The socio-economic gradient in early child outcomes: evidence from the Millennium Cohort Study

Lorraine Dearden
Institute for Fiscal Studies and Institute of Education, University of London
ldearden@ifs.org.uk
Luke Sibieta
Institute for Fiscal Studies
Kathy Sylva
University of Oxford

Abstract

This paper shows that there are large differences in cognitive development between children from rich and poor backgrounds at the age of 3, and that this gap widens by the age of 5. Children from poor backgrounds also face much less advantageous "early childhood caring environments" than children from better off families. For example we identify differences in poor children's and their mothers' health and well-being (e.g. birth-weight, breast-feeding, and maternal depression); family interactions (e.g. mother-child closeness); the home learning environment (e.g. reading regularly to the child); parenting styles and rules (e.g. regular bed-times and meal-times), and experiences of childcare by ages 3 and 5. Differences in the home learning environment, particularly at the age of 3 have an important role to play in explaining why children from poorer backgrounds have lower test scores than children from better off families. However, a much larger proportion of the gap remains unexplained, or appears directly related to other aspects of family background (such as mothers' age, and family size) that are not mediated through the early childhood caring environment.

1. Introduction

Children growing up in poor families emerge from our schools with substantially lower levels of educational attainment. These educational deficits emerge early in children's lives, even before entry into school, and widen throughout childhood (Feinstein 2003, 2004). In this article, we examine some of the routes through which family socio-economic position (SEP) affects cognitive development in the early years up to age 5.

Other articles in this Special Issue examine the socio-economic gap in the primary years (Gregg and Washbrook), the secondary years (Chowdry, Crawford and Goodman) and inter-generational factors (Crawford, Goodman and Joyce). The overall theoretical and empirical approach taken in these

four articles is discussed in detail in Goodman, Gregg and Washbrook, also in this Special Issue.

In seeking to explain the socio-economic gap in early child cognitive outcomes, we focus on a range of parenting behaviours, health and well-being factors, and family interactions. We refer to these factors collectively as the "early childhood caring environment". These have seldom been measured together in large datasets, preventing comprehensive empirical analysis of their roles as predictors and consequences of poor child outcomes. In order to conduct this analysis we will follow children over time within the Millennium Cohort Study (MCS), a recent and detailed study of children born at the turn of the millennium. Here

we focus on cognitive outcomes, but in our more detailed working paper (Dearden, Sibieta and Sylva, 2010) we also examine socio-emotional development, the mediation of other gradients and the determination of the early childhood caring environment.

The rest of this paper proceeds as follows: section 2 describes our theoretical approach, and section 3 describes the data and summary statistics, including the socio-economic gradients in early child outcomes and the early childhood caring environment. In section 4, we use multivariate regression techniques to decompose the socio-economic gaps in child outcomes as per the empirical methodology set out in Goodman, Gregg and Washbrook in this Special Issue. Finally, section 5 concludes the paper.

2. Theoretical Approach

In this paper, we use the same theoretical approach described in detail in Goodman, Gregg and Washbrook in this volume. Our starting point is the socio-economic gradient in child outcomes in the early years. We then explore some of the potential transmission mechanisms through which socio-economic disadvantage may lead to poor child outcomes. In doing so, we attempt to explore the role played by a set of 'proximal' factors that may drive the relationship between poor child outcomes and 'distal' influences such as socio-economic disadvantage and other aspects of family background. Generally speaking, proximal factors are those that drive the observed association between a distal factor and an outcome (as can be seen in Figure 1).
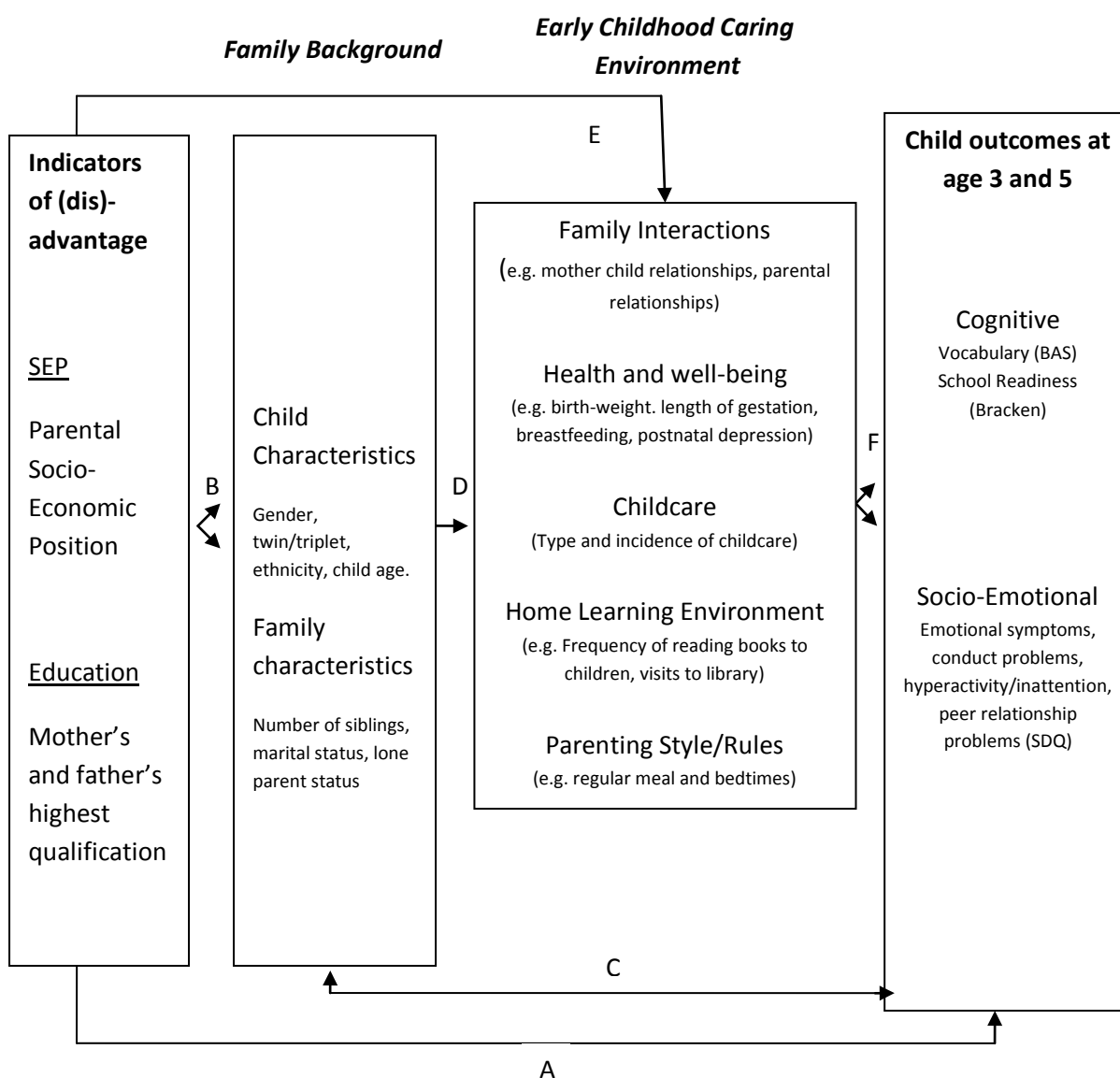
In this paper, we refer to the full set of proximal influences we observe as 'the early childhood caring environment'. This includes the nature of family interactions, health and well-being factors, childcare arrangements, the home learning environment (HLE), and parenting style and rules. These factors are inter-related, and seem to be the most likely to influence child development between birth and the

age of five. These factors will mediate the SEP gradient both via the effects of SEP on the early childhood caring environment (arrow E), and the early childhood caring environment on child outcomes (arrow F). However, these factors might also mediate the effects of other distal factors such as family background (arrow C) via a combination of arrows D and F.

Our choice of transmission factors reflects the diverse social science literature on the determinants of educational success, as a well as data availability. Empirical research in this area is far from definitive, but a number of potential pathways are frequently discussed in the literature (e.g. the home environment, quality of child care, parent-child relationships etc), many of which have been included in the present theoretical framework.

Several studies have found that differences in the home environment, as measured by the HOME scale (which includes items on household resources, such as reading materials and toys, and parental practices, such as discipline methods), account for a substantial portion of the effect of income on the cognitive development of pre-school children and on the achievement scores of elementary school children (Brooks-Gunn et al 1993; Duncan, Brooks-Gunn and Klebanov 1994; Korenman, Miller and Sjaastad 1995). Similar findings have been obtained by the Effective Provision of Pre-school Education (EPPE) Project in the United Kingdom, which found that although other family factors are also highly significant, the early years HLE (i.e. frequency of self-reported parental involvement in a range of activities, such as reading, library visits, playing with letters/numbers etc.) exerts a significant and independent influence on several cognitive (i.e. attainment at ages 3, 5 and 11) and behavioural (i.e. self-regulation, pro-social behaviour, and hyperactivity) outcomes. However, the EPPE HLE index is only moderately correlated (r=0.33) with family socio-economic status (SES) or mother's qualification levels (Sylva et al 2008; Melhuish et al 2008).

Figure 1. Simple model linking financial position to early education outcomes

Family Background   *Early Childhood Caring Environment*

Indicators of (dis-) advantage	Child Characteristics	Family Interactions / Health and well-being / Childcare / Home Learning Environment / Parenting Style/Rules	Child outcomes at age 3 and 5

Indicators of (dis-)advantage

SEP

Parental Socio-Economic Position

Education

Mother's and father's highest qualification

Child Characteristics

Gender, twin/triplet, ethnicity, child age.

Family characteristics

Number of siblings, marital status, lone parent status

E

Family Interactions

(e.g. mother child relationships, parental relationships)

Health and well-being

(e.g. birth-weight. length of gestation, breastfeeding, postnatal depression)

Childcare

(Type and incidence of childcare)

Home Learning Environment

(e.g. Frequency of reading books to children, visits to library)

Parenting Style/Rules

(e.g. regular meal and bedtimes)

Child outcomes at age 3 and 5

Cognitive
Vocabulary (BAS)
School Readiness
(Bracken)

Socio-Emotional
Emotional symptoms, conduct problems, hyperactivity/inattention, peer relationship problems (SDQ)

B   D   F   C   A

Another important pathway involves the health and well-being of the child and parents (e.g. birth-weight, gestation, breast-feeding patterns and indicators of post-natal depression). A 1990 analysis in the USA indicated that the poverty-related heath factors such as low birth weight, elevated blood lead levels, anaemia, and recurrent ear infections and hearing loss contributed to the differences in IQ scores between poor and non-poor four-year-olds (Goldstein 1990). The findings suggest that the cumulative health disadvantage experienced by poor children on these health measures may have accounted for as much as 13% to 20% of the difference in IQ between the poor and non-poor children during the 1970s and 1980s (Goldstein 1990). Parents who are poor are also likely to be less healthy themselves, both emotionally and

physically (Adler, Boyce, Chesney, Folkman and Syme 1993), and this could result in impaired parent-child interactions and fewer home learning experiences (Bornstein 1995). For example, a study conducted by the USA National Longitudinal Survey of Youth (NLSY) found that currently poor mothers spanked their children more often than non-poor mothers, and this harsh behaviour was an important component of the effect of poverty on children's mental health (McLeod and Shanahan 1993).

A third possible pathway is through the care young children receive outside the home, as much research has shown that high-quality, developmentally appropriate child care in the pre-school years is associated with enhanced social, emotional, and linguistic competence (Howes 1988; Hofferth and

Phillips 1991; NICHD Early Child Care Research Network 1997, 1998; Ramey and Ramey 1998). In addition, randomized trials have demonstrated that intensive early childhood programmes for poor children can increase verbal ability and reasoning skills through early elementary school (Belfield, Nores, Barnett and Schweinhart 2006; Brooks-Gunn et al 1994; Burchinal, Campbell, Bryant, Wasik and Ramey 1997; Garces, Thomas and Currie (2002); Lazar and Darlington 1982; Ludwig and Miller 2007; Ramey and Ramey 1998).

One crucial mechanism we are unable to consider is the inter-generational transmission of cognitive ability and social skills. To the extent that parental ability and social skills are also correlated with the measures of the early childhood environment we do observe, their omission may lead to biased conclusions about the extent to which the observed factors can explain the socio-economic gaps in educational attainment. Crawford, Goodman and Joyce in this Special Issue use the British Cohort Study to examine the influence of parental cognitive ability and social skills on children's cognitive skills. They conclude that parental cognitive ability (measured during the parent's childhood) is a statistically significant predictor of children's cognitive skills, and that it explains about one sixth of the SEP gap in those skills, even after controlling for a rich set of demographic, attitudinal and behavioural factors. However, the addition of these parental characteristics to their model does very little to alter their impression of the relative importance of other proximal factors.

3. Data and descriptive statistics

The Millennium Cohort Study (MCS) began as a longitudinal study of approximately 18,000 children born in the UK in 2000. The first sweep of the study was conducted when MCS children were about 9 months old. This over-sampled individuals from ethnic minorities and individuals living in disadvantaged areas of the country. Three further sweeps of data have since been collected when the children were aged about 36 months (sweep 2), when they were about 5 years old (sweep 3) and when they were 7 years old (sweep 4). Further sweeps of data will be collected at future key milestone ages. For our analysis, we chose to sample those who responded to the first three surveys and those where the mother is the main respondent[i]. We also excluded individuals who had missing data for some key characteristics, such as education, measures of the HLE, child outcomes and ethnicity. This left us with approximately 11,100 observations.

We proxy for each child's socio-economic background using two measures: a socio-economic position index and parent(s) highest educational qualifications. As described in detail in Gregg et al (2010), we construct our index of socio-economic position by performing principal-component analysis on a number of potential proxies for socio-economic background (average log equivalised income at sweeps 2 and 3, mother's class, father's class, housing tenure and whether the family have experienced financial difficulties). We then take the first principal component and use this as an index of socio-economic position. The factors used are summarised in Table 1 below, across quintiles of this index of socio-economic position. The data is weighted to take account of both the sampling design and non-response, as it is throughout this paper.

Table 2. Factors included in the index of socio-economic position

	SEP	SEP	SEP	SEP	SEP	Gap
Avg. Log Equivalised Income (MCS2 and MCS3)	4.93	5.26	5.55	5.87	6.39	1.46**
Mother – Professional/Managerial (MCS1)	1.4%	7.4%	22.7%	50.6%	84.1%	82.6 ppts**
Mother – Intermediate Occupation (MCS1)	4.6%	14.9%	30.9%	32.7%	14.6%	10 ppts**
Mother – Small Employer (MCS1)	1.2%	6.0%	7.5%	6.6%	1.1%	-0.1 ppts
Mother – Lower Supervisory (MCS1)	7.0%	8.7%	8.0%	4.1%	0.1%	-6.8 ppts**
Mother – Routine Occupation (MCS1)	85.8%	63.1%	31.0%	6.1%	0.1%	-85.6 ppts
Father – Professional/Managerial (MCS1)	2.6%	12.8%	28.7%	55.6%	94.7%	92.1 ppts**
Father – Intermediate Occupation (MCS1)	1.4%	3.9%	8.2%	9.0%	2.8%	1.4 ppts**
Father – Small Employer (MCS1)	7.1%	17.4%	19.9%	15.3%	2.2%	-4.9 ppts**
Father – Lower Supervisory (MCS1)	18.1%	25.4%	21.8%	14.0%	0.2%	-17.9 ppts**
Father – Routine Occupation (MCS1)	70.8%	40.5%	21.5%	6.1%	0.1%	-70.7 ppts**
Rented Privately in one or more sweep	14.0%	12.2%	7.5%	2.9%	0.5%	-13.4 ppts**
In Social Housing in one or more sweep	86.0%	31.4%	4.5%	0.5%	0.0%	-86 ppts**
Experienced financial difficulties	41.9%	19.2%	12.3%	5.4%	0.1%	-41.8 ppts**

*** and * indicate significant differences between Q1 and Q5 at the 1% and 5% levels, respectively.*
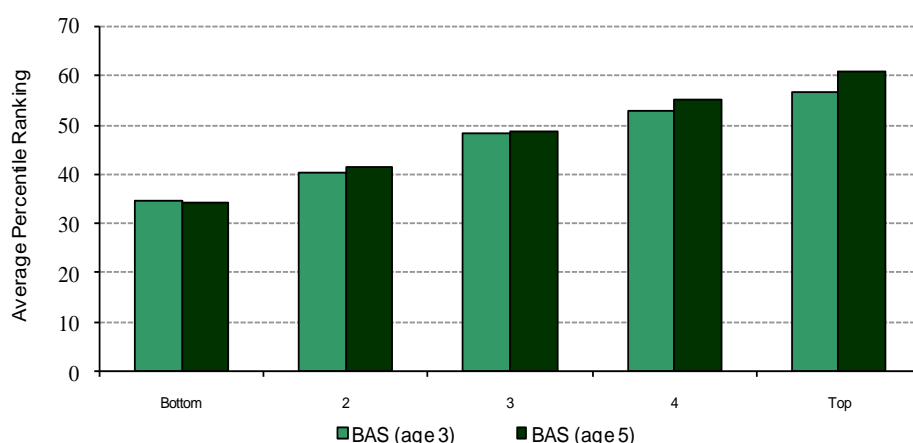
We focus on socio-economic differences in a specific measure of cognitive development: the British Ability Scales (BAS) (Early Years version) Naming Vocabulary. In our more detailed working paper (Dearden et al 2010), we also examine differences in the Bracken School Readiness measure, as well as differences in socio-emotional development as captured by the Goodman Strength and Difficulties questionnaire.

The BAS is well recognised as an excellent measure of children's vocabulary, and is highly correlated with other language measures as well. It is also one of the best predictors of children's general intelligence, which is highly stable, as there may well be a strong genetic component to both cognitive and linguistic skills. On the other hand, a measure of school readiness, such as the Bracken School Readiness measure, is much more amenable to positive change through early childhood education programmes. For ease of interpretation, we have converted all outcome measures into percentile ranks.

Figure 2 shows the average percentile rankings of BAS scores at ages 3 and 5 for these five SEP quintiles. At age 3, the gap between the top and bottom quintiles for the BAS stands at about 22 percentile points, but by age 5, the gap has widened to over 26 percentile points.

Figure 2. BAS Scores by SEP Quintile

In Table 3, we investigate the dynamics of the BAS scores in a bit more detail. As we would expect, children in lower SEP quintiles are much more likely than those in higher quintiles to be in the bottom 40% of BAS scores at age 3. However, lower quintiles are also much less likely to escape the bottom 40% by age 5, as only around 25% of these children were able to do so (compared to about 60% in the top quintile group). Again, as expected, children from lower quintiles are less likely to be in the top 40% of BAS scores at age 3, but those who attained that level of vocabulary skills are more likely than children in higher quintiles to drop out later[ii]. So not only do children from lower SEP quintiles have lower average outcomes at age 3, but they are also more likely to stay in the bottom 40% of achievers and more likely to drop out of the top 40%. The aim of the rest of this paper is to try and explain the socio-economic gaps in child outcomes at age 3 and age 5, as well as the widening of the gap between ages 3 and 5.

Table 3. Cognitive ability at ages 3 and 5, by parents' SEP

Proportion of children:	SEP Q1	SEP Q2	SEP Q3	SEP Q4	SEP Q5
Bottom 40% at Age 3 (BAS)	58.5	50.0	37.8	31.5	27.0
Escape from bottom 40% by Age 5 (BAS)	25.4	33.1	41.8	53.7	60.7
Top 40% at Age 3 (BAS)	24.6	33.4	45.8	52.1	58.3
Drop out of top 40% by Age 5 (BAS)	57.9	48.3	40.8	34.1	29.8
Bottom 40% at Age 5 (BAS)	61.9	51.8	39.8	30.3	22.3
Top 40% at Age 5 (BAS)	21.7	30.6	42.2	51.8	59.5

Transmission Mechanisms

Here we describe the groups of proximal influences that we have chosen to examine as potential transmission mechanisms.

Parental Education consists of variables that measure parent(s) highest educational qualification. We measure this as the highest equivalent NVQ level for both mothers and, where present, fathers. Dummy variables are included for other qualifications that could not be categorised into an NVQ level.

Family Background consists of variables pertaining to characteristics of the child and the family. Child characteristics include: the sex of the child; their age; whether they are a twin or triple; whether the child was in special care unit just after birth; and their ethnicity. Characteristics of the rest of the family include: mother's age at birth (plus a quadratic term); a separate teenage mother dummy variable; parental employment at sweeps 1, 2 and 3; whether only English is spoken in the household at sweep 1; marital status and partnering of the parents at sweeps 1, 2 and 3; the number of siblings and older siblings at sweeps 2 and 3.

The next groups of variables are those we collectively refer to as measure of the early childhood caring environment.

Family Interactions are measured by indicators of relationship quality at sweeps 1 and 2, based on the number of questions where respondents indicate a problem in their relationship with their partner (using a shortened version of the Golombok Rust Inventory of Marital State with 7 items at sweep 1 and 4 items at sweep 2, e.g. Are you on the brink of separation?), which is standardised to have a mean of zero and a standard deviation of one; interviewer-assessed measure of mother-child closeness at sweep 2, defined by number of items where mother and child appeared close out of a total of six items (e.g. whether the mother spontaneously praised their child during the interview), which is then standardised to have a mean of zero and standard deviation of one); mother-child relationship and conflict problems, measured as the number of problems identified on two subscales of the Pianta index (7 item parent-child relationship scale, e.g. do you share an affectionate/warm relationship with your child, and the 7-item parent-child conflicts scale, e.g. Does your child easily become angry with you), which are both standardised to have a mean of zero and standard deviation of one; and, finally, whether mothers and fathers felt that they spent plenty of time with their children at the second sweep.

Health and Well-Being is measured by the number of cigarettes smoked by the mother per day before and after becoming pregnant; units of alcohol mothers reported drinking whilst pregnant and at the time of sweep 1; length of gestation (days); birth-weight (kg); did the mother try to breast-feed; age at which breast-feeding ceased (0-26 weeks); still breast-feeding after 26 weeks; whether the mother was suffering from post-natal depression during the child's early infancy (defined as agreeing to four of the nine statements from a shortened version of the Malaise inventory, e.g. do you often get into a violent rage?); mother reported indicators of infant temperament with three sub-scales relating to children's mood (5-tems, e.g. Is your child your child content during interruptions of milk or solids), the regularity of day-to-day routines (e.g. does your child nap for about the same length from day to day) and adaptability to new situations (5-items, e.g. Is your child still wary or frightened of strangers after 15 minutes); the mother's height (cm), weight (kg) and body mass index before birth; and the father's height (cm), weight (kg) and body mass index at sweep 1.

(Non-Maternal) Childcare indicators are constructed from sweeps 2 and 3 information to show whether parents mainly used informal care settings (grand-parents, other relatives, other non-relatives), and whether children had attended nursery school/class, playgroup, pre-school or childminder by the ages of 3 (sweep 2) and 5 (sweep 3).

Home Learning Environment at age 3 is measured by an index constructed from the reported frequency with which parents undertake 7 specific activities with their children (frequency of reading to child, library visits, play with ABCs/letters, teaches numbers/shapes, songs/nursery rhymes, drawing/painting). These 7 items are added together to give an index with a maximum score of 49, which is then divided into five equally-sized quintiles. We construct a similar index of the home-learning environment at age 5, which is based on a similar set of parenting activities (frequency of reading to child, tells stories to their child, takes them to an outdoor playground, parenting play with toys with their children, parents play sport with their children, songs/nursery rhymes, drawing/painting). These measures of the HLE are comparable to those used in

Melhuish et al (2008). We also separately control for self-reported parenting competency; frequency of reading to child (which is also included in HLE index). Including frequency of reading in addition to the overall HLE index allows us to see whether reading has a differential effect to the others elements. All measures are available at both sweeps 2 and 3.

Parenting styles/rules measures whether: the parents have lots of rules (at sweep 2); parents strictly enforce rules (sweep 2); children have regular bed-times (sweeps 2 and 3); children have regular meal times (sweeps 2 and 3); family eats breakfast together (sweep 3); children watch more than 3 hours of TV a day (sweeps 2 and 3); children play computer for more than 1 hour per day (sweep 3). Table 4 then shows socio-economic differences in parental education, family background and these measures of the early childhood caring environment. As one might expect, mothers and fathers in the bottom quintile are less highly qualified than those from higher quintiles. For example, less than 7% of mothers in the bottom quintile have educational qualifications equivalent to NVQ level 4 or higher (i.e. university or equivalent), as compared with over 70% of the highest quintile.

Children from lower SEP quintiles are generally more likely to be from ethnic minorities (though Indian children are most likely to come from higher SEP quintiles). There are bigger differences across the SEP quintiles when we move on to examine differences in characteristics of the rest of the family. Mothers in the lowest SEP quintile were, on average, younger at the time of birth than those from higher quintiles. Mothers (and fathers, where present) are more likely to have been in employment across the three sweeps of the MCS. Nearly three quarters of mothers from the richest SEP quintile were in work at the time of the third sweep, as compared with about 30% of the poorest quintile. Children from the poorest quintile are less likely to speak only English at home as compared with richest quintile.

Looking at measures of family size and structure, we observe that children in the bottom SEP quintile are more likely to have come from lone-parent families than those from higher quintiles at the first MCS sweep (about 37%, as compared with just under 1% of the top SEP quintile). If two parents are present, parents in the bottom SEP quintile are also more likely to be cohabiting than those from higher quintiles (about a third at the first sweep, as compared with a little over 12% of the richest

quintile). Children from the bottom SEP quintile, on average, also have a greater number of siblings (numbers shown in table are for the third sweep).

The table then moves on to examine differences in our measures of the early childhood caring environment. The first panel relates to family interactions. Lower SEP quintiles seem to experience more mother-child relationship and conflict problems than richer SEP quintiles. Where they are together, there is also a greater degree of relationship problems amongst parents in poorer SEP quintiles than in the richest SEP quintile.

Mothers in the richest SEP quintile are more likely to try breast-feeding than those from poorer quintiles (90% of the richest quintile compared with a little over 50% of the poorest quintile). If they do breast-feed, mothers from the richest SEP quintile are likely to breast-feed for longer as well. However, mothers in the bottom quintile do not, on average, report consuming more units of alcohol during and after pregnancy Mothers in the bottom quintile are, however, more likely to smoke during and prior to pregnancy. Mothers from the poorest quintile also have a shorter length of gestation and their children have lower birth-weights[iii]. Children from poorer families are also less likely to display a happy mood, appear less adaptable to new situations and display less regularity in daily routines. It is also noteworthy than whilst 7% of mothers from the richest quintile suffered from maternal depression during early infancy, about 22% of mothers from the poorest quintile did so. Parents from the richest quintile also report being taller at birth than parents from the poorest quintile, and report being slightly heavier. When we examine differences in Body Mass Index, we find that mothers and fathers in the poorest SEP quintiles are both more likely to be obese and more likely to be under-weight, i.e. are less likely to be classed as normal weight.

Examining childcare patterns up to sweep 2, we find that children from the poorest SEP quintile are more likely to have been to a nursery school or class than those from richer SEP quintile, but less likely to have been to a pre-school, childminder, day nursery or crèche. Children from the middle SEP quintile are those most likely to have been to a playgroup. These patterns continue up to sweep 3, though (as we would expect) noticeably more children have been to a nursery school or class between sweeps 2 and 3. At both sweeps, poorer parents are more likely to be

relying on sources of informal childcare, grandparents in particular.

The next part of the table examines socio-economic differences in the home-learning environment. We find that children from the poorest SEP quintiles are least likely to be in the richest quintile of our HLE index at sweeps 2 and 3. They are less likely to be read to every day at sweeps 2 and 3. Nearly 80% of children in the richest SEP quintile are read to every day at sweep 2, compared with 43% of the poorest SEP quintile.

The last block of the table shows that children from the richest quintile are more likely to have lots of rules and for these to be strictly enforced at sweep 2. They are also the most likely to have regular bed-times or meal times at sweep 2 (about 92% of these children have a regular bed-time compared with around 69% of those from the poorest SEP quintile). These differences continue up to sweep 3, though are less dramatic. Children from the poorest SEP quintile are also more likely to watch more than 3 hours of TV a day at sweeps 2 and 3 than those from higher quintiles, are more likely to play more than one hour of computer a day at sweep 3.

Therefore, there are a wide variety of socio-economic differences across a range of factors, from number of siblings, birth-weight and reading frequency, to regularity of bedtimes, mother's age at birth and childcare patterns. But which of these factors explain the socio-economic gaps in child outcomes at ages 3 and 5 that we saw earlier. This is the focus of the next section.

Table 4 - Socio-economic differences in family and child characteristics/behaviours

	SEP Q1 (Low)	SEP Q2	SEP Q3	SEP Q4	SEP Q5 (High)	Gap (High – Low)
Parental Education (MCS1)						
Mother - no qualifications	32.2%	15.8%	6.8%	2.4%	0.8%	-31.3 ppts**
Mother NVQ level 1	16.8%	12.4%	6.6%	3.3%	0.9%	-15.8 ppts**
Mother NVQ level 2	34.4%	41.3%	37.2%	25.9%	13.5%	-20.8 ppts**
Mother NVQ level 3	10.2%	15.3%	19.4%	17.7%	11.0%	0.8 ppts
Mother NVQ level 4/5	6.5%	15.1%	30.1%	50.7%	73.7%	67.2 ppts**
Mother - other qualifications	3.4%	3.2%	1.2%	1.4%	0.5%	-2.8 ppts**
Father - no qualifications	44.8%	31.9%	18.1%	9.3%	3.9%	-40.9 ppts**
Father NVQ level 1	12.3%	8.5%	6.2%	3.9%	1.1%	-11.1 ppts**
Father NVQ level 2	25.6%	32.0%	31.4%	23.7%	11.4%	-14.1 ppts**
Father NVQ level 3	10.8%	15.1%	18.3%	17.6%	11.4%	0.6 ppts
Father NVQ level 4/5	6.5%	12.5%	26.0%	45.5%	72.2%	65.6 ppts**
Father - other qualifications	0.0%	0.1%	0.0%	0.0%	0.0%	0 ppts
Child Characteristics						
Male	48.0%	49.1%	48.9%	51.4%	49.7%	1.6 ppts
Child's age at MCS3 (years)	5.21	5.19	5.20	5.20	5.20	-0.01
Multiple birth	2.1%	3.0%	2.3%	3.0%	2.8%	0.7 ppts
Special care unit	8.7%	8.8%	9.5%	9.6%	9.3%	0.6 ppts
MCS1 White	85.5%	86.7%	92.9%	94.3%	94.0%	8.5 ppts**
MCS1 Indian	0.6%	2.3%	1.8%	1.6%	1.5%	0.8 ppts**
MCS1 Pakistani	3.1%	4.3%	1.5%	1.0%	0.4%	-2.7 ppts**
MCS1 Bangladeshi	1.3%	0.5%	0.2%	0.2%	0.0%	-1.2 ppts**
MCS1 Black Caribbean	1.9%	1.1%	0.6%	0.4%	0.4%	-1.4 ppts**
MCS1 Black African/Other	2.2%	1.2%	0.4%	0.3%	0.3%	-1.8 ppts**
MCS1 Other ethnicity	0.7%	1.1%	0.7%	0.6%	0.7%	0 ppts
MCS1 Mixed ethnicity	4.8%	2.7%	2.0%	1.8%	2.7%	-2.1 ppts**

(*table 4 cont'd*)

Family Characteristics

Mother age at MCS1 (years)	24.9	27.4	29.7	31.0	32.2	7.29**
Mother worked at MCS1 or MCS2	32.7%	59.8%	74.8%	77.7%	81.2%	48.5 ppts**
Mother worked at MCS3	30.5%	53.8%	71.6%	76.3%	74.2%	43.7 ppts**
Father worked at MCS1 or MCS2	77.7%	95.5%	98.9%	99.4%	99.7%	22.0 ppts**
Father worked at MCS3	49.7%	71.4%	83.7%	89.0%	93.6%	43.8 ppts**
Only English spoken at home	91.2%	90.1%	94.2%	94.4%	94.5%	3.3 ppts**
Lone parent at MCS1	36.7%	16.6%	6.6%	2.9%	1.3%	-35.3 ppts**
Teenage mother	13.9%	6.4%	1.6%	0.2%	0.0%	-13.8 ppts**
Cohabiting at MCS1	34.3%	33.2%	22.7%	18.8%	12.2%	-22.0 ppts**
Lone parent at MCS2	41.4%	18.8%	8.3%	4.3%	2.5%	-38.9 ppts**
Cohabiting at MCS2	25.2%	23.1%	16.1%	12.3%	8.1%	-17.0 ppts**
Lone parent at MCS3	41.4%	21.8%	11.1%	6.7%	4.3%	-37.1 ppts**
Number of siblings at MCS3	1.60	1.42	1.30	1.21	1.14	-.46**
Number of older siblings at MCS3	1.09	0.98	0.85	0.73	0.61	-.48**

Family Interactions

Mother-child relationship problems (sds)	0.21	0.04	-0.03	-0.11	-0.11	-.31**
Mother-child conflict problems (sds)	0.10	-0.01	-0.01	-0.03	-0.04	-.14**
Interviewer assessed measure of closeness (sds)	-0.23	-0.05	0.04	0.10	0.14	.37**
Relationship problems at MCS1 (sds)	0.31	0.17	0.09	-0.03	-0.07	-.37**
Relationship problems at MCS2 (sds)	0.32	0.15	0.08	-0.01	-0.06	-.38**
Mother spends plenty of time with child (MCS2)	78.7%	71.7%	62.5%	58.3%	51.1%	-27.6 ppts**
Father spends plenty of time with child (MCS2)	33.3%	20.5%	20.9%	18.4%	13.5%	-19.8 ppts**

Health and Well-Being

Tried tobreast-feed Child	50.6%	64.8%	73.7%	82.9%	89.9%	39.2 ppts**
Age at which breast-feeding stopped (weeks)	9.69	11.74	13.35	15.26	16.75	7.06**
Still breast-feeding at MCS1	10.2%	17.8%	24.4%	32.3%	38.8%	28.6 ppts**
Alcohol units consumed by mother in pregnancy	0.45	0.43	0.34	0.39	0.48	2.8 ppts
Mother alcohol consumption at wave 1 (units)	2.06	2.03	2.17	3.06	3.97	1.9**
Smoking by mother during pregnancy (# cigs)	4.61	2.47	1.19	0.50	0.29	-4.32**
Smoking by mother after pregnancy (# cigs)	8.60	5.69	3.45	1.96	1.19	-7.4**
Gestation length in days	276.1	276.8	277.2	277.3	277.5	1.43**
Birth Weight (kg)	3.26	3.33	3.39	3.42	3.45	0.18**
Infant temperament - mood (sds)	0.03	-0.01	0.06	-0.01	-0.08	-0.11**
Infant temperament - regularity (sds)	-0.32	-0.09	0.09	0.11	0.20	0.51**
Infant temperament - adaptability (sds)	-0.13	-0.05	0.03	0.02	0.13	0.26**
Mother suffered post-natal depression	21.8%	14.5%	11.8%	8.3%	6.7%	-15.1 ppts**
Mother height at birth (cm)	163.15	163.80	164.19	164.63	165.49	2.33**
Father height at birth (cm)	177.15	177.45	178.49	179.13	179.74	2.59**
Mother weight at birth (kg)	63.15	64.07	64.65	64.39	63.90	0.75 ppts
Father weight at birth (kg)	80.24	81.98	84.21	84.25	83.79	3.55**
Father under-weight	1.8%	0.8%	0.2%	0.4%	0.3%	-1.4 ppts**
Father normal-weight	37.0%	35.3%	32.7%	34.9%	41.4%	4.4 ppts**
Father over-weight	27.1%	33.0%	42.0%	45.9%	45.5%	18.3 ppts**
Father obese	11.3%	11.8%	12.9%	11.6%	9.0%	-2.3 ppts*
Mother under-weight	8.9%	6.6%	3.4%	3.4%	2.6%	-6.3 ppts**
Mother normal-weight	54.4%	57.9%	63.8%	67.5%	72.3%	17.9 ppts**
Mother over-weight	17.3%	20.9%	19.9%	18.7%	16.8%	-.5 ppts
Mother obese	10.9%	9.7%	9.5%	7.7%	5.6%	-5.2 ppts**

(Table 4 cont'd)

Childcare

Cared for by grandparents MCS2	17.0%	24.3%	31.4%	32.6%	25.7%	8.7 ppts**
Cared for by other relatives MCS2	4.5%	7.3%	5.7%	5.0%	3.7%	-.8 ppts
Cared for by friends/neighbours MCS2	1.7%	0.9%	1.1%	1.1%	2.0%	.3 ppts
Has been to nursery school/class MCS2	27.3%	21.8%	21.1%	22.0%	23.7%	-3.5 ppts**
Has been to playgroup MCS2	25.2%	30.0%	31.1%	30.5%	25.7%	.4 ppts
Has been to pre-school MCS2	9.8%	13.6%	18.8%	17.7%	17.5%	7.7 ppts**
Has been to childminder MCS2	3.1%	5.6%	7.4%	11.3%	11.2%	8.1 ppts**
Has been to day nursery or creche MCS2	7.1%	8.8%	11.9%	17.3%	23.6%	16.5 ppts**
Cared for by grandparents MCS3	22.0%	27.5%	29.7%	32.6%	27.5%	5.5 ppts**
Cared for by relatives MCS3	10.2%	10.3%	7.8%	6.3%	3.5%	-6.7 ppts**
Cared for by friends MCS3	5.0%	6.2%	6.4%	8.0%	11.4%	6.4 ppts**
Has Been to nursery school/class MCS3	66.4%	60.1%	53.1%	53.7%	49.4%	-16.9 ppts**
Has Been to playgroup MCS3	20.6%	25.5%	26.9%	25.9%	21.6%	.9 ppts
Has Been to pre-school MCS2	13.9%	21.7%	28.5%	30.7%	29.9%	15.9 ppts**
Has Been to childminder MCS3	3.0%	4.2%	5.8%	7.4%	5.6%	2.5 ppts**
Has Been to day nursery or creche MCS3	6.2%	7.6%	10.6%	14.7%	20.4%	14.2 ppts**

Home-Learning Environment

Bottom HLE quintile at MCS2	31.8%	25.7%	22.1%	20.7%	17.5%	-14.3 ppts**
2nd HLE quintile at MCS2	16.8%	18.7%	17.8%	17.0%	16.5%	-.3 ppts
3rd HLE quintile at MCS2	19.5%	20.5%	21.2%	20.5%	22.8%	3.3 ppts**
4th HLE quintile at MCS2	17.4%	16.6%	17.4%	20.3%	20.5%	3.1 ppts**
Top HLE quintile at MCS2	14.5%	18.5%	21.5%	21.5%	22.7%	8.1 ppts**
Read to everyday at MCS2	43.1%	51.8%	63.3%	72.4%	78.6%	35.5 ppts**
Read to some days at MCS2	45.5%	40.1%	32.9%	25.2%	20.0%	-25.4 ppts**
Bottom HLE quintile at MCS3	27.2%	25.2%	24.2%	19.8%	18.2%	-8.9 ppts**
2nd HLE quintile at MCS3	18.5%	16.8%	19.8%	18.3%	18.4%	-.1 ppts
3rd HLE quintile at MCS3	19.5%	21.1%	21.7%	23.7%	22.6%	3 ppts*
4th HLE quintile at MCS3	16.0%	18.5%	17.3%	20.3%	21.1%	5 ppts**
Top HLE quintile at MCS3	18.7%	18.5%	17.0%	17.9%	19.7%	.9 ppts
Read to everyday at MCS3	42.1%	47.7%	51.4%	57.0%	62.7%	20.5 ppts**
Read to some days at MCS3	48.8%	47.1%	43.8%	40.5%	36.1%	-12.7 ppts**
Mother rates herself as good parent MCS2	15.6%	20.8%	25.3%	28.5%	37.0%	21.3 ppts**
Mother rates herself as very good parent MCS2	26.9%	29.6%	28.0%	28.7%	26.7%	-.1 ppts
Father rates himself as good parent MCS2	17.5%	19.2%	25.9%	29.9%	36.7%	19.2 ppts**
Father rates himself as very good parent MCS2	27.1%	30.7%	31.7%	30.6%	30.4%	3.2 ppts

Parenting Style/Rules

Lots of rules - MCS2	27.4%	28.0%	31.1%	32.9%	34.3%	6.8 ppts**
Strictly enforced rules - MCS2	42.7%	46.0%	47.2%	52.4%	58.0%	15.3 ppts**
Regular bed-times at MCS2	68.7%	76.0%	82.1%	87.7%	92.1%	23.4 ppts**
Regular meal-times at MCS2	85.1%	89.4%	94.1%	96.2%	98.0%	12.8 ppts**
Watches > 3 hours TV a day - MCS2	29.9%	22.0%	14.5%	8.6%	6.0%	-23.8 ppts**
Watches > 3 hours TV a day - MCS3	21.4%	16.6%	13.2%	11.3%	8.2%	-13.2 ppts**
Plays computer > 1 hour a day - MCS3	31.0%	25.0%	20.8%	16.6%	11.3%	-19.7 ppts**
Regular bed-times at MCS3	84.2%	88.6%	92.2%	95.2%	96.4%	12.2 ppts**
Regular meal-times at MCS3	88.0%	92.4%	94.8%	97.3%	97.6%	9.6 ppts**
Eat breakfast together at MCS3	86.6%	90.6%	93.9%	96.1%	97.1%	10.4 ppts**

*** and * indicate significant differences between Q1 and Q5 at the 1% and 5% levels, respectively.*

4. Explaining socio-economic gradients in child outcomes

In this section, we examine how much of the observed socio-economic differences in child cognitive outcomes at the ages of 3 and 5 can be explained by measures of the early childhood caring environment, including: family interactions; health and well-being; (non-maternal) childcare; the home-learning environment; and parenting style and rules; and how much remains related to distal factors such as socio-economic position and other aspects of family background.

Empirical Methodology

In order to decompose the socio-economic gap in early cognitive outcomes into the contribution from different sources, we use the decomposition methodology described in detail in Goodman, Gregg and Washbrook in this Special Issue.

We first set out the raw differences in each educational outcome, according to socio-economic position of the parents, focusing on the gap between the top and bottom quintiles (top-bottom gap) as well as the gap between the middle and bottom quintiles (middle-bottom gap). All gaps are expressed in percentile point terms.

We then attempt to isolate the role of each factor in explaining the socio-economic gradient in child outcomes, after controlling for all observable characteristics, by decomposing the raw gap in child outcomes into the amount explained by each variable. This is calculated as the coefficient on each variable (taken from a regression including all observable characteristics) multiplied by unconditional differences in that variable across quintiles of our socio-economic position index. We group similar mediating factors together. The sum of the amount explained by all groups of potential mediating factors, represents the total amount explained by observable characteristics; the rest of the raw gap is thus unexplained.

The relationships we estimate are unlikely to be causal. For our estimates to be the causal impact of these factors, we would have to argue (among other things) that our measures of family characteristics, and in particular our mediating factors, cannot be affected by the child outcomes of interest. However, it is highly likely that factors, such as the child's home learning environment, are going to be influenced by the child's cognitive development. We do not attempt to take into account this possible

simultaneity. Furthermore, we would also have to argue that there are no other unobserved characteristics of the child or family, that influence these mediating factors, as well as the child outcomes we measure. Again, this is unlikely to be true and this means that our estimates of the impact of different mediating factors are likely to be biased.[iv]

While our work (along with most other work in this area) cannot robustly establish the presence of direct causal links between these factors, we are fortunate to have an extremely rich dataset at our disposal, that allows us to observe in great detail a wide range of family background variables, family health and well-being, as well as parenting activities, relationships and behaviours, that serve as plausible transmission mechanisms between child poverty and poor early educational attainment. This should give us some clues as to possible policy responses to address the socio-economic gaps in early child outcomes, as well as avenues for future research.

We start by examining BAS vocabulary scores at age 3, before moving on to age 5 BAS vocabulary scores (with and without controls for prior scores at age 3). The full regression results on which these decompositions are based are shown in Table A1 in the Appendix.

Age 3 - BAS (Vocabulary)

In the first row of Table 5, we show the raw gap at age 3 in BAS average percentile ranks between the top and bottom quintiles of our index of socio-economic position (21.9 ppts), as well as the gap between the middle and bottom quintiles (13.5 ppts). Both are statistically significant at the 1% level. The next two rows then show that about 69% of the socio-economic gaps in BAS scores at age 3 can be explained by differences in observable characteristics; whilst 31% remains unexplained. The rest of the table then decomposes the 69% explained by all observable characteristics, into how much is explained by differences in the early childhood caring environment, and how much remains related to parental education and family background (as well as a separate group for missing dummy variables). Within each group, we also further decompose the amount explained by each group into the amount explained by individual (or very similar) variables. We observe that:

- **Parental Education** explains about one sixth of the socio-economic gaps in BAS scores,

after controlling for all other observable characteristics. This is the amount that remains related to differences in parental education, which cannot be accounted for by differences in the early childhood caring environment.

- **Family Background** characteristics still explain about a quarter of the socio-economic gaps in BAS scores at age 3, conditional on all other observable characteristics. Looking at the detailed results, we see that 15-17% of the gaps are explained by the fact that mothers of children from higher quintiles are, on average, older than those from lower quintiles. Similarly, the facts that the top quintiles contains greater proportions of working fathers and that these children have fewer siblings, explain 7-10% and 6-8% of the socio-economic gaps, respectively. The age of children explains about 10% of the gaps, which means that the poorer children must have been tested at slightly later ages, since age has a positive effect on BAS percentiles.
- **Family Interactions** explain 4-5% of the socio-economic gaps, with the largest single amount being explained by greater levels of mother-child closeness among families in higher quintile.
- **Health and Well-Being** factors similarly explain about 3-5% of the gaps, with most of the positive contribution coming from breast-feeding and birth-weight, offset by some negative factors (smoking before pregnancy and parental height/weight).
- **Childcare** patterns explain about 1% of the top-bottom gap, but about 4% of the middle-bottom gap.
- **The Home-Learning Environment** explains one sixth of the socio-economic gaps, most of which is done by the HLE index and reading frequency at age 3.
- **Parental Style/Rules** explains very little of the socio-economic gaps at age 3 in BAS percentile ranks.

Therefore, this decomposition shows that about a quarter of the socio-economic gaps in BAS scores at age 3 can be explained by differences in the early childhood caring environment, with the home-learning environment explaining about one sixth, and family interactions and health/well-being factors explaining about 5% each. However, around a quarter of the gap remains related to family background characteristics, mostly down to the fact that in higher quintiles, mothers are older, children have fewer siblings and fathers are more likely to be in work. And about a sixth remains related to differences in parental education, controlling for other observable characteristics. Around 30% of the gap cannot be explained by differences in any of the observable characteristics.

Table 5. Explaining socio-economic gaps in cognitive ability at age 3 (BAS)

	Percentile point gap		As % total gap	
	Q5-Q1	Q3-Q1	Q5-Q1	Q3-Q1
Raw Gap	**21.90**	**13.54**	**100%**	**100%**
Total: Explained	**15.02**	**9.66**	**69%**	**71%**
Total: Unexplained	**6.88**	**3.88**	**31%**	**29%**
Amount Explained by Factors:				
Parental Education	***3.81***	***2.02***	***17%***	***15%***
Family Background	***5.19***	***3.72***	***24%***	***27%***
Gender	-0.11	-0.06	-1%	0%
Age of child	-1.89	-1.42	-9%	-10%
Twin/triplet	0.00	0.00	0%	0%
Special care unit after birth	-0.01	-0.01	0%	0%
Ethnicity	0.55	0.39	3%	3%
Only English spoken at home	0.38	0.35	2%	3%
Country of residence	-0.06	0.02	0%	0%
Mother works	0.39	0.34	2%	3%
Father works	1.46	1.31	7%	10%
Mother's age at birth	3.35	2.35	15%	17%
Marital/partner Status	-0.52	-0.38	-2%	-3%
Siblings	1.67	0.84	8%	6%
Family Interactions	***0.83***	***0.62***	***4%***	***5%***
Mother-child closeness	0.99	0.73	5%	5%
Parental harmony	0.00	0.00	0%	0%
Parental time	-0.16	-0.11	-1%	-1%
Health and Well-Being	***1.02***	***0.43***	***5%***	***3%***
Breast-feeding	0.72	0.40	3%	3%
Alcohol consumption	0.13	0.01	1%	0%
Smoking patterns	-0.13	-0.01	-1%	0%
Gestation length	-0.04	-0.04	0%	0%
Birth weight	0.54	0.39	2%	3%
Infant temperament	-0.01	-0.09	0%	-1%
Maternal depression	-0.09	-0.06	0%	0%
Parental height/weight	-0.10	-0.17	0%	-1%
Childcare	***0.33***	***0.52***	***1%***	***4%***
Home-Learning Environment	***3.60***	***2.18***	***16%***	***16%***
HLE and reading at Age 3	2.64	1.62	12%	12%
Self-reported parental competence	0.96	0.55	4%	4%
Parenting Style/Rules	***-0.04***	***-0.08***	***0%***	***-1%***
Amount/strictness of rules	0.02	-0.01	0%	0%
Regular bed times at age 3	0.47	0.27	2%	2%
Regular meal times at age 3	-0.05	-0.03	0%	0%
Watches lots of TV at age 3	-0.49	-0.32	-2%	-2%
Missing Dummies	***0.28***	***0.26***	***1%***	***2%***

Table 6. Explaining socio-economic gaps in cognitive ability at age 5 (BAS)

	No controls for prior ability				Controlling for prior ability			
	ppt terms		As % total gap		ppt terms		As % total gap	
	Q5-Q1	Q3-Q1	Q5-Q1	Q3-Q1	Q5-Q1	Q3-Q1	Q5-Q1	Q3-Q1
Raw Gap	*26.57*	*14.19*	100%	100%	26.57	14.19	100%	100%
Total Explained	19.87	11.94	75%	84%	23.58	14.05	89%	99%
Total Unexplained	6.70	2.25	25%	16%	2.99	0.14	11%	1%
Amount Explained by Factors:								
Prior Cognitive Ability					*13.40*	*8.14*	*50%*	*57%*
Prior Non-Cognitive Ability					*0.31*	*0.21*	*1%*	*1%*
Parental Education	*6.68*	*3.09*	*25%*	*22%*	*4.09*	*1.89*	*15%*	*13%*
Family Background	*9.42*	*6.70*	*35%*	*47%*	*5.29*	*3.69*	*20%*	*26%*
Gender	-0.01	-0.01	0%	0%	0.04	0.02	0%	0%
Age of child	0.02	0.03	0%	0%	0.02	0.03	0%	0%
Twin/triplet	-0.03	-0.01	0%	0%	-0.02	-0.01	0%	0%
Special care unit after birth	0.00	0.00	0%	0%	0.00	0.00	0%	0%
Ethnicity	0.33	0.28	1%	2%	0.07	0.10	0%	1%
Only English spoken at home	0.44	0.40	2%	3%	0.27	0.25	1%	2%
Country of residence	0.04	0.03	0%	0%	0.07	0.04	0%	0%
Mother works	0.41	0.35	2%	2%	0.11	0.08	0%	1%
Father works	0.69	0.55	3%	4%	-0.30	-0.27	-1%	-2%
Mother's age at birth	4.07	2.86	15%	20%	2.52	1.75	9%	12%
Marital/partner Status	1.35	1.12	5%	8%	1.42	1.13	5%	8%
Siblings	2.11	1.11	8%	8%	1.10	0.58	4%	4%
Family Interactions	*0.77*	*0.60*	*3%*	*4%*	*0.14*	*0.16*	*1%*	*1%*
Mother-child closeness	0.74	0.55	3%	4%	0.16	0.12	1%	1%
Parental harmony	-0.01	0.01	0%	0%	-0.02	0.02	0%	0%
Parental time	0.04	0.04	0%	0%	0.00	0.02	0%	0%
Health and Well-Being	*-1.10*	*-1.11*	*-4%*	*-8%*	*-1.04*	*-0.89*	*-4%*	*-6%*
Breast-feeding	1.04	0.52	4%	4%	0.60	0.29	2%	2%
Alcohol consumption	0.15	0.07	1%	0%	0.10	0.06	0%	0%
Smoking patterns	-1.09	-0.73	-4%	-5%	-0.90	-0.63	-3%	-4%
Gestation length	-0.07	-0.07	0%	0%	-0.09	-0.08	0%	-1%
Birth weight	0.29	0.21	1%	1%	0.11	0.08	0%	1%
Infant temperament	0.26	0.13	1%	1%	0.17	0.09	1%	1%
Maternal depression	-0.03	-0.02	0%	0%	0.05	0.03	0%	0%
Parental height/weight	-1.64	-1.22	-6%	-9%	-1.08	-0.74	-4%	-5%
Childcare	*0.55*	*0.47*	*2%*	*3%*	0.25	0.26	1%	2%
Home-Learning Environment	*2.06*	*1.16*	*8%*	*8%*	*0.18*	*-0.03*	*1%*	*0%*
HLE and reading at Age 3	1.97	1.21	7%	9%	0.47	0.24	2%	2%
HLE and reading at Age 5	0.04	0.03	0%	0%	0.11	0.04	0%	0%
Self-reported parental competence	0.04	-0.08	0%	-1%	-0.40	-0.31	-1%	-2%
Parenting Style/Rules	*1.24*	*0.71*	*5%*	*5%*	*0.88*	*0.51*	*3%*	*4%*
Amount/strictness of rules	0.15	0.04	1%	0%	0.10	0.03	0%	0%
Regular bed times at age 3	0.31	0.18	1%	1%	0.06	0.03	0%	0%
Regular bed times at age 5	0.33	0.22	1%	2%	0.37	0.24	1%	2%
Regular meal times	0.34	0.24	1%	2%	0.08	0.06	0%	0%
Watches lots of TV/Computer	0.10	0.04	0%	0%	0.27	0.15	1%	1%
Missing Data	*0.27*	*0.32*	*1%*	*2%*	*0.09*	*0.12*	*0%*	*1%*

Age 5 – BAS (Vocabulary) – Static Model

Table 6 then shows the results of our decomposition methodology for analysing the socio-economic differences in BAS scores at age 5. As we have already shown, the top-bottom gap has widened to 26.6 percentile points, and the middle-bottom gap to 14.2 percentile points. The four columns under the heading "No controls for prior ability" show the results of our static decomposition. Specifically, it shows that 84% of the middle-bottom gap and 75% of top-bottom gap can be accounted for by differences in observable characteristics. The specific amounts explained by different variables are as follows:

- **Parental Education** still explains about a quarter of the socio-economic gaps, after controlling for other observable characteristics. This is more than was the case at age 3.
- **Family Background** characteristics still explain a further 35% of the top-bottom gap, but nearly 47% of the middle-bottom gap. The most important individual variables were mother's age (15-20% of the gaps), number of siblings (8%), whether the father was in work (3-4%) and marital/partner status (6-9%). Only the latter was not important in explaining gaps at age 3.
- **Family Interactions** explain about 3-4% of the socio-economic gaps in BAS scores at age 5, with the most important variables again being those related to mother-child closeness. This is similar to what we found for BAS scores at age 3.
- **Health and Well-Being** factors explain a negative amount of the gap (i.e. they increase it) with positive contributions (from breast-feeding patterns) more than offset against negative ones (smoking patterns and parental height/weight).
- **Childcare** only explains about 2-3% of the socio-economic gaps.
- **The Home-Learning Environment** explains 8% of the socio-economic gaps. Interestingly, this is solely down to differences in the home-learning environment and reading frequency at age 3. The differences at age 5 are found to be unimportant.
- **Parenting Style/Rules** make a further small contribution of 5%, with small contributions from all of the individual variables.

Age 5 - BAS (Vocabulary) – Value added model

The last four columns of Table 6 show the results of our decomposition methodology when we also include controls for prior cognitive ability and socio-emotional development (specifically, the Strengths and Difficulties questionnaire). This shows that observable characteristics explain 89% of the top-bottom gap and fully explain the middle-bottom gap. Specifically:

- **Prior cognitive ability** explains the majority of both socio-economic gaps. It explains 57% of the middle-bottom gap and 50% of the top-bottom gap.
- **Prior socio-emotional development**, however, only explains 1% of the gaps.
- **Parental Education** still explains a further 13-15% of the socio-economic gaps, conditional on all other observable characteristics and prior outcomes.
- **Family Background** characteristics still explain a further 20% of the top-bottom gap and 26% of the middle-bottom gap, conditional on other observable characteristics and prior outcomes. The individual variables making the largest contributions, are mother's age at birth, marital/partner status and number of siblings. However, the importance of mother's age at birth and number of siblings are halved as compared with the static decomposition at age 5, suggesting that part of their impact at age 5 is via higher levels of cognitive ability at age 3.
- **Family Interactions** only explain about 1% of the socio-economic gaps, conditional on other observable characteristics and prior outcomes.
- **Health and Well-Being** factors make a small negative contribution, overall. There are positive contributions from breast-feeding patterns, but negative ones from smoking patterns and parental height/weight.
- **Childcare** only contributes a further 1-2% to the socio-economic gap.
- **The Home-Learning Environment** hardly contributes anything at all, conditional on other observable characteristics and prior outcomes. Since it was important in the static decomposition, this suggests that differences in the home-learning environment only explain gaps at age 5 via improving cognitive ability at age 3.
- **Parenting Style/Rules** explain a further 3-4% of the gap, though no one variable appears to be particularly important.

Summary of Results

In total, observable characteristics explain about 65-75% of the socio-economic gaps at age 3, and 77-87 % of socio-economic differences in cognitive outcomes at age 5.

At age 3, differences in the early childhood caring environment explain about 25% of the socio-economic gap. This can mostly be accounted for by differences in the home-learning environment and reading frequency. Other proximal factors, like family interactions and health and well-being, explain a small proportion of the socio-economic gap. However after conditioning on this rich set of proximal factors, it is family background factors that still explain the largest portion of socio-economic differences. Looking at individual family background factors, it is differences in mother's age, number of siblings and working patterns, that are found to explain the largest element of these socio-economic differences. The next largest contribution comes from the remaining effects of parental education that do not occur via the observable characteristics.

When we look at age 5 cognitive outcomes, we find that prior cognitive outcomes explain over 50% of the socio-economic differences at age 5, whilst prior socio-emotional development explains very little, if anything. The only other factors that explain a large proportion of the socio-economic gaps, after controlling for prior ability, are parental education and family background (again, mother's age, number of siblings and marital/partner status being most important). The influence of these items is much reduced compared with the static model, suggesting some of their impact occurs via their effect on cognitive ability at age 3. The home-learning environment is found to be unimportant in this decomposition, suggesting that it only explains age 5 outcomes through its influence on age 3 cognitive outcomes.

Conclusions

In conclusion, we have confirmed that there are substantial differences in cognitive and socio-emotional development between children from rich and poor backgrounds even at the age of 3, and that this gap widens by the age of 5. Children from poor backgrounds also face much less advantageous "early childhood caring environments" than children from better off families. For example, we have observed significant differences in poor children's and their mothers' health and well-being and the home learning environment.

We also find that differences in the home learning environment at age 3 have an important role to play, explaining about a sixth of the gap in cognitive outcomes between children from richer and poorer backgrounds. However, a much bigger proportion of the socio-economic gap remains directly related to other aspects of family background (such as mothers' age, and family size) that are not mediated through the early childhood caring environment, and a significant element remains explained.

It is noteworthy that it is the home-learning environment measured at age 3, that is found to be important in explaining outcomes at ages 3 and 5, the latter working through its impact on cognitive ability at age 3. The home-learning environment measured at age 5 is not estimated to impact on cognitive outcomes at age 5, or thus the gap in cognitive outcomes at age 5. This stresses the importance of early intervention. However, it is difficult to know with certainty whether policymakers can, a) change the home-learning environment, and (b) whether any shifts in the home-learning environment will reduce the gap in early child outcomes. This is partly because it is difficult to put a definite causal interpretation to our finding, and because the malleability of the home-learning environment to outside policy intervention is currently unknown. We therefore believe it to be essential that different methods of shifting the home-learning environment at early ages are trialled and evaluated in the UK at the earliest opportunity.

However, it is also worth noting that pre-determined factors still explain the largest element of the socio-economic divides in cognitive outcomes at age 3 and age 5, even when we condition on a rich set of measures of the early childhood caring environment. The most important factors being mother's age at birth, number of siblings, parental education and prior cognitive development (at age 3). With a view to closing socio-economic gaps in cognitive outcomes, these results underline the importance of early intervention, at least before age 3 and perhaps even prior to birth, if one believed the results that would suggest encouraging poorer mothers to delay the birth of their first child might narrow some of the socio-economic gap in early cognitive development.

Taken together, our findings suggest that policies to improve parenting skills and home learning environments in isolation cannot possibly eliminate the cognitive skills gap between rich and poor young children, though such policies could go some way towards reducing it. On the other hand, in our working paper we show that many aspects of the early childhood caring environment do have a positive effect on children's social and emotional development, suggesting that policies aimed at improving health, parenting skills and the home-learning environment could have other important short- and long-term pay-offs.

Acknowledgements

References

Adler NE, Boyce WT, Chesney MA, Folkman S and Syme SL. (1993) Socioeconomic inequalities in health: No easy solution. *Journal of the American Medical Association*, 269, 3140-3145.

Belfield CR, Nores M, Barnett S and Schweinhart L. (2006) The High/Scope Perry Preschool Program: Cost–Benefit Analysis Using Data from the Age-40 Followup, *Journal of Human Resources*, University of Wisconsin Press, 41, 162-190.

Brooks-Gunn J, Duncan GJ, Klebanov PK and Sealand N. (1993) Do neighborhoods influence child and adolescent behavior? *American Journal of Sociology*, 99, 335-95.

Brooks-Gunn J, McCarton C, Casey P, McCormick M, Bauer C, Bernbaum J, Tyson J, Swanson M, Bennett F, Scott D, Tonascia J and Meinert C. (1994) Early intervention in low birth weight, premature infants: Results through age 5 years from the Infant Health and Development Program. *Journal of the American Medical Association*, 272, 1257-1262.

Burchinal MR, Campbell FA, Bryant DM, Wasik BH and Ramey CT. (1997) Early intervention and mediating processes in cognitive performance of children of low-income African-American families. *Child Development*, 68, 935-954.

Dearden L, Mesnard A and Shaw J. (2006) Ethnic differences in birth outcomes in England. *Fiscal Studies*, 27, 17-46.

Dearden L, Sibieta L and Sylva K. (2010) The socio-economic gradient in early child outcomes: evidence from the Millennium Cohort Study. *IFS Working Paper*.

Duncan GJ, Brooks-Gunn J and Klebanov PK. (1994) Economic deprivation and early-childhood development. *Child Development*, 65, 296-318.

Feinstein L. (2003) Inequality in the Early Cognitive Development of British Children in the 1970 Cohort. *Economica,* 70, 73-97.

Feinstein L. (2004) Mobility in Pupils' Cognitive Attainment During School Life. *Oxford Review of Economic Policy*, 20, No. 2: Education.

Garces E, Thomas D and Currie J. (2002) Longer-Term Effects of Head Start. *American Economic Review*, American Economic Association, 92, 999-1012.

Goldstein H. (1990) Explaining socioeconomic differences in children's cognitive test scores. *Working Paper No. H-90-1. Cambridge, MA*: Malcolm Wiener Center for Social Policy, John F Kennedy School of Government, Harvard University.

Goodman A and Gregg P. (eds) (2010) Poorer children's educational attainment: how important are attitudes and behaviour? Joseph Rowntree Foundation: York.

Hofferth S and Phillips DA. (1991) Childcare policy research. *Journal of Social Issues*, 47 1-13.

Howes C. (1988) Relations between early child care and schooling. *Developmental Psychology*, 24, 53-57.

Korenman S, Miller JE and Sjaastad JE. (1995) Long-term poverty and child development in the United States: Results from the National Longitudinal Survey of Youth. *Children and Youth Services Review,* 17, 127-51.

Lazar I and Darlington RB. (1982) Lasting effects of early education: A report from the consortium for longitudinal studies. *Monographs of the Society for Research in Child Development*, 47(2-3, Serial No. 195).

Ludwig J and Miller DL. (2007) Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design. *Quarterly Journal of Economics*, MIT Press, 122, 159-208.

Melhuish EC, Sylva K, Sammons P, Siraj-Blatchford I, Taggart B and Phan M. (2008) Effects of the Home Learning Environment and preschool center experience upon literacy and numeracy development in early primary school. *Journal of Social Issues, 64*, 157-188.

McLeod JD and Shanahan MJ. (1993) Poverty, parenting and children's mental health. *American Sociological Review,* 58, 351-66.

NICHD Early Child Care Research Network. (1997) The effects of infant child care on infant-mother attachment security: Results of the NICHD study of early child care. *Child Development*, 68, 860-879.

NICHD Early Child Care Research Network. (1998) Relations between family predictors and child outcomes: Are they weaker for children in child care? *Developmental Psychology*, 34, 1119-1128.

Ramey CT and Ramey SL. (1998) Prevention of intellectual disabilities: Early interventions to improve cognitive development. *Preventive Medicine*, 27, 224-232.

Sylva K, Melhuish E, Sammons P, Siraj-Blatchford I and Taggart B. (2008) Effective Pre-school and Primary Education 3-11 Project (EPPE 3-11): Final Report from the Primary Phase: Pre-school, School and Family Influences on Children's Development during Key Stage 2 (Age 7-11). Department for Children, Schools and Families, London.

Appendix A

Table A1 – Full Specification Regression Results for BAS at age 3 and age 5

	BAS (age 3)	BAS (age 5)	
		Levels	Value Added
Prior Ability			
Bracken (age 3)	n/a	n/a	6.796***
BAS (age 3)	n/a	n/a	7.985***
SDQ (age 3)	n/a	n/a	0.384
SEP Quintile			
2nd SEP quintile	1.145	1.118	0.385
3rd SEP quintile	4.397***	2.254**	0.14
4th SEP quintile	6.331***	4.636***	1.397
Top SEP quintile	7.714***	6.696***	2.992**
Parental Education			
Mother NVQ level 1	1.92	1.372	0.492
Mother NVQ level 2	3.150***	2.344**	0.788
Mother NVQ level 3	3.183**	1.87	0.165
Mother NVQ level 4/5	5.141***	5.549***	2.650**
Mother - other qualifications	0.865	1.489	1.152
Father NVQ level 1	-0.767	1.34	0.98
Father NVQ level 2	0.496	1.997*	1.911*
Father NVQ level 3	1.298	3.582***	2.764***
Father NVQ level 4/5	1.794*	5.528***	3.903***
Father - other qualifications	5.473*	-10.889	-8.646
Child Characteristics			
Male	-6.764***	-0.543	6.796***
Child's age (months/100)	0.911***	-0.539*	7.985***
Multiple birth	-0.44	-3.917	0.384
Special care unit	-1.658	-0.384	2.441***
MCS1 Indian	-5.970**	1.207	-0.463
MCS1 Pakistani	-9.127***	-3.855*	-2.582
MCS1 Bangladeshi	-11.527***	-0.599	0.127
MCS1 Black Caribbean	-8.799***	-6.565**	2.611
MCS1 Black African/Other d	-2.936	-5.719**	1.691
MCS1 Other ethnicity	-9.867***	-3.637	6.148
MCS1 Mixed ethnicity	-0.811	-0.378	-2.657
Family Characteristics	11.358***		
Mother's age at birth	0.373	1.661***	0.940**
Mother's age at birth squared	4.795***	-0.020***	-0.011*
Only English at Home	4.129***	13.051***	-5.678**
Lives in Wales	0.802	-2.936***	-2.239
Lives in Scotland	n/a	2.840**	-0.999
Lives in Northern Ireland	3.084**	0.941	8.078***
Mother worked at one of waves	n/a	1.119	-3.520***
Mother worked at wave 3		-0.307	0.395
Father worked at one of waves	1.329***	0.291	-0.634
Father worked at Wave 3	-0.016*	0.968	0.721
Lone parent at MCS1	0.665	-2.335	-2.318
Had baby in teens	0.226	-0.09	-0.562
Two cohabiting parents at MCS1	1.192	-0.963	-1.592**
Got married by MCS2	-0.063	1.3	1.251
Split up by MCS2	-2.493**	-2.173	-2.867*
New partner by MCS2	-1.088	0.505	2.079
Split up by MCS3	n/a	-1.689	-0.055
New partner by MCS3	n/a	0.941	-0.064
Number of siblings at MCS 2/3	0.199	-0.408	-0.189
Number of older siblings at MCS 2/3	-3.679***	-3.984***	-2.100***

Family Interactions

Mother-child relationship problems (sds)	-2.596***	-1.755	1.55
Mother-child conflict problems (sds)	-0.496***	2.389	2.778
Interviewer assessed measure of closeness (sds)	0.872***	-0.267	0.482
Relationship problems at MCS1 (sds)	-0.028	-4.535**	-1.946
Relationship problems at MCS2 (sds)	n/a	0.735**	0.065
Mother spends plenty of time with child - MCS2	1.008	-3.253	-2.91
Father spends plenty of time with child - MCS2	-0.713	-0.212	-0.933

Health and Well-Being

Tried to breast-feed Child	1.102	0.26	0.696
Age at which breast-feeding stopped (weeks)	0.002	0.091	2.733*
Still breast-feeding at MCS1	0.992	0.055	-0.316
Mother alcohol consumption during pregnancy (units)	0.159	-0.023	-0.108
Mother alcohol consumption (small amount)	0.581	1.681***	0.073
Mother alcohol consumption at Wave 1 (units)	0.033	-0.026	-0.347
Number of cigarettes smoked by Mother during	-0.279*	-0.195	-0.121
(squared)	0.007	0.008	1.379**
Number of cigarettes smoked by mother before	0.222*	0.287**	-0.033
(squared)	-0.005	-0.005	-0.076
Gestation length in days	0.639**	0.488	0.006
Gestation length in days (squared)	-0.001**	-0.001*	0.17
Birth weight (kg)	2.893***	1.584**	-0.003
Infant temperamen - mood (sds)	-0.619**	-0.518*	0.211
Infant temperament - regularity (sds)	-0.208	0.278	-0.001
Infant temperament - adaptability (sds)	0.724**	0.222	0.616
Mother suffered post-natal depression	0.83	0.223	0.279
Mother height at birth (cm)	-0.018	0.017	5.148
Father height at birth (cm)	-0.036*	-0.03	-0.322
Mother weight at birth (kg)	0.005	-0.027	-4.542
Father weight at birth (kg)	0.082*	0.039	0.007
Father under-weight	-1.113	2.407	-0.008
Father over-weight	-2.125**	-2.325***	-0.004
Father obese	-2.653	-0.647	-0.002
Mother under-weight	-0.961	-0.473	2.735
Mother over-weight	-1.431	-0.105	-0.998
Mother obese	-0.512	1.538	1.013

Childcare

Cared for by grandparents MCS2	1.351*	1.096	0.953
Cared for by other relatives MCS2	0.334	-1.01	-0.879
Cared for by friends/Neighbour MCS2	-0.233	5.077*	4.416*
Has Been to nursery school/class MCS2	0.243	-0.414	-2.104***
Has Been to playgroup MCS2	1.045	-1.454*	-2.168***
Has Been to pre-school MCS2	2.239**	-0.778	-2.113**
Has Been to childminder MCS2	1.094	1.763	1.087
Has Been to day nursery or creche MCS2	-0.621	1.624	0.382
Cared for by grandparents MCS3	n/a	0.491	-0.094
Cared for by relatives MCS3	n/a	0.55	0.917
Cared for by friends MCS3	n/a	-0.72	-0.635
Has Been to nursery school/class MCS3	n/a	-1.215	-0.219
Has Been to playgroup MCS3	n/a	1.274	2.051**
Has Been to pre-school MCS2	n/a	0.664	0.696
Has Been to childminder MCS3	n/a	1.869	2.733*
Has Been to day nursery or creche MCS3	n/a	-1.753	-0.316

Home-Learning Environment

2nd HLE quintile at MCS2	1.966**	1.144	-0.473
3rd HLE quintile at MCS2	3.239***	1.419	-0.687
4th HLE quintile at MCS2	4.045***	2.290**	-0.65
5th HLE quintile at MCS2	5.292***	3.847***	-0.515
Read to everyday at MCS2	7.487***	5.488***	0.812
Read to some days at MCS2	2.495*	1.605	-1.024
2nd HLE quintile at MCS3	n/a	0.241	0.731
3rd HLE quintile at MCS3	n/a	0.514	1.138
4th HLE quintile at MCS3	n/a	-1.213	-0.744
5th HLE quintile at MCS3	n/a	-0.054	1.307
Read to everyday at MCS3	n/a	-0.309	-0.186
Read to some days at MCS3	n/a	-1.212	-1.08
Mother rates herself as good parent – MCS 2/3	2.075***	1.366*	0.173
Mother rates herself as very good parent – MCS 2/3	0.68	-1.304*	-1.923***
Father rates himself as good parent – MCS 2/3	1.268*	-0.386	-0.999
Father rates himself as very good parent – MCS 2/3	1.240*	-1.083	-1.249*
Parenting Style/Rules			
Lots of rules - MCS2	-0.875	-0.292	-0.024
Strictly enforced rules - MCS2	0.554	1.138**	0.643
Regular bed-times at MCS2	1.685**	1.326	0.251
Regular meal-times at MCS2	-0.277	0.914	0.307
Watches > 3 hours TV a day - MCS2	2.291***	0.394	-0.159
Watches > 3 hours TV a day - MCS3	1.145	0.146	-0.084
Plays computer > 1 hour a day - MCS3	n/a	-1.102	-1.140*
Regular bed-times at MCS3	n/a	2.699**	3.005***
Regular meal-times at MCS3	n/a	1.166	0.176
Eat breakfast together at MCS3	n/a	1.02	0.228
Observations	11054	11054	11054
R-Squared	0.23	0.21	0.36

***, ** and * indicate significance at the 1%, 5% and 10% levels, respectively.

Endnotes

[i] At the time of writing, the fourth sweep has only very recently become available.

[ii] Please note that the proportion of children in the "bottom 40%" is not exactly 40%. This is due to the fact that test scores are only semi-continuous.

[iii] This may be related to ethnic differences in birth weight. Dearden, Mesnard and Shaw 2006, show that ethnic minorities tend to have lower birth weights.

[iv] Without the help of any experimental variation in SEP, parental income, or indeed any mediating factors of interest in this project, we cannot rule out that there are unobservable factors (such as genetics, or typically unobserved attributes such as motivation of parent) that instead explain low incomes, poor achievement, and the potential transmission pathways we have highlighted. Structural Equation Models or Instrumental Variables Methods could be used to correct for this. However, the assumptions they rely on (e.g. exclusion restrictions and non-linearities) are not credible in this context and it is easy to show that results are very sensitive to the particular assumptions made.