullying at ages 7 and 11, focusing on those who responded that they were bullied 'all the time' (age 7) or 'most days' (age 11).

Table 10: Bivariate Associations between Predictors and Bullying Outcomes

	Bullied 'all of the time' age 7	Bullied 'most days' age 11
	Age 5	Age 7
BMI	-ve	Ns
BMI squared	+ve	Ns
BAS Naming Vocabulary T-score	-ve	-ve
Externalising behaviour	+ve	+ve
Internalising behaviour	+ve	+ve
Peer problems	+ve	+ve
Mother's mental health - Kessler	+ve	+ve
Month of birth		
September	(base)	(base)
October	Ns	+ve
November	Ns	Ns
December	Ns	Ns
January	Ns	Ns
February	Ns	Ns
March	Ns	Ns
April	+ve	+ve
May	Ns	Ns
June	+ve	+ve
July	Ns	+ve
August	Ns	+ve
Female	+ve	Ns
Child's ethnicity		
White	(base)	(base)
Mixed	Ns	Ns
Indian	Ns	+ve
Pakistani and Bangladeshi	+ve	Ns
Black or Black British	Ns	Ns
Other Ethnic group (inc Chinese,Other)	Ns	Ns
Number of siblings		
None	(base)	(base)
1	+ve	Ns
2	Ns	Ns
3+	+ve	Ns
Income quintile		
Highest	(base)	(base)
Fourth quintile	Ns	Ns
Third quintile	+ve	Ns
Second quintile	+ve	+ve
Lowest quintile	+ve	+ve

Mother's highest level of education		
None of these	(base)	(base)
NVQ level 1	Ns	Ns
NVQ level 2	+ve	Ns
NVQ level 3	+ve	+ve
NVQ level 4	+ve	+ve
NVQ level 5	+ve	Ns
Overseas qual only	+ve	Ns
Mother work status		
In work	(base)	(base)
Not in work	+ve	+ve
Family structure		
2 parents/carers in household	(base)	(base)
One parent/carer	+ve	+ve
Child weight		
Not overweight	(base)	(base)
Overweight	Ns	Ns
Obese	+ve	+ve
Child has longstanding illness	+ve	+ve
Special educational needs	(not available in sweep)	+ve
How often mother shouts at child when naughty		
Never	(base)	(base)
Rarely	Ns	Ns
Sometimes (about once a month)	Ns	Ns
Often (about once a week or more)	Ns	Ns
Daily	+ve	Ns

Summary of Findings

- Many significant bivariate associations observed at age 7 disappear by age 11: BMI, being female, being Pakistani or Bangladeshi, Mother's education being low, number of siblings, and mother's shouting often at the child.
- All high scores on problem subscales of SDQ increases likelihood of being bullied at both age 7 and age 11. These associations are strong.
- Higher scores on verbal cognitive tests are associated with lower chances of being bullied.
- Living with a single parent, being overweight, having a long standing illness, having low income, and mother's mental health being poor are all associated with a greater likelihood of being bullied in both age groups. These associations have large effect sizes for the risks of being bullied.

Peer Problems Behaviour

This is a subscale from the Strength and Difficulties Questionnaire. Peer problems' average values is around 2 in a scale of 10. Table 11 below reports bivariate associations between predictors and peer problem behaviour.

Table 11: Bivariate Associations between Predictors and SDQ- Peer Problems

	Age 5	Age 7	Age 11
	coefficient	*coefficient*	*coefficient*
Smoking			
yes	+ve	+ve	+ve
Drinking			
Never	(base)	(base)	(base)
Less than once a month	-ve	-ve	-ve
1-2 times a month	-ve	-ve	-ve
1-2 times a week	-ve	-ve	-ve
3-4 times a week	-ve	Ns	Ns
5-6 times a week	Ns	Ns	Ns
Every day	-ve	Ns	Ns
Low birth weight			
Yes	+ve	+ve	+ve
Gestational age preterm			
Yes	+ve	+ve	+ve
Felt low or sad			
Yes	+ve	+ve	+ve
Depression - long standing illness			
Depression	+ve	+ve	+ve
Other illnesses	+ve	+ve	+ve
Depression - as diagnosed by Dr.			
Yes	+ve	+ve	+ve
Mother's mental health - Kessler			
coefficient	+ve	+ve	+ve
Mother's mental health - Kessler - categorical			
0 - low	(base)	(base)	(base)
1	+ve	+ve	+ve
2	+ve	+ve	+ve
3	+ve	+ve	+ve
4	+ve	+ve	+ve
5	+ve	+ve	+ve
6	+ve	+ve	+ve
7 - high	+ve	+ve	Ns
Income Quintiles			
Highest quintile	(base)	(base)	(base)
Fourth quintile	+ve	+ve	+ve
Third quintile	+ve	+ve	+ve
Second quintile	+ve	+ve	+ve
Lowest quintile	+ve	+ve	+ve
Whether in work or not			
No	+ve	+ve	+ve

Education			
NVQ level 5	(base)	(base)	(base)
NVQ level 4	+ve	+ve	Ns
NVQ level 3	+ve	+ve	+ve
NVQ level 2	+ve	+ve	+ve
NVQ level 1	+ve	+ve	+ve
None of these	+ve	+ve	+ve
Overseas qual only	+ve	+ve	+ve
Family structure			
One parent/carer	+ve	+ve	+ve
Mother's age at child's birth			
20 to 29	(base)	(base)	(base)
12 to 19	+ve	+ve	Ns
30 to 39	-ve	-ve	-ve
40 plus	Ns	Ns	Ns
Gender of the child			
Female	-ve	-ve	-ve
Number of siblings			
0	(base)	(base)	(base)
1	-ve	-ve	-ve
2	-ve	-ve	-ve
3+	Ns	-ve	-ve
Ethnicity			
White	(base)	(base)	(base)
Mixed	+ve	+ve	Ns
Indian	+ve	+ve	Ns
Pakistani and Bangladeshi	+ve	+ve	+ve
Black or Black British	+ve	+ve	Ns
Other Ethnic group (inc Chinese,Other)	+ve	+ve	Ns
Relationship with partner - happiness			
coefficient	+ve	+ve	+ve
Importance of stimulation of the child			
Strongly agree	(base)	(base)	(base)
Agree	+ve	+ve	+ve
Neither agree nor disagree	+ve	+ve	Ns
Disagree	Ns	Ns	Ns
Strongly disagree	Ns	Ns	-ve
Neighbourhood effects [base:safe]			
coefficient	+ve	+ve	+ve

Summary of Findings

A) Maternal Predictors

- If Mother smokes, child behavioural problems increase by around 0.5 points.
- Mother drinking during pregnancy is negatively associated with peer problems, but the effect size is small.
- Low both weight and preterm gestational age is associated with child behavioural problems and the size of coefficient is around 0.3 point.

- All measures of depression: asked of the respondent, diagnosed by doctor, evaluated using the Kessler scale, and asking the respondent whether she felt low or sad lately indicate a large and negative association between maternal depression and child's peer problem behaviour.

B) *Family and Environmental Predictors*

- Income is highly and negatively correlated with peer problems.
- Mother having no education or low education compared to highest level is associated with more peer problems.
- Children in one parent/carer families experience around 0.5 point more of total difficulties.
- Compared to children born to mothers aged 20-29 years old, children from younger mothers experience more difficulties, and the effects are larger at age 5 and 7 than in age 11.The opposite is true for children born to older mothers. The coefficient for mothers between 30-39 years old stays constant across child ages. But for mothers aged 40 plus, no associations are significant.
- Girls have fewer peer problems than boys.
- Having 1 or 2 siblings is significantly associated with fewer peer problems than no siblings. Instead, having three or more siblings is barely significantly different to having no siblings.
- At ages 5 and 7, peer problems are greater for those from a minority ethnic group, especially for Pakistani or Bangladeshi children.
- By age 11 the association is only significant for Pakistani or Bangladeshi children, and with a lower coefficient than in other sweeps at around 0.4 points.
- Children of parents who are unhappy in their relationship have higher rates of peer problems.
- An increase in peer problems for children is observed for those whose parents think child stimulation is not important.
- The neighbourhood not being safe is associated with more peer problems especially at ages 7 and 11.

LSE Enterprise

4.2. Multivariate Models

Above, we demonstrated a large number of bivariate associations between predictor variables and outcomes. However, in many cases they may be confounded with each other. For example, lone parenthood effects may work through lower income or through poorer maternal mental health. Therefore in this section, we turn to multivariate analysis to investigate the net effect of predictors of interest controlling for the others. In the tables that follow we summarise the direction, significance and size of the effects in stylized form for the convenience of the reader. This makes it possible to see at a glance, which factors are independently associated with the outcome measures, at which ages and how strongly. Full tables for all models at all ages are provided in the Appendix.

Health Outcomes

BMI and Obesity

Table 12 summarises the multivariate models relating to BMI and the risk of being overweight. One or more + symbols indicates that the variable is positively associated with BMI / overweight, e.g. is linked to higher BMI and a greater risk of overweight, while a – sign indicates the relationship is in the opposite direction. The number of +s and –s indicate the size of the association, with more indicating a bigger effect, as discussed in the note to the table. Finally, cells are left blank where no statistically significant association was found.

Table 12: Summary of Multivariate Analyses for Health Outcomes

	BMI			Overweight		
MATERNAL PREDICTORS	**Age 5**	**Age 7**	**Age 11**	**Age 5**	**Age 7**	**Age 11**
Birth Weight	+ (girls)			+ (girls)		
Month of birth						
Smoked during pregnancy	+	+(boys) ++(girls)	++		+	+
Maternal weight before pregnancy	+	+	+	+ (girls)	+ (girls)	+ (girls)
Working mother					- (boys)	
Working through pregnancy						
Mothers age at birth						
Maternal BMI at sweep 1	+	+	+	+	+	+
Maternal BMI at interview						
SOCIO-ECONOMIC FACTORS						
Family structure: One-parent family / sweep 1						
Income Quintile is below the highest sweep 1						
Income Quintile is below the highest sweep 3				+(boys)		
Breastfeeding up to 6 months (Base = no breastfeeding)					- (Boys)	
Breastfeeding more than 6 months (Base = no breastfeeding)						
Mother's Education Level is Lowest					+ (girls)	+(girls)
Mothers age at birth				+		
Regular bed times						
Exercise with child (Frequency)		(n/a)				
Child exercise (for age 7)		- - - (girls)				
Breakfast (Frequency)						+++ (boys)
Fruits and vegetables (Portions a day)						

Note: Table above provides a summary of variables that are significant in the multivariate models balanced samples for each outcome variable. See unbalanced samples in the appendix.

Number of plus or minus signs indicate the size of coefficients in absolute terms according to the following criteria: Coefficients with an absolute size $0<b<0.4$ is shown as + or - , Coefficients with an absolute size $0.3<b<1$ is shown as ++ or - - , Coefficients with an absolute size $1<b$ are shown as +++ or - - -.

The predictors that are not significant in the multivariate models are represented with an empty cell. Multivariate regression output tables for all models with the estimated coefficients (marginal effects) are provided in the Appendix.

Summary of Findings

- Few predictors are associated with BMI and overweight for both boys and girls and at all ages.
- Maternal weight before pregnancy is associated with higher child BMI at all ages but the effect is larger among girls. Maternal BMI at age nine months is consistently significant too and larger for girls (0.07) than for boys (0.05) for all ages.
- Smoking during pregnancy is associated with higher likelihood of obesity and overweight at ages 7 and 11.
- Socio-economic status at age 9 months is significant at age 7 for girls' BMI and boys' overweight
- Mother's age and maternal employment, along with single parenthood do not influence child's BMI.
- As in other models, the effect of eating habits is either non-significant or not robust, and we find as in the bivariate correlations a protective effect of moderate physical exercise on girls BMI.
- Lack of physical exercise is associated with higher levels of BMI for girls at age 7.

**Interpretation of Findings and Effect Size**

The models explain about 18- 12% of the total variation in the outcome measure, which is in line with the level of variation typically explained in models using micro data. i.e. typically more than 80 per cent of variance is linked to factors that it is impossible to measure, that are unobserved or are simply due to idiosyncratic individual variation.

To put the effect sizes in perspective, we plot below the predicted values for overweight at age 5 using the full multivariate specification described above. One of the main determinants of overweight with the largest effect sizes is maternal BMI. Whilst overweight probability at age 5 is 20% (0.2) for those whose mother has a BMI of 25, it is 30% for boys whose mothers' BMI is 35. The effect among girls is even steeper as chances of being overweight for girls with a mother with a BMI 35 are 40% (0.4).

Figure 1: Predicted values of child overweight by maternal BMI, Girls
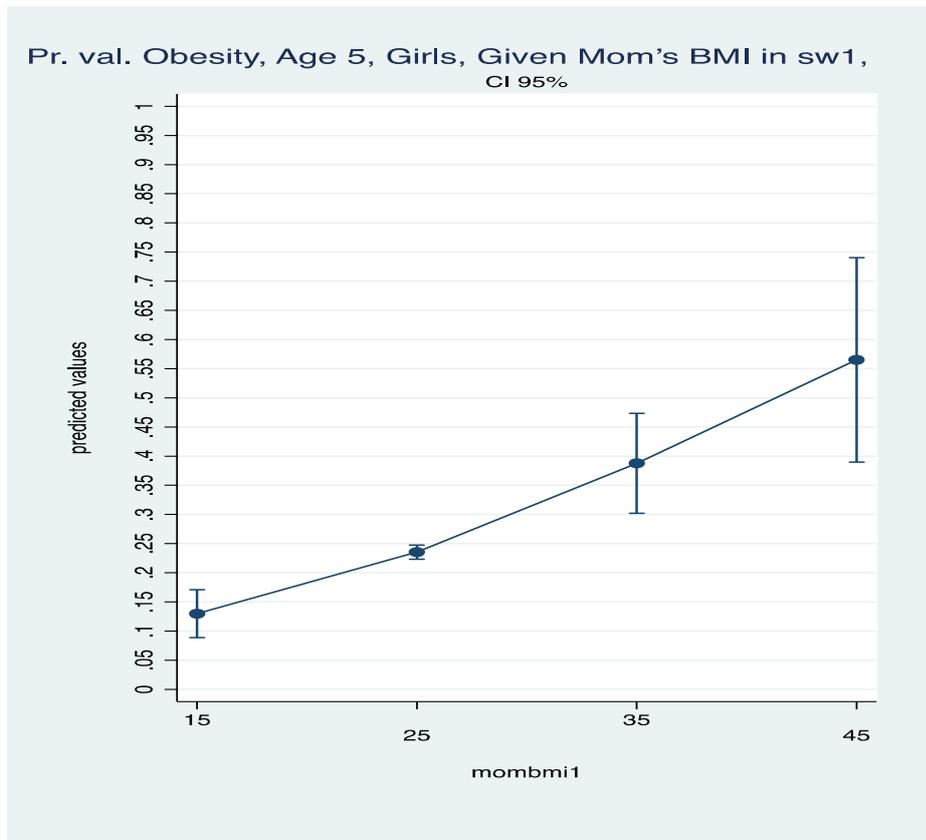
Figure 2: Predicted values of child overweight by maternal BMI, boys

LSE Enterprise
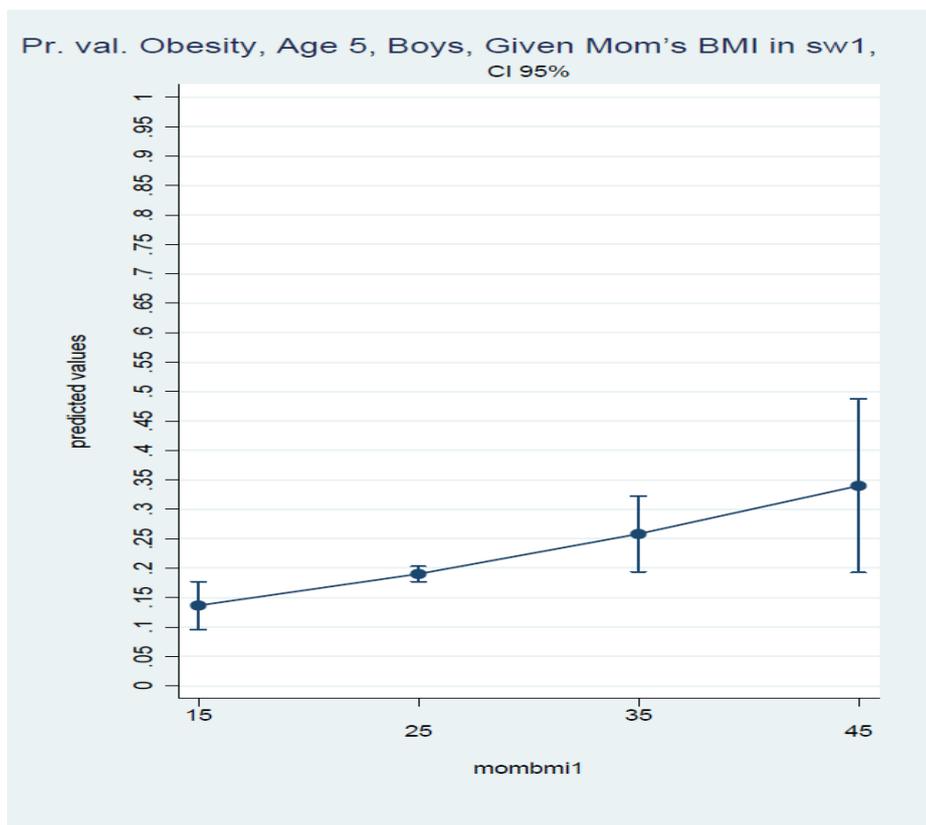
Self-Reported Child Health

We next turn to SRCH. Table 13, summarises, in a similar way to Table 12, the significant independent associations between the different factors and child health. Note that all these associations reported below are now conditional on all the other measures in the model – they represent the net or independent effect of the predictors.

Table 13: Summary of Multivariate Results for Self-Reported Child Health

	Self-reported health		
MATERNAL PREDICTORS	**Age5**	**Age7**	**Age 11**
Maternal Smoking	-		
Drinking (Every day)	-		
Low Birth Weight	-	-	-
Felt low or sad	-	-	
Breastfeeding			
Month of Birth			
Mothers' health	+		
Mother's long standing illness	-		
Mother's mental health		-	
SOCIO-DEMOGRAPHIC FACTORS	**Age5**	**Age7**	**Age 11**
Income Quintile is below the highest sweep 1	+	+	+
Mother's Education Level is Lowest			+
Mothers age at birth			
Female	+	+	
Ethnicity – Indian		-	-
Ethnicity – Pakistani and Bangladeshi	--	-	-
Siblings			
Headaches	-	-	-
Child picked on/Bullied	-	-	-
Child exercise (Frequency)	n/a	+	+

Note: Table above provides a summary of variables that are significant in the multivariate models balanced samples for each outcome variable See unbalanced samples in the appendix. Number of plus or minus signs indicate the size of coefficients in absolute terms according to the following criteria:
Coefficients with an absolute size 0<b<0.4 is shown as + or - ; Coefficients with an absolute size 0.3<b<1 is shown as ++ or - - ; Coefficients with an absolute size 1<b is shown as +++ or - - -
The predictors that are not significant in the multivariate models are represented with an empty cell.
Multivariate regression output tables for all models with marginal effects are provided in the Appendix.

Summary of Findings

- Low birth weight exhibits a significant effect at all ages and associated with worse child health, especially at age 11 where the coefficient shows a coefficient of -0.20.
- Maternal physical health is again a significant predictor of child health at age 5 with coefficients being the largest (0.43-45 at age 5).
- Higher income is consistently associated with better self-reported health
- Girls have better health than boys (coefficient being 0.05-0.06), but the effect does not survive to the age of 11.

- Pakistani children systematically exhibit lower self-reported health, although the coefficient declines from -0.42 at age 5, to -0.18 at age 11. The health of children of Indian ethnicity is lower compared to whites but only at age 7 and 11, but the effect compares to that of Pakistani children at those ages.
- Psychosomatic symptoms or being picked on by other children is significantly independently associated with worse child health at all ages, and the effect remains over child age, exhibiting a coefficient of -0.18.
- The effect of being bullied reduces child health by a coefficient of –0.08 at age 5, but exhibits the largest effect -0.018-0.019 at age 7 and remains at age 11 with a coefficient of -0.15
- Exercising 'several times a week' rather than 'less than once a year/never' is associated with better child health. However, frequency of child exercise is significantly associated with better child health at ages 7 and 11, as is eating more portions of fruit per day. The effect of exercise five times a week is to increase the coefficient by 0.20 and exercise three times a week increases the coefficient by 0.18.

Interpretation of Findings and Effect Size

Figure 3 below plots the predicted values of self-reported health at age 11 on changes in maternal self-reported health. The figure shows a large effect of a change from fair maternal health to excellent health of about 0.5 score points, given the high overall level of reported child health. Even though the coefficients are significant the overall effect of all other predictors on the values of self-reported health are very small [See figures for other predictors in the Appendix].

Figure 3: Predicted values of self-reported child health by mother's health
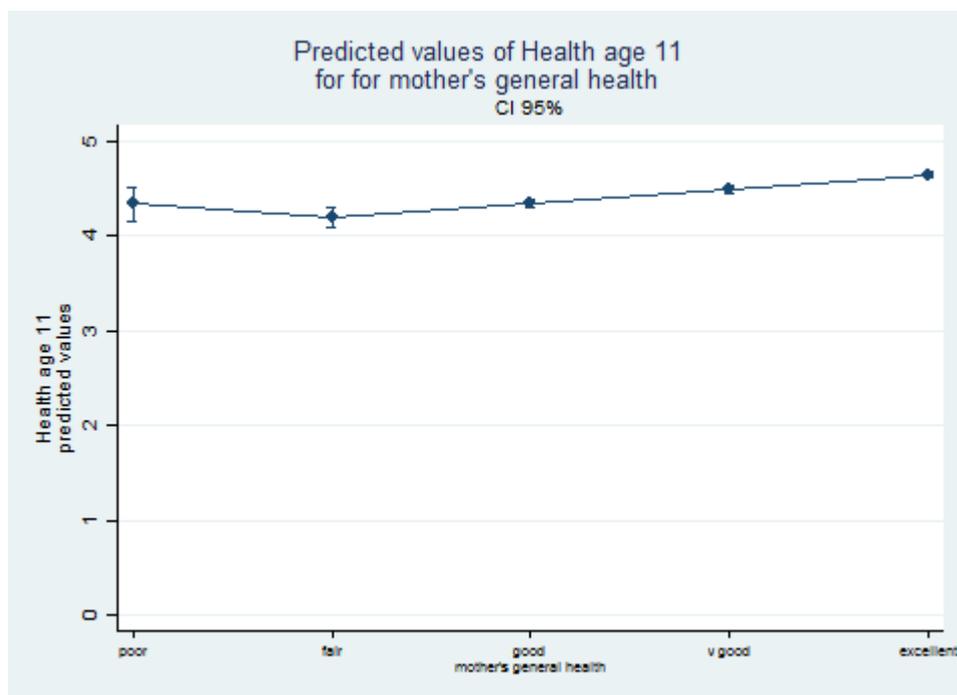
Cognitive Outcomes

We now turn to cognitive outcomes. Table 14 summarises the independent effects of the various predictors on child scores the various tests carried out age ages 5, 7 and 11. As with the bivariate results, given the differences in the tests at the different ages, we summarise the results by age.

Table 14: Summary of Multivariate Results for Cognitive Outcomes

	Age 5			Age 7			Age 11
MATERNAL PREDICTORS	**Naming Vocab.**	**Pattern Construc.**	**Picture Similar.**	**NFER MATH Skills**	**Pattern Construc.**	**Word Reading**	**Verbal Similarity**
Married/Cohabiting at birth							---
Birth Weight							+
Child care sweep 1							
Raised blood and preeclampsia							
Bleeding and thread to miscarriage		---					
SOCIO-ECONOMIC FACTORS	**Naming Vocab.**	**Pattern Construc.**	**Picture Similar.**	**NFER MATH Skills**	**Pattern Construc.**	**Word Reading**	**Verbal Similarity**
Breastfeeding up to 6 months			+++	+++	+++	+++	+++
Income at age 9m	+++	++	+++	+++	+++	+++	+++
Income at preceding survey	+++	++	+++	+++		+++	+++
Work status during pregnancy	+++					+++	+++
Work at age 9m	---			---	---	---	---
Work at birth	+++				+++		
Maternal Education	+++	+++	+++	+++	+++	+++	+++
Fam. structure age 9m							
Fam. structure age 3y						---	
Mothers age at birth				-	-		
Female	+++	+++	+++		++	+++	
Child care							+
Reception				+++		+++	

Note: Table above provides a summary of variables that are significant in the multivariate models balanced samples for each outcome variable See unbalanced samples in the appendix. Number of plus or minus signs indicate the size of coefficients in absolute terms according to the following criteria:
    Coefficients with an absolute size 0<b<0.4 is shown as + or -
    Coefficients with an absolute size 0.3<b<1 is shown as ++ or - -
    Coefficients with an absolute size 1<b is shown as +++ or - - -
The predictors that are not significant in the multivariate models are represented with an empty cell. Multivariate regression output tables for all models with the estimated coefficients (marginal effects) are provided in the Appendix.

Age 5

- Girls exhibit better naming vocabulary skills than boys by 1.2 points on the overall score; maternal work during pregnancy is associated with 1.9 points on vocabulary scores, net of other factors.
- Children of mothers who were working at age 9 months exhibit lower vocabulary skills by one unit on the score. Lower socio-economic status in the first and second sweep reduces monotonically vocabulary scores. Lower maternal education reduced naming vocabulary scores. However, working at time of the interview increases naming vocabulary scores.
- When we examine the predictors of picture similarity two variables emerge as important, namely socio-economic status at age nine months and age 3, and breastfeeding to 6 months. The latter increases picture similarity score skills by 1.36 in the balanced sample. Finally, low maternal education exerts a negative independent effect on picture similarity scores.
- On pattern construction skills the main significant determinants are socio-economic status at age nine months and age 3 and low maternal education and breastfeeding until age 6.

Age 7

- Breastfeeding to more than six months increases maths scores by 2.1 to1.75 points. Other very important influences are higher socio-economic status at age9 months and age 5 and maternal education. The lowest level of maternal education is associated with an 8 point reduction in maths score. Children of working mothers at sweep 1 exhibit a disadvantage in maths score, and having started school at reception is probably the single most important variable increasing scores by 17 points.
- When we examine pattern construction skills we find that breastfeeding up to 6 months increases skills. Low socio-economic status reduces pattern construction scores and so does having a working mothers at age nine months - by 1.38 score points. In contrast, mother working during pregnancy increases word reading by 2.6 score points.
- Low socio-economic status reduces reading scores, and specifically the lowest income quintile reduced scores by 8 score points. Similarly, maternal education improves reading scores.
- Finally, single parenthood at sweep 3 reduces scores by 1.7 and starting school at reception increases scores by 6.8 scores in the balanced sample.

Age 11

- Birth weight is associated with higher verbal similarity by 0.011 points per pound.
- Children whose mothers worked during pregnancy exhibit 1.2 score points increase in verbal similarity.
- Low socio-economic status at ages nine months and 7 years has a very significant effect, lowest socio-economic status at age nine months is associated with 3 points lower scores, and economic status at age 7 is associated with 2 points lower scores.
- Breast feeding up to 6 months is associated with a 1 point higher score in verbal similarities at age 11.
- Lowest level of maternal education is associated with lower scores by 5.3 points.

Interpretation of Findings and Effect Sizes

Maternal education has a clear impact on her child's verbal abilities, according to our multivariate analyses described in the table above. Figure 4 shows that the predicted score is 60 points for children of higher educated women and 55 for children of women with the lowest education attainment. This is approximately a 2 per cent difference in the test score outcomes (and around half a standard deviation).

Figure 4: Predicted values of verbal similarities score by maternal education

LSE Enterprise
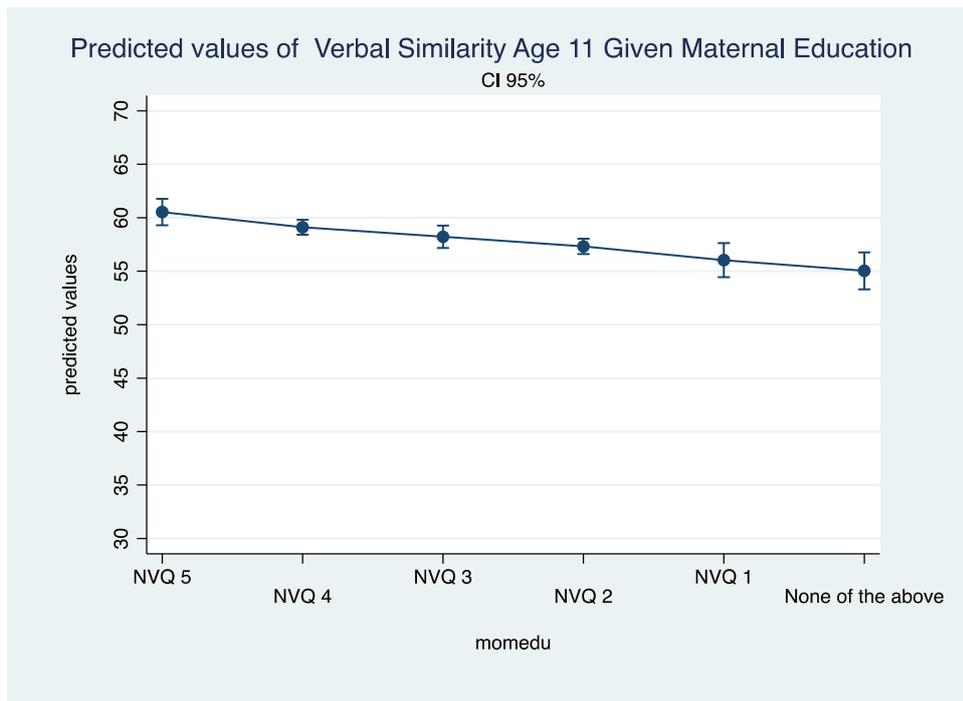
Multivariate results for KS2 Outcomes

Table 15 summarizes the results for the four KS2 outcomes. As it illustrates, the independent predictors of English (and reading) test scores are somewhat different to those for Maths (and arithmetic).

Table 15: Summary of multivariate results for KS2 outcomes

VARIABLES	Reading	English	Maths	Arithmetic
Female	+++	+++	---	--
Birth Weight			+	+
Months of birth	+++	+++	+++	+++
Childcare at Sweep 1 Base- No child care at sweep 1			+++	+++
Working At Pregnancy Base- Not working at pregnancy				
Working At Sweep 1 Base – Not working at sweep 1	---	---	---	
One Parent Carer At Sweep 1 Base -More than one parent at sweep 1				
Married Or Cohabiting At Birth Base- Not married at sweep 1				
Deprivation Index Mult Deprivation 10-20% Sweep 1 Index Mult Deprivation 20-30% Sweep 1 Index Mult Deprivation 30-40% Sweep 1 Index Mult Deprivation 40-50% Sweep 1 Index Mult Deprivation 50-60% Sweep 1 Index Mult Deprivation 60-70% Sweep 1 Index Mult Deprivation 70-80% Sweep 1 Index Mult Deprivation 80-90% Sweep 1 Index Mult Deprivation Highest Decile Sweep 1 (Base Highest)				
Second Highest Income Quintile age 9m	---	---		--
Third Highest Income Quintile Sweep 1	---	---		---
Fourth Highest Income Quintile Sweep 1	---	---		---
Fifth Highest Income Quintile Sweep 1	---	---		---
Breastfeeding period				
Base - No Breastfeeding				
Breastfeed Between 1 And 3 Months				
Breastfeed Between 13 Weeks And 6 Months	+++	+++	+++	++
Breastfeed More Than 6 Months				
Health conditions at pregnancy Bleeding And Threat Of Miscarriage Early In Pregnancy Raised Blood Pressure, Eclampsia/Preeclampsia Or Toxaemia Other Illness During Pregnancy				

LSE Enterprise

|				
Deprivation Index Mult Deprivation 10-20% Sweep 4 Index Mult Deprivation 20-30% Sweep 4 Index Mult Deprivation 30-40% Sweep 4 Index Mult Deprivation 40-50% Sweep 4 Index Mult Deprivation 50-60% Sweep 4 Index Mult Deprivation 60-70% Sweep 4 Index Mult Deprivation 70-80% Sweep 4 Index Mult Deprivation 80-90% Sweep 4 Index Mult Deprivation Highest Decile – Base Sweep 4	+	+		
Income quantile sweep 4 Highest Quantile –Base Second Highest Income Quintile Sweep 4 Third Highest Income Quintile Sweep 4 Fourth Highest Income Quintile Sweep 4 Fifth Highest Income Quintile Sweep 4			-- --- ---	-- --- ---
Maternal Education National Vocational Qualification 5 - Base National Vocational Qualification 4 National Vocational Qualification 3 National Vocational Qualification 2 National Vocational Qualification 1 None Of The Above	--- --- --- ---	--- --- --- ---	--- ---	--- ---
One Parent Carer At Sweep 4 More than one parent- Base				
Working Mother No working mother - Base				
Mothers Age	+	+	+	+
Whether Child Attends After School Club Does not attend school club - Base				

Summary of Findings

- Multivariate regressions suggests consistently with the bivariate findings that girls do better in English and reading and boys in maths and arithmetic
- Birthweight only increases school results in maths and arithmetic, and maternal work at age nine months and women's age are linked to lower English and reading scores but not for mathematics and arithmetic's.
- Low maternal education and low income at nine months are associated with lower English, reading, arithmetic's, reading and mathematics scores.
- Breastfeeding between 13 weeks to 6 months is associated with higher KS2 scores across the board.
- Low income at age 7 is linked to lower test scores except for English

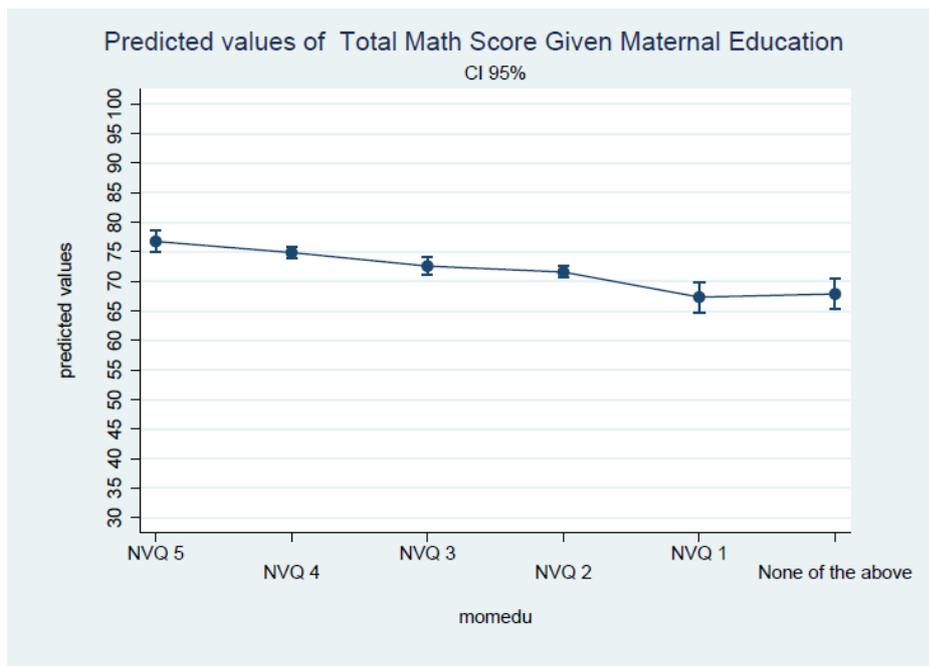**Interpretation of Findings and Effect Sizes**

Largest effects sizes on KS2 outcomes are again observed for maternal education, according to our multivariate analyses described in the table above. Figure 5 shows that the predicted score for total English is 79 points for children of the highest educated women and around 70 points for children of women with the lowest education attainment.

This is approximately a one standard deviation difference in test scores. Figure 6 shows a similar pattern for maths.

Figure 5: KS2 English test scores by maternal education

Predicted values of Total English Score Given Maternal Education
CI 95%
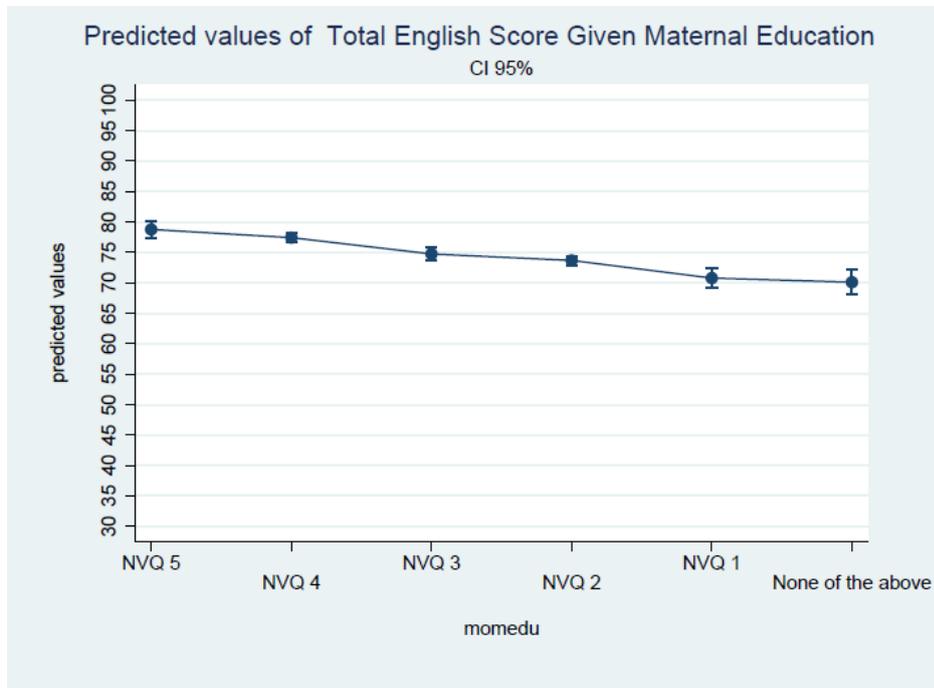
Figure 6: KS2 Maths scores by maternal education

Predicted values of Total Math Score Given Maternal Education
CI 95%

Behavioural Outcomes

Table 16: Results from multivariate analysis of Behavioural Outcomes (SDQ)

	Total difficulties			Internalising problems			Externalising problems			Hyperactivity problems		
MATERNAL PREDICTORS	Age 5	Age 7	Age 11	Age 5	Age 7	Age 11	Age 5	Age 7	Age 11	Age 5	Age 7	Age 11
Smoking during Pregnancy	++	++	++				++	++	++	++	++	++
Drinking during Pregnancy												
Low Birth Weight	++	++	++	+	++	+	++	++	++	++	++	++
Premature birth	++			++			++			+		
Mother feeling low or sad	++	++	++	+	+	+	+	+	+	+	+	+
Depression Diagnosed						++						
Mother's mental health not well [Kessler]	+	+	+	+	+	+	+	+	+	+	+	+
SOCIO-ECONOMIC FACTORS												
Income Quintile is below the highest	++ ++(the lowest)	++ ++(the lowest)	++ ++(the lowest)	+	+	+	++	++	++	+	+	+
Mother is NOT Working			++	+	+	+						
Mother's Education Level is Lowest	++	++	++				++	++	++		+	+
Family structure: One-parent family							++					
Teenage mother												
Mothers age at birth is 30-39	--	--	--	-	-	-	-	- -	- -	-	-	-
Mothers age at birth is 40+							- -	- -	- -	- -		
Gender child: Female	- - -	- - -	- - -	-	-	-	--	- - -	- - -	- -	- -	- -
Number of siblings [ref: no child]	- -	- - -	- -	- -	- -	- -			--	- -	- -	- -
Ethnicity: Indian			- - -						- -			--
Ethnicity: Pakistani or Bangladeshi	+++		- -	+++								--
Ethnicity: Black or Black British		- -						- -				
Mother agrees stimulation important	++	++	++	+			++-	+-		+	+	
Mother strongly disagrees stimulation important				---	---	---						
Neighbourhood is NOT safe	+	++		+	+	+	+	+	+	+	+	+

Note: Table above provides a summary of variables that are significant in the multivariate models balanced samples for each outcome variable See unbalanced samples in the appendix. Number of plus or minus signs indicate the size of coefficients in absolute terms according to the following criteria: Coefficients with an absolute size 0<b<0.4 is shown as + or - , Coefficients with an absolute size 0.3<b<1 is shown as ++ or - - , Coefficients with an absolute size 1<b is shown as +++ or - - - .The predictors that are not significant in the multivariate models are represented with an empty cell. Multivariate regression output tables for all models with the estimated coefficients (marginal effects) are provided in the Appendix.

Summary of Findings

In the following summary, the predictors from multivariate models where dependent variables are "SDQ-total difficulties scale" is discussed. However, when appropriate, we discuss how these predictors performed when we look at component subscales. The full tables and comparison tables from the balanced panels can be found in the Appendix.

- Smoking during pregnancy is associated with poorer child behavioural outcomes, especially higher total difficulties score, externalising behaviour and hyperactivity but not with more internalising problems.
- Low birth weight affects negatively mostly externalising behaviour and specifically hyperactivity/attention deficit problems.
- Premature birth only has an effect on overall effect on total difficulties and only at age 5. Possibly driven both by internalising and externalising problems.
- Feeling low or sad has a negative effect on total difficulties.
- Mother's depression has only significant internalising problems at age 11 at the 10% significance level.
- Mother's mental health being poor increases total difficulties. Each problem scale is affected in a similar way and the effect is consistent across time.
- The lower the income, the more total difficulties experienced by child - and stronger effect, especially for second and lowest quintile. The effect is present for all levels of income below the highest quintile for internalising problems. Instead, for externalising the effect of income is visible for the second-to bottom and lowest quintile.
- The lower the education level, the more total difficulties experienced by child – this is a strong effect especially for the two lowest educational levels. For internalising problems the effect of education is visible only for those with no education and it becomes less statistically significant at age 11 (though the size of the coefficient is similar).
- Not having a working mother increases total difficulties. However, this effect is only significant at age 11. The effect is significant only for internalising problems among subcomponents of total difficulties, with similar sizes and significance levels across age.
- Family structure is significant at age 3 only for externalising and hyperactivity problems, but not for internalising problems.
- Mother being age 30-39 at the birth decreases significantly total difficulties outcomes, especially at ages 7 and 11 (the size of the coefficient is larger at these ages). The effect is stronger for externalising than internalising outcomes, and the effect is also strong for hyperactivity problems.
    - Girls have fewer total difficulties. This effect is driven by externalising and hyperactivity outcomes.
    - The effect of having more siblings is positive; it decreases total difficulties outcomes significantly and it seems that the effect is linear.
    - Looking at the components, the number of siblings seems to decrease internalising and hyperactivity problems more.
    - Total difficulties are significantly higher for Indian children at age 5 compared to white children. The effect is driven mostly by internalising outcomes.

- At age 7, the effect is not significant for any outcome. At age 11 the effect changes sign for all outcomes, with Pakistanis or Bangladeshi children experiencing fewer difficulties (overall and for each component) than white children.

    - If the parent has a weaker agreement on the importance of stimulation for the child, this increases behavioural problems at ages 5 and 7. The effect is driven both by internalising and externalising problems. The effect sizes are larger at age 11.

    - Living in a neighbourhood considered unsafe increases total difficulties and each component across child ages.

### Interpretation of Findings and Effect Sizes

In the following Figures 7 and 8 we report predicted probabilities estimated from multivariate models to illustrate the effect size of a given variable on the final outcome score while holding constant all other variables at their mean values. We chose two variables for illustration: mother's mental health and mother's education. As can be seen from Figure 7, mother's mental health has the largest effect on total difficulties when all other variables are taken into account. Mother's education is associated with variation in average difficulties of 2 points, holding constant all other predictors at their means (Figure 8). These graphs illustrate the range of total difficulties

Graphs for all other predictors and outcome measures are included in the Appendix. See Appendix also for the effects of the significant predictors on the estimated levels of SDQ subscales.

Figure 7: Predicted values of total difficulties by maternal mental health (Kessler)

Predicted values of SDQ total difficulties score for values of mother's mental health
CI 95%
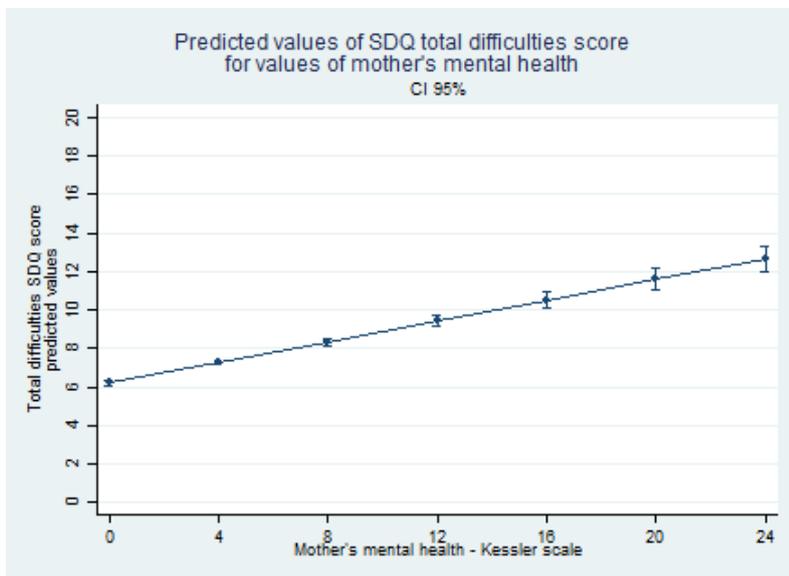
Figure 8: Predicted values of total difficulties by maternal education
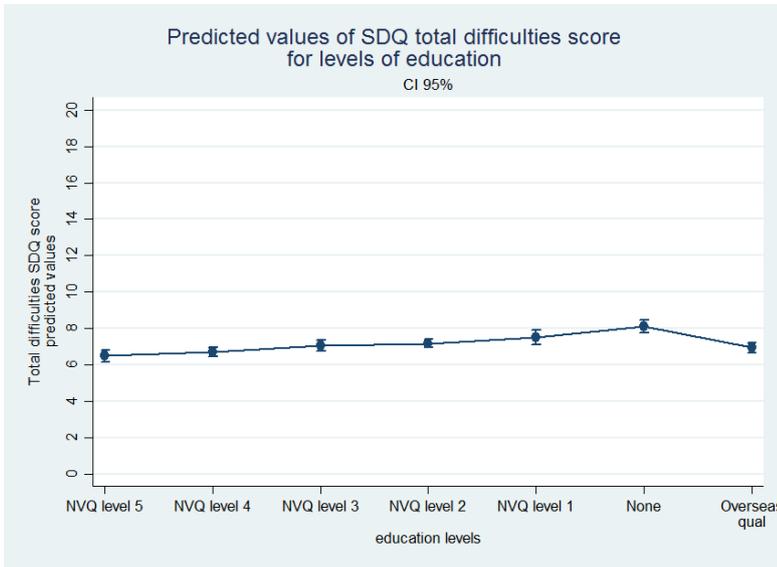
4.3. Child Fixed-Effects Models

Below we provide a summary of the results of child fixed-effects estimations. Many of time varying predictors do not change much over short period observed for the child, rendering them non-significant. Non-significant results might also indicate that some of these variables' true effect is not statistically significant.

These models are estimated only for a limited set of outcomes: Fixed effects could not be estimated for bullying outcomes, nor for most of the cognitive outcomes, including KS2. The most complete set of estimations is supplied for Behavioural Outcomes, which are reported below.

Health Outcomes

BMI, Obesity and Overweight: We find that an additional maternal year of age increases BMI by 0.56 and overweight by 1.3% among boys and 0.75 and 1.7% among girls. However, highest maternal education level (NVQ 4) reduced BMI by 0.7 units and overweight by 5.6% among boys and only 0.56 units of BMI for girls and did not reduce their probability of overweight. The rest of the results are either non-significant or non-robust.

Self-reported health: Very few predictors remain significant when within-individual variation is analysed. Having three or more siblings is associated with worse child health and having more portions of fruit per day is associated with better child health.

Cognitive Outcomes

For most cognitive skill outcome measures, child fixed-effects are not possible to estimate because these measures do not repeat in the same way across ages. The only outcome measure comparable over time is "pattern construction". We find that "pattern construction" is only available for ages 5 and 7, so the panel has two observation points. For this variable, we have been able to identify the following time varying variables, which have significant effects on pattern constriction: higher maternal age leads to higher scores by 0.74 points, and having a working mothers increase scores by 1.3 points.

Behavioural Outcomes

Changes in work status of the mother from the baseline category of working to not working have a negative effect on total difficulties and internalising behaviour. This is clearly a strong result as this effect is net of all fixed factors at the family and child level.

Increases in number of siblings consistently affect all behavioural outcomes. Additional siblings increase on average the scores on total difficulties and internalising behaviour and decrease externalising behaviour and hyperactivity and attention deficit problems, other things being equal.

LSE Enterprise

Table 17: Summary of result from Child Fixed effects specification for Behavioural Outcomes

	Total difficulties	Internalising behaviour	Externalising behaviour	Hyperactivity/ Inattention	Prosocial behaviour
Mother's mental health					
Income [base: highest]					
Fourth quintile					
Third quintile					
Second quintile					
Lowest quintile					
Work status [base: work]					
Education [base: level 5]					
NVQ level 4	+	+			+
NVQ level 3					
NVQ level 2					
NVQ level 1					
None of these		- -			-
Overseas education	- -	- -			-
Fam. Structure [base: 2 par.]					
Num. siblings [base: none]					
One	++	++		- -	
Two	++	++	-	- -	
Three or more		++	-	- -	
N	17,267	17,267	17,267	17,267	17,267

Note: Table above provides a summary of time varying variables that are significant in the fixed effects models [balanced samples] for each outcome variable. Number of plus or minus signs indicate the size of coefficients in absolute terms according to the following criteria:

> *Coefficients with an absolute size 0<b<0.4 is shown as + or -*
> *Coefficients with an absolute size 0.3<b<1 is shown as ++ or - -*
> *Coefficients with an absolute size 1<b is shown as +++ or - - -*

The predictors that are not significant in the multivariate models are represented with an empty cell. Multivariate regression output tables for all models with the estimated coefficients (marginal effects) are provided in the Appendix.

4.4. Robustness Checks

2.4.1 Attrition

In addition to use appropriate weights to ensure representativeness and to correct for attrition, all multivariate models in the appendix are reported using both the balanced and unbalanced samples in two consecutive columns for all outcomes. This allows us to see change in coefficients due to attrition. This is clearer in linear models. In non-linear models differences in samples may generate comparability problems.

Health Outcomes - Self-reported health and Obesity

Attrition is not a major problem for most of the variables considered: namely, the balanced and unbalanced samples report the same coefficients. The difference in the number of observations between balanced and unbalanced panels is 1,110 less at age 5 compared to the initial sample of MCS children; but 521 less at age 7 than age 5, and only 1 observation less at age 11 compared to age 7. Hence, samples are the same between ages 7 and 11, are similar to age 5, meaning that attrition not a major problem for comparison over these ages.

Smoking during pregnancy is one of the few variables that have a significant effect on girls' BMI only on a balanced panel sample but not in the unbalanced sample. The differences between balanced and unbalanced samples were mostly on the effect sizes, and the effect on sample size was limited to 700 observations out of 2,800.

Cognitive Outcomes

For cognitive outcomes, the sample sizes are smaller for balanced, we lose 2,000 observations (out of approximately 7,000) for age 5 but the attrition drops to only 1,200 for age 11. However, it does not affect the statistical significance or signs of the coefficients, just as in the health case, it merely affects the size of coefficients (mostly it increases them).

Behavioural Outcomes - SDQ

In general there seems to be no major differences between the balanced and the unbalanced panels, and therefore, problems of attrition seem to be small.

Two exceptions are as follows:

- Internalising behaviour - Black or British: at age 5, the coefficient for black individuals is lower and ceases to be significant when we estimate our models on the balanced sample.
- Prosocial behaviour – number of siblings: at age 5, the coefficient for one and three or more siblings is smaller and not significant when we take only the balanced sample into account.

Social outcomes

For bullying outcomes, two observations at age 7 and 11 are used. Attrition is negligible between these two ages therefore coefficients were identical for all variables between balanced and unbalanced samples.

In sum, the results suggest that attrition is not a major concern for our analysis and that the results presented in the main text are by and large robust to this check.

4.4.2 Bonferroni Corrections

Our multivariate specifications include a large set of predictors for each outcome and we cover many child outcome measures. When a large number of variables are included in a model specification to predict one outcome measure, it is possible to observe at least one significant association due to random chance. There are various methods to balance the effects of all factors working on the outcomes at the same time. One could start with focusing on the outcomes first and apply the Bonferroni method, which is a simple method for correcting for multiple comparisons, especially when several dependent or independent statistical tests are being performed simultaneously on a single data set (using the same sample). In other words, the probability of identifying at least one significant result due to random chance increases as more hypotheses are tested using the same dataset. Bonferroni correction sets the significance cut-off at a more conservative threshold than, say typical 95%. The threshold for significance cut off becomes more conservative as the number of predictors increase as it is adjusted by the number of predictors included the specification. This is especially the case for our multivariate analyses. Bonferroni adjustment can be used to correct any set of P values for multiple comparisons, after ANOVA or OLS regression for this type of type I error. The advantages of this method are that it is simple to understand and could be applied to various types of multivariate analyses. The disadvantage is that it tends to be too conservative, when the number of predictors is high (i.e. more than 20).

Below, we report the predictors identified in the multivariate that <u>did</u> survive the Bonferroni corrections, and are therefore the factors that they have statistically significant associations not produced by random chance. The implication here is that the variables in multivariate analyses that are not reported in tables below should be treated with caution.

Table 18: Predictors that Survived Bonferroni Correction: Health Outcomes (Boys and Girls)

	BMI			Overweight		
MATERNAL PREDICTORS	Age5	Age7	Age 11	Age 5	Age 7	Age 11
Birth Weight				✓		
Smoked during pregnancy					✓	✓
Working mother					✓	
Maternal BMI at sweep 1		✓	✓	✓	✓	✓
SOCIO-ECONOMIC FACTORS						
Income Quintile is below the highest sweep 1		✓	✓		✓	
Income Quintile is below the highest sweep 2				✓		
Breastfeeding up to 6 months (Base = no breastfeeding)					✓	
Mothers age at birth	✓			✓		
Regular bed times			✓	✓		
Exercise with child (Frequency)		✓			✓	

Many variables that were significant in bi-variate associations for the BMI outcome did not survive Bonferroni corrections, especially at age 5. More predictors survived for the overweight outcome.

LSE Enterprise

Maternal BMI remains one of the consistent predictor of child BMI and probability of overweight. Children start their life at the highest income quintiles also advantaged at later ages regarding the risk of overweight and increases in BMI. Mother's age at birth only remained significant for both outcomes at age 5 but not later. Mother smoking during pregnancy has an association with child being overweight at age 7, 11, which survived Bonferroni correction.

Table 19: Predictors that Survived Bonferroni Correction: Self-Reported Health

	Self-reported health		
MATERNAL PREDICTORS	**Age 5**	**Age 7**	**Age 11**
Maternal Smoking	✓		
Drinking (Every day)	✓		
Felt low or sad	✓	✓	
Breastfeeding			
Mothers' health	✓		
Mother's mental health		✓	
SOCIO-DEMOGRAPHIC FACTORS			
Income Quintile is below the highest sweep 1			✓
Mother's Education Level is Lowest			✓
Mothers age at birth		✓	
Female	✓	✓	✓
Ethnicity – Indian		✓	✓
Ethnicity – Pakistani and Bangladeshi	✓	✓	✓
Headaches	✓	✓	✓
Child picked on/Bullied	✓	✓	✓
Exercise with child (Frequency)	✓	✓	

Bonferroni corrections on the multivariate analyses showed that the ethnicity variables (e.g., being Pakistani or Indian) are consistently associated with mother's self-reported health outcomes of children at all ages. Most maternal predictors only survived for Age 5 outcomes.

LSE Enterprise

Table 20: Predictors that Survived Bonferroni Correction: Cognitive Outcomes

MATERNAL PREDICTORS	Age 5			Age 7			Age 11
	Naming Vocab.	Pattern Construc.	Picture Similar.	NFER MATH Skills	Pattern Construc.	Word Reading	Verbal Similarity
Married/Cohabiting at birth							✓
Birth Weight							✓
Bleeding and thread to miscarriage		✓					
SOCIO-DEMOGRAPHIC FACTORS							
Breastfeeding up to 6 months			✓	✓	✓	✓	✓
Income at age 9m				✓	✓	✓	✓
Income at age 3y				✓	✓	✓	✓
Work status during pregnancy						✓	✓
Work at age 9m	✓			✓	✓	✓	✓
Work at birth	✓						
Maternal Education	✓	✓	✓				
Mothers age at birth				✓	✓		✓
Female	✓	✓	✓		✓	✓	
Child care							✓
Reception				✓	✓		

Predictors such as, Income, Mother's work status at age 9 month, Breastfeeding up to 6 months have associations that survived Bonferroni corrections at age 7 and age 11. Being a Girl is associated with better cognitive outcomes at all ages, even after Bonferroni correction, except for Maths skills at age 7.

Table 21: Predictors that Survived Bonferroni Correction: KS2 Test Scores

VARIABLES	Reading	English	Maths	Arithmetic
Female	✓	✓	✓	✓
Months of birth	✓	✓	✓	✓
Working At Sweep 1	✓	✓		
Income Quantile at Sweep 1	✓	✓	✓	✓
Income quantile at Sweep 4	✓	✓	✓	✓
Maternal Education	✓	✓	✓	✓
Mothers Age	✓	✓	✓	

After testing form the Bonferroni correction we find that gender of child, month of birth, maternal work, income at sweep 1 and 4, maternal education and age survive the correction.

Behavioural Outcomes

We report below the variables that survived Bonferroni Adjustment by each age and behavioural outcome.

Table 22: Predictors that Survived Bonferroni Correction: Total Difficulties

	Age 5 Bonferroni-Adjusted	Age 7 Bonferroni-Adjusted	Age 11 Bonferroni-Adjusted
Smoking	✓	✓	✓
Drinking	✓		
Felt low or sad	✓	✓	✓
Illness	✓	✓	✓
Mother's mental health	✓	✓	✓
Income	✓	✓	✓
Education	✓	✓	✓
Mother's age at birth	✓	✓	✓
Gender	✓	✓	✓
Number of siblings	✓		✓
Ethnicity	✓		✓
Importance stimulation child	✓	✓	
Neighbourhood	✓	✓	✓

Table 23: Predictors that Survived Bonferroni Correction: Internalising

	Age 5 Bonferroni-Adjusted	Age 7 Bonferroni-Adjusted	Age 11 Bonferroni-Adjusted
Felt low or sad	✓	✓	✓
Illness	✓	✓	✓
Mother's mental health	✓	✓	✓
Income	✓	✓	✓
Work status	✓	✓	✓
Mother's age at birth		✓	✓
Number of siblings	✓	✓	✓
Ethnicity	✓		✓
Importance stimulation child			✓
Neighbourhood	✓	✓	✓

LSE Enterprise

Table 24: Predictors that Survived Bonferroni Correction: Externalising Problems

	Age 5 Bonferroni-Adjusted	Age 7 Bonferroni-Adjusted	Age 11 Bonferroni-Adjusted
Smoking	✓	✓	✓
Felt low or sad	✓	✓	✓
Mother's mental health	✓	✓	✓
Income (base: high)	✓	✓	✓
Education (base: high)	✓	✓	✓
Mother's age at birth	✓	✓	✓
Gender	✓	✓	✓
Ethnicity			✓
Importance stimulation child *Agree*	✓ ✓	✓ ✓	
Neighbourhood		✓	✓

Table 25: Predictors that Survived Bonferroni Correction: Hyperactivity/ AD Problems

	Age 5 Bonferroni-Adjusted	Age 7 Bonferroni-Adjusted	Age 11 Bonferroni-Adjusted
Smoking	✓	✓	✓
LBW	✓	✓	
Felt low or sad	✓	✓	
Mother's mental health	✓	✓	✓
Income	✓		
Education	✓	✓	
Mother's age at birth	✓	✓	✓
Gender	✓	✓	✓
Number of siblings	✓	✓	✓
Ethnicity	✓		✓
Importance stimulation child	✓	✓	
Neighbourhood		✓	✓

Maternal predictors, such as Mother's smoking and Drinking behaviour, mental health, feeling low or sad have consistently survived Bonferroni corrections and turned out to be associated with many of the Behavioural problem outcomes at all three ages. Gender differences in internalising behaviour did not survive Bonferroni correction, however, gender still is an important predictor for behavioural problems at all ages for three of the four outcome measures.

Social Outcomes: Being Bullied and Peer Problems

Table 26: Predictors that Survived Bonferroni Correction: Being Bullied

	Age 7	Age 11
Month of birth [base: Sep]	✓	✓
Female	✓	
Number of Siblings	✓	
Income		✓
Mother's education		✓
Mother's mental health	✓	
Child ethnicity [base: White]		✓
BMI	✓	
Obese		✓
Child had longstanding illness [base; no]	✓	✓
Internalising behaviour - emotions & peer	✓	✓
Peer Problems	✓	✓

	Age 5 Bonferroni-Adjusted	Age 7 Bonferroni-Adjusted	Age 11 Bonferroni-Adjusted
Smoking	✓		
Premature			
Felt low or sad			
Illness		✓	✓
Mother's mental health	✓	✓	✓
Income	✓	✓	✓
Work status			✓
Education level			
Mother's age at birth			✓
Gender child	✓	✓	✓
Number of siblings	✓	✓	✓
Ethnicity	✓		
Importance stimulation child			
Neighbourhood	✓	✓	✓

Conclusions

In this report we set out to shed light on those factors that are associated with child outcomes across multiple domains of their lives and across their early years up to the end of primary school. We considered the following domains of children's lives: health, cognitive and educational, behavioural, and social, and we looked at a variety of indicators covering these domains for ages 5, 7 and 11. We had already ascertained what factors were identified from the literature as potentially relevant to the different child outcomes. We then tested whether we found expected (or counterintuitive) associations using careful empirical analysis of the Millennium Cohort Study.

We identified, which factors (or predictors) were found to be significantly associated with the different outcomes, at which ages, and we compared the strength of the associations. We proceeded in stages. First, we identified the simple associations between an outcome at one of the three ages and a 'predictor' measured in an earlier survey. We then considered the various predictors together in order to identify which ones had 'independent' effects on the selected outcomes. We then not only compared the strength of these associations, but we also evaluated the robustness of our results using a series of additional tests. We compared our analytical samples with those that only retained all children observed at all sweeps to get some insight into whether attrition was affecting our results; we used child fixed effects models to more robustly identify 'causal' relationships; and, given the number of our estimations and the number of our predictors, we used Bonferroni tests to assess the extent to which some of our observed relationships might be due to chance.

Overall we found a small number of highly robust relationships, many of which persisted across all three child ages. These were as follows:

For **BMI / overweight** we found that maternal weight before pregnancy and maternal BMI at age nine months were consistent predictors of higher BMI and of an increased risk of overweight. While the effect sizes were not large once other factors were controlled for they were robust and persisted up to age 11, though maternal weight before pregnancy was a strong predictor only for girls. We also found that smoking during pregnancy was associated with higher risks of overweight at ages 7 and 11, revealing the potential long-run impacts of pre-natal and early childhood behaviours. While a range of other predictors were associated in some analyses and at some sweeps none of the other findings were as robust. That is they did not 'survive' all our additional tests.

For **Self-Reported Child Health** we found consistent negative effects on health of low birth weight and being of Pakistani or Bangladeshi ethnic origin, while higher income was associated with better health at age 11. In addition, being picked on by other children or being bullied was consistently associated with worse health. Again a number of other variables were significant in some models and at certain ages, and so may have some relevance, but were not fully robust to all our tests.

Turning to **Cognitive and educational outcomes,** we identified a female advantage in the age 5 tests, for Pattern Construction and word reading at age 7, but by age 11 this advantage was no longer present, though in the KS2 outcomes they performed better on English and Reading and worse on Maths and Arithmetic. These findings were robust to our additional tests. In addition, having a mother who worked at age nine months had a consistently negative independent association with cognitive outcomes in the longer term (age 7 and 11), though interestingly not at age 5. Similarly

income had a consistent effect on cognitive outcomes at ages 7 and 11. Lower maternal education mattered for all cognitive outcomes at all ages in multivariate and after the Bonferroni correction the association at age 5 strengthened. These factors also matter for age 11 KS2 outcomes for English and Reading but, interestingly, not for Maths and Arithmetic. Breastfeeding up to 6 months was also robustly associated with all test results at ages 7 and 11.

Finally, in terms of **Behavioural outcomes**, we found that mother's mental health (measured as the Kessler score or as feeling low or sad) was associated with all child's behavioural difficulties, both total difficulties and all sub-domains. Although the size of the effects is not large, the associations are highly consistent, even if the child fixed effects models were not able to identify significant associations given the lack of variation in responses over time. Maternal education was also significantly associated with total difficulties at all ages, while having more siblings was associated with fewer problems. Again this was a highly consistent finding, though in the fixed effects analysis, the picture was more complicated with additional children over time increasing total difficulties and internalising behaviour scores and decreasing externalising behaviour scores. The reduction in difficulties associated with more siblings was also supported by the Bonferroni analysis. Overall our results suggest that higher number of siblings consistently reduce at least two behavioural outcomes, hyperactivity and internalising problems.

From this analysis we can see how many factors that have a bearing on child outcomes are in fact interlinked. In addition, the findings demonstrate how important behaviours and characteristics from early childhood - or even from before birth - can be for outcomes 6-10 years down the line. They also reinforce some of the associations that were clear from the earlier PREview study, such as the importance of maternal mental health for behavioural outcomes and of maternal education and family income for a range of outcomes. Finally, it is important to recognise that in many cases, even if the associations are clear-cut, they only go a small way to explaining the great diversity in outcomes across children. Children differ in multiple ways for reasons that we cannot capture, and even the strongest relationships are not deterministic.