Article:
Hernández Alava, M. and Popli, G. (2013) Children's development and parental input: evidence from the UK Millennium Cohort Study. HEDS Discussion Paper 13/03. (Unpublished)

The University Of Sheffield.

School Of Health And Related Research.

Health Economics and Decision Science (HEDS)

Discussion Paper

Children's development and parental input: Evidence from the UK Millennium Cohort Study

Mónica Hernández Alava and Gurleen Popli

DP 13/03

Feb 2013

HEDS Discussion Paper

No. 13.03

Children's development and parental input: Evidence from the UK Millennium Cohort Study

Mónica Hernández Alava and Gurleen Popli

Children's development and parental input: evidence from the UK Millennium Cohort Study*

Mónica Hernández Alava[1], **Gurleen Popli**[2]

[1]School of Health and Related Research, University of Sheffield
[2]Department of Economics, University of Sheffield

February 2013

ABSTRACT: The experiences in the first few years of children are critical in shaping their future lives and ensuring that they achieve their full potential. Research has shown that both cognitive and non-cognitive abilities are important determinants of future socioeconomic outcomes, health and wellbeing. Furthermore, the activities that parents carry out with children at home (parental input or investment) have a significant effect in children's development. Using the Millennium Cohort Study we estimate a dynamic factor model of child development in the UK applying the framework of Cunha and Heckman (2008). Exploiting the wealth of information in the dataset, our model of development follows children from birth until seven years of age. We find a significant self-productivity effect in both cognitive and non-cognitive development. We also find evidence of dynamic dependence across different abilities; non-cognitive development increases cognitive development in the following period but cognitive development only appears to influence significantly non-cognitive development in the pre-school years. Parental investment is another significant influence in children's developmental trajectories. Furthermore, we find substantial evidence of two distinct parental investment latent variables which evolve over time.

KEYWORDS: Cognitive and non-cognitive development. Parental investment. Home learning environment. Dynamic factor model.

ADDRESS FOR CORRESPONDENCE: Mónica Hernández Alava. Health Economics and Decision Science. School of Health and Related Research. University of Sheffield. Regent Court. 30 Regent Street Sheffield S1 4DA. E-mail: monica.hernandez@sheffield.ac.uk. Tel: +44 (0)114 2220736. Fax: +44 (0)114 2724095

* We are grateful to The Centre for Longitudinal Studies, Institute of Education for the use of the Millenium Cohort Study data and to the UK Data Archive and Economic and Social Data Service for making them available. All responsibility for the analysis and interpretation of these data lies with the authors.

1 Introduction

The critical importance of the first few years of children's life in shaping not only their future achievements but also their future health and wellbeing is widely recognised. Research has shown that the levels of cognitive and non-cognitive development influence schooling decisions, employment and wages, teenage pregnancy, smoking, participation in illegal activities and incarceration (Feinstein 2003, Heckman et al. 2006, Blanden et al. 2007) and also drive in part health inequalities in adulthood (Conti and Heckman 2010)

Differences between children's cognitive and non-cognitive abilities form very early in life. By the time children reach school age, there are wide differences in their abilities and the evidence shows that children from disadvantaged backgrounds do worse in terms of cognitive and behavioural development (Shonkoff and Phillips 2000, NICHD 2005, Heckman 2006, Kiernan and Huerta 2008, Hobcraft and Kiernan 2010). These early gaps are highly persistent over time with disadvantaged children having lower life-coping abilities (see Neal and Johnson 1996, Feinstein 2000, Carneiro et al. 2005, Cunha et al. 2006 amongst others).

Parental input into child development is the focus of much recent research but there is a long history of related work in various fields. For example, in developmental psychology, epidemiology, sociology, child health and development (Mercy and Steelman 1982, Bergeman and Plomin 1988, Bradley et al. 1989, 2006, Kiernan and Huerta 2008, Lugo-Gil and Tamis-LeMonda 2008, Melhuish et al. 2008, Kelly et al. 2011, Byford et al. 2012) and more recently in economics (Cunha and Heckman 2008, Ermisch 2008, Currie 2009, Melhuish 2011, Aizer and Cunha 2012) the emphasis of research is on the importance of the parenting quality and the home environment defined in terms of the quality of stimulation and support available to the child. Parental investment in children has been found to have its greatest impact on cognitive development in the early years, but at a later stage for non-cognitive development (Cunha and Heckman 2008, Cunha et al. 2010). In the UK, a number of studies using the Millenium Cohort Study (see Ermisch 2008, Kiernan and Mensah 2009, 2011, Schoon et al. 2010, 2011, Kelly et al. 2011 amongst others) have shown that parental input plays a significant role in explaining child development.

Although research in child development across different disciplines has been largely independent, important common findings are emerging. Skill development is a dynamic process which builds on earlier skill levels. This process depends not only on genetics but also builds on experiences beginning in the prenatal period (Currie and Moretti 2007, Currie 2011). At the same time different skills are intertwined fostering higher levels of complementary skills in the future (Knudsen et al. 2006).

The highest priority recommendation of the Marmot Review (2010) was "to give every child the best start in life". The document setting out the Government's policy framework (DfE 2011) emphasizes the fundamental importance of children's development during the foundation years so that children are ready for school and can take full advantage and fulfill their potential. At the same time, the document stresses the influence of parenting behaviour and the impact of the home learning environment on children's development in those early years. It is at this early age that interventions to mitigate disadvantage is likely to be most effective (Shonkoff and Phillips 2000, Heckman 2006, Knudsen et al. 2006, Allen 2011). For this reason there are many policy initiatives aimed at reducing inequalities in the formative years. Despite recent efforts the "Two Years On" data recently released by the UCL Institute of Health Equity shows that, although there has been a small improvement in the proportion of children achieving a good level of development by the age of 5, this still leaves 40% who fall short, and the marked social gradient in child development still remains (UCL Institute of Health Equity 2012).

To be able to design effective policy interventions it is necessary to understand the complex dynamic interactions between the development of children's skills, both cognitive and behavioural and their home environment. In this paper using the framework of Cunha and Heckman (2008) we estimate a dynamic factor model of child development for the UK. We use longitudinal data from the Millennium Cohort Study, which has been following a cohort of children born in 2000-01. To our knowledge, this is the first comprehensive study which models ability formation in very young children in the UK. The dataset allows us to model cognitive and non-cognitive development from birth up to the age of 7. In addition, the present paper departs from the current literature in one important respect as we investigate parental input and its measures in more detail.

The structure of the paper is as follows. Section 2 outlines a dynamic model of child development. Section 3 discusses in detail the datataset and variables used in the analysis. Section 4 presents the results and section 5 concludes.

2 A dynamic model of child development

The framework of analysis uses the theoretical model of development presented in Cuhna and Heckman (2008) which is a generalization of the model by Todd and Wolpin (2007). The model formalises the growing body of evidence that cognitive and non-cognitive skills are interrelated and evolve jointly over time. The levels of cognitive and non-cognitive skills depend on the initial endowments a child is born with, the socioeconomic circumstances the child grows up in and the continuous influence of the parents through their investment. Parental investment is broadly defined as the set of activities carried out with the child and their more direct contributions to their education.

More formally, let $\boldsymbol{\theta}_t = \left(\theta_t^C, \theta_t^N\right)'$ represent the vector of latent developmental variables in period t; θ_t^C and θ_t^N are the stocks of latent cognitive and non-cognitive skills in period t respectively. Childhood can be divided in T time periods, $t = 0, 1, ...T-1$, not necessarily equivalent to a year. This developmental state evolves over time according to the following dynamic process:

$$\boldsymbol{\theta}_t = \mathbf{A}_t\boldsymbol{\theta}_{t-1} + \mathbf{B}_t\boldsymbol{\lambda}_{t-1} + \mathbf{C}_t\mathbf{X}_t + \boldsymbol{\eta}_t, \qquad t = 1, 2, \ldots, T-1 \tag{1}$$

where $\boldsymbol{\lambda}_t$ is a $(r \times 1)$ latent vector representing the parental investment at time t towards the development of the child. The model allows for the possibility that more than one latent factor $(r \neq 1)$ underlies parental investment. $\mathbf{A}_t$, $\mathbf{B}_t$ and $\mathbf{C}_t$ contain time-varying coefficients, $\mathbf{X}_t$ is a matrix of observed variables representing the child's current socioeconomic environment and $\boldsymbol{\eta}_t$ is a random error assumed to be independent across individuals and independent over time for the same individual with contemporaneous covariance matrix $\Sigma_{\eta\eta}^t$. The relationship in (1) allows for a cumulative effect of parental investment, with past investments built into the current developmental state and new investments influencing development into the next period. In addition, the latent vector of parental investment at time t is also a function of a matrix of observable variables $\mathbf{X}_t^I$

$$\boldsymbol{\lambda}_t = \mathbf{C}_t^I \mathbf{X}_t^I + \boldsymbol{\epsilon}_t^I, \qquad t = 1, 2, \ldots, T-2 \tag{2}$$

where $\mathbf{C}_t^I$ is a matrix of coefficients and $\boldsymbol{\epsilon}_t^I$ are random errors independent across individuals and over time and independent of $\boldsymbol{\eta}_t$.

We assume the following equations for the initial period of observation

$$\boldsymbol{\theta}_0 = \mathbf{C}_0 \mathbf{X}_0 + \boldsymbol{\xi} \tag{3}$$

$$\boldsymbol{\lambda}_0 = \mathbf{C}_0^I \mathbf{X}_0^I + \boldsymbol{\epsilon}_0^I \tag{4}$$

where the matrices $\mathbf{X}_0$ and $\mathbf{X}_0^I$ include variables representing family background as well as natal and prenatal circumstances and immediate post-natal factors such as breastfeeding. These two matrices do not necessarily contain the same set of variables. The random error $\boldsymbol{\xi}$ is assumed independent of $\boldsymbol{\eta}_t$, $\boldsymbol{\epsilon}_0^I$ and $\boldsymbol{\epsilon}_t^I$ with covariance matrix $\Sigma_{\xi\xi}$. All random errors are assumed to be normally distributed.

2.1 Model for the measurements

Both the vector of the developmental state, $\boldsymbol{\theta}_t$, and the vector of parental investment in the child, $\boldsymbol{\lambda}_t$, are latent variables and cannot be observed directly. However, the survey contains many observable outcome indicators or measures, denoted by $Y_{i,t}^k, k \in \{C, N, I\}, i = 1, \ldots, m_t^k$, which can be used as imperfect measures of the unobserved variables with m_t^k denoting the number of indicators of each latent variable at time t. These measures can be continuous or categorical.

Assuming that the outcome measures for cognitive and non-cognitive development are continuous[1] and the outcome measures used to identify parental investment are categorical, the measurement system can be written as:

$$Y_t^s = \mathbf{D}_t^s \theta_t^s + \mathbf{G}_t^s \mathbf{Z}_t^s + \boldsymbol{\varepsilon}_t^s \qquad s \in \{C, N\} \tag{5}$$

$$\left(Y_t^I\right)^* = \mathbf{D}_t^I \boldsymbol{\lambda}_t + \mathbf{G}_t^I \mathbf{Z}_t^I + \boldsymbol{\varepsilon}_t^I \tag{6}$$

$$y_{j,t}^I = r \text{ iff } \alpha_{r-1}^j \leq \left(y_{j,t}^I\right)^* \leq \alpha_r^j, \qquad j = 1, \ldots, m_t^I \text{ and } r = 1, \ldots, R_t^j \tag{7}$$

where $\alpha_0^j = -\infty$ and $\alpha_{R_t^j}^j = +\infty$, R_t^j is the number of response categories for the outcome measure at time t; $\mathbf{Z}_t^k$, $k \in \{C, N, I\}$, are matrices of variables specific to particular measures; $\mathbf{D}_t^k$ and $\mathbf{G}_t^k$ are matrices of coefficients and $\boldsymbol{\varepsilon}_t^k$ are normally distributed random errors of the measurement equations with a diagonal covariance matrix $\Sigma_{\varepsilon\varepsilon}^t$. Thus, all the correlation between the outcome measures at time t is due to the underlying effect of the latent variables and the covariates $\mathbf{Z}_t^k$. The matrices of coefficients $\mathbf{D}_t^k$ are known as the matrices of factor loadings and the relative size of these coefficients give an indication of the importance of the measure in the underlying latent variable.

The dynamic factor model of equations (1) to (7) is estimated using Mplus version 6.1 (Muthén and Muthén, 2010).

3 Data

The Millennium Cohort Study (MCS) is a longitudinal cohort survey that follows the families of around 19,000 children who were born in 2000-2001 in the UK. Its initial design oversampled families living in areas with high proportions of ethnic minorities in England, areas where child poverty was high and the three smaller countries of the UK (Wales, Scotland, and Northern Ireland). For a more thorough discussion of the MCS sampling design see Plewis (2007). Given the stratified design of the survey we use sampling weights in all the analyses reported in this paper.

[1] The outcome measures for both cognitive and non-cognitive development in the initial period are binary variables and treated as such but for simplicity of exposition we ignore it here.

The data are collected via both direct interview and by self-completion. In the majority of cases the main individual providing information about the household and child is the mother (or mother figure) but fathers/father figures are also interviewed. Four waves are available for 2001-2002, 2003-2004, 2006 and 2008 when the children were 9 months, 3 years, 5 years and 7 years old respectively. The MCS offers a wealth of socioeconomic variables and a guide to the datasets can be found in Hansen (2012).

Originally 20,646 families were contacted, of which 18,552 families responded (almost 90% response rate). The total number of cohort members in this wave was 18,818 children including 256 and 10 sets of twins and triplets respectively. In wave 2, some 'new-families' were added to the survey, however they cannot be included in our analysis as the information at 9 months is missing. We excluded from the analysis all twins and triplets (522), babies admitted to a special care unit immediately after they were born (1604) and babies with birth weight under 2500 grams (660). This leaves us with 16,032 children in our first wave sample. The final sample used in analysis contains to 9,089 children due to missing data on essential covariates and attrition (the fourth wave collected data on 13,857 families and 14,043 children). No additional observations are lost due to missing values in the measurement equations.

Details of the variables used as measures of cognitive development, non-cognitive development, parental investment and other covariates are given below.

3.1 Cognitive development

A wide range of measures of cognitive ability are used in the MCS dataset although not all them are used in each wave of the survey.

Cognitive development at age 9 months is measured using the Denver Developmental Screening Tests (DDST). DDST is an assessment widely used for examining the development of children from birth till age 6 years. The test was first devised in 1960 (Frankenburg et al. 1967) with major revisions in 1992 (Frankenburg et al. 1992). The test assesses children on 125 items grouped into four different areas. The MCS only uses a subset of the items covering three areas: *Fine motor function*: eye/hand co-ordination, and manipulation of small objects, e.g. grasping, passing, picking, putting hands together; *Gross motor functions*: motor control, e.g. sitting, walking, standing and other movements; *Communicative gestures*: production of sounds, ability to recognise, understand and use of language, e.g. smiles, gives toy, waves bye-bye, extends arms, nods for yes. A baby is classified as having a delay in an item if s/he is unable to perform a task 90% of babies in the same age group can. If a baby shows delays in more than one item in an area, s/he is classified as having a delay in that area. This classification is based on answers from the main respondent of the survey.

There is a range of standard tests of cognitive development in the other waves administered to the child by a trained interviewer: British Ability Scales Naming Vocabulary (BAS-NV); BAS Word Reading (BAS-WR); BAS Picture Similarity (BAS-PS); BAS Picture Construction (BAS-PC); Bracken School Readiness Assessment (BSRA); and Progress in Maths (PiM) test.

BAS-NV is a verbal scale which assesses spoken vocabulary. The children are shown a series of coloured pictures of objects one at a time which they are asked to name. The scale measures the children's expressive language ability. This test was administered to the children at ages 3 and 5 years. In BAS-PC, the child constructs a design by putting together flat squares or solid cubes with black and yellow patterns on each side. The child's score is based on both speed and accuracy in the task. This test was carried out at ages 5 and 7 years. The BAS-PS test assesses pictorial reasoning; the test was carried out only at age 3 years. Finally, in the BAS-WR the child reads aloud a series of words presented on a card. This test is an age appropriate version of BAS-NV, and was administered to the children when they were 7 years old. Further details on BAS tests can be found in Elliott et al. (1996, 1997).

BSRA is used to assess the conceptual development of young children across a wide range of categories in separate subtests. The MCS employs six of the subtests which specifically evaluate: colours, letters, numbers/counting, sizes, comparisons and shapes. The BSRA test result used is a composite score based on the total number of correct answers across all six subtests. This test was carried out when the children were 3 years old. For information on the BSRA see Bracken (2002).

Children's numerical and analytical skills at age 7 years were assessed using a variant of the

National Foundation for Educational Research (NFER) standard Progress in Maths (PiM) test in which children are examined on a range of tasks covering number, shape, space, measures and data.

Table A1 in the appendix shows sample summary statistics of these cognitive measures. The measures at 9 months are binary and show very low proportions of children with motor function delays. Delays in communicative gestures appear more common in the sample. The scores of the cognitive tests at all other ages except the PiM test have been standardised and are age adjusted in 3 months intervals. The sample means of these tests are all above 50 showing a slightly higher score than the norming groups used in the BAS tests to normalise the test scores. There are a larger number of missing values in the tests when the children are 3 years old.

3.2 Non-cognitive development

Non-cognitive or behavioural development at 9 months is measured using selected items from the Carey Infant Temperament scale (Carey 1972; Carey and McDevitt 1978) which capture the temperament of the children (reported by the mother) across three dimensions: mood, adaptability to new situations and regularity. Each dimension includes a range of questions and a total score for each of the three dimensions is obtained by adding the individual scores. High scores on the first two dimensions indicate distress and withdrawal while high scores on the last dimension indicate regularity. A child is classified as being 'difficult' if his/her score on a dimension is lower than the average score for the cohort.

The non-cognitive abilities of the children in the rest of the waves are assessed using the Strength and Difficulties Questionnaire (SDQ) (Goodman 1997). The SDQ is located in the self-completion questionnaire filled by the mother. It comprises 25 questions designed to capture the behavioural attributes of 3 to 16 year olds. The 25 questions are then grouped to assess children on five different scales: *Emotional Symptoms Scale*: complains of headaches/stomach aches/sickness, often seems worried, often unhappy, nervous or clingy in new situations, easily scared; *Conduct Problems Scale*: often has temper tantrums, generally obedient, fights with or bullies other children, can be spiteful to others, often argumentative with adults; *Hyperactivity Scale*: restless, overactive, cannot stay still for long, constantly fidgeting, easily distracted, cannot stop and think before acting, sees tasks through to the end; *Peer Problems Scale*: tends to play alone, has at least one good friend, generally liked by other children, picked on or bullied by other children, gets on better with adults. *Pro-social Scale*: considerate of others' feelings, shares readily with others, helpful if someone is hurt, upset or ill, kind to younger children, often volunteers to help others.

A higher number indicates worse behavioural problems for the first four scales, the reverse is true for the pro-social scale. The first four scales are often combined to obtain a total difficulties score for the child and it has been argued that the pro-social scale captures a different concept: "..the absence of prosocial behaviours is conceptually different from the presence of psychological difficulties." (Goodman 1997, page 582). For this reason the measure is excluded from the present analysis.

The bottom panel of Table A1 in the appendix presents summary statistics of the measures of non-cognitive development in our sample. The hyperactivity, emotional symptoms and peer problems scales are relatively stable over time with perhaps a small tendency for the mean to increase at the age of 7. In contrast, there is a clear movement towards less conduct problems as children get older with a significant drop between the ages of 3 and 5 (mean from 3.307 to 2.473) but it tends to stabilise between the ages of 5 and 7.

3.3 Parental investment

Parental investment at 9 months is measured using the mother's attitudes towards child rearing. Responses to four questions about the importance for development of talking to the baby, cuddling the baby, stimulating the baby and having a regular sleeping and eating time for the baby are used.

Responses of the mother to a wide range of questions are used at the ages of 3 and 5 to measure parental investment. Table A2 in the appendix shows the sample summary statistics and the coding of the different measures. The majority of the questions are identical in both waves but the wording of some of the questions changes in order to reflect the developmental stage. For example, when the children are 3, mothers are asked about the frequency their child is helped with the alphabet. At age 5 they are asked instead about the frequency their child is helped with reading and, separately, with writing.

The questions cover a wide range of activities parents may carry out with their children. These span activities which are directly related to preschool/school (for example " How often does someone at home help *[Cohort child's name]* with numbers, counting and adding up?") and other leisure activities indirectly related (such as "How often do you draw, paint or make things with *[Cohort child's name]*?"). Many of the questions are not specific about the person doing the activity with the child but some are. One such question is "How often do you read to *[Cohort child's name]*?" which is asked separately to both the mother and the father. This question gives us information about the degree of involvement of the father in these parenting activities.

There are also more general questions about everyday routines like regular bedtimes and hours spent watching television.

Table A2 in the appendix presents some summary statistics of these variables. The means of the answers to these questions are similar at 3 and 5 years suggesting that parenting styles are quite persistent over time. The only exception is the downward trend in the mean of the frequency of drawing/painting/making things which might reflect the children's increasing independence in choosing these type of activities as they grow up.

3.4 Covariates

A number of covariates are used in the analysis. To control for differences in the starting developmental position of children, equation (3), we use birthweight, the number of months the baby was breastfed, the age of the mother at birth and parental socioeconomic status. The initial level of parental investment, equation (4) is a function of parental socioeconomic status (NS-SEC 5 classes), level of education of the mother (a dummy for an NVQ level 5 or higher) and the household composition measured by the number of siblings in the family and the absence of a partner in the household. These same covariates are also used in the latent parental investment equations in each time period (2) with the addition of a dummy variable at the age of 5 to indicate a specially arranged meeting with the school teachers. In the dynamic developmental equation for cognitive development at age 5 which is part of the system of equations in (1) we control for the number of months the child had been attending school when the cognitive assessment was made.

Covariates in the measurement equations (5) and (6) capture systematic differences in the outcome measures for the same level of the latent variable.

Meltzer et al (2000) found that more behavioural problems are reported for boys and Goodman (1997) suggested using a different (higher) threshold for boys. The evidence on differences in reporting according to ethnicity is mixed. In a UK study Hackett and Hackett (1993) found that Gujarati parents have a more stringent concept than English parents of what constitutes 'acceptable behaviour'. Studies such as Miner et al. (2008), Atzaba-Poria et al. (2004) and Zwirs et al. (2006) found differences in reporting according to ethnicity, however, Goodman et al (2011) reported no significant differences. To allow for differences in reporting two dummy variables, one for male and one for white ethnic background are included in the measurement equations for non-cognitive ability.

The measurement equations for cognitive ability include the age of the child at the time the assessment was made in all time periods. Even though the cognitive measures used are defined relative to age groups, these age groups are defined in three months intervals and on a norming sample different to the MCS sample. Therefore, it is recommended for age to be included as a control variable (Hansen 2012). Other variables included in the cognitive measurement equations at 3, 5 and 7 years are a dummy for white ethnic background and a set of dummies identifying children who speak only English at home, English with some other language or other language but no English. In addition, when the children were 5 the interviewer recorded his/her perception of the child's fatigue at the time of the assessment. Two dummies, one for children who were 'a little tired' and another for children who were 'very tired' are used as covariates with the baseline category being 'wide awake'. No covariates are used in the measurement equations for parental investment in the final model.

Table A3 in the appendix presents summary statistics of all the covariates. The sample is split equally between boys and girls, the majority of the children are of a white background and only speak English at home. The average age of the mother at birth is 29 and it ranges from 13 to 48. In terms of socioeconomic status, the largest group is in managerial/professional occupations, followed by the baseline group, semi-routine/routine. The number of mothers with an NVQ 5 or higher remains unchanged in the first two waves with a small increase in the third wave.

Single parent households are a relatively small group and there is a tendency for this group to become more numerous with time. The number of siblings as expected also tends to increase with time. Perhaps surprisingly the parents of 16 percent of our sample had an specially arranged meeting with the teachers when the children were 5 years old. Unfortunately, the survey does not provide any information about the reasons for the meeting and these could be very wide ranging.

4 Results

This section discusses the estimation results. First, the parameter estimates for the different parts of the model are presented. Afterwards, we illustrate the dynamics implied by the model comparing the relative developmental positions of three children from the sample over time. The initial period of observation ($t = 0$ in the model in Section 2) corresponds to 9 months of age and includes data and variables relating to the time when the child was born. The next three periods ($t = 2, 3, 4$) correspond to ages 3, 5 and 7.

4.1 Parameter estimates

A model with contemporaneous correlations between the errors in the latent cognitive and non-cognitive development equations was estimated initially but all of the correlations were found to be highly insignificant. The restrictions of no contemporaneous correlations cannot be rejected with χ^2 test statistic of 2.747 with 4 degrees of freedom (p-value 0.6010). Therefore the model with no correlations is reported below.

Initial period latent equations

The estimated coefficients $\mathbf{C}_0$ and $\mathbf{C}_0^I$ of the initial period equations (3) and (4) are shown in Table 1. The findings are in line with previous studies. Birthweight is positively and significantly related to the levels of both cognitive and non-cognitive development in the initial period. The duration of breastfeeding is also positively related to both however it is not statistically significant at standard levels in the equation of cognitive development. The age of the mother at birth is also significant in both equations and exhibits a concave relationship; the maximum size effect is reached at around 28 years of age in the cognitive development equation and at around 37 in the non-cognitive development equation. The socioeconomic status of the household at the time of the birth shows also a significant relationship to the level of development of the child with increasingly higher coefficients found for higher socioeconomic groups. The only exception is the coefficient of the group of intermediate occupations in the cognitive development equation.

The signs of the coefficients of the parental investment equation are mainly as expected with higher mother's education and socioeconomic status of the household having a positive relationship to the level of parental investment and the number of siblings having a negative relationship. Perhaps surprisingly, in this initial period being a single parent household is associated with a higher level of parental investment. Nevertheless, only the coefficients of the highest two socioeconomic groups are statistically significant at standard levels.

Table 1: Estimated parameters of the initial latent equations (standard errors in brackets)

Covariate	Cognitive development - θ_0^C	Non-cognitive development - θ_0^{NC}	Parental investment - λ_0
Birthweight	0.169***	0.021**	
	(0.039)	(0.008)	
Breastfeeding (months)	0.024	0.014**	
	(0.016)	(0.006)	
Breastfeeding (months) squared	0.000	-0.001**	
	(0.002)	(0.001)	
Age of mother at birth/10	0.409*	0.300***	
	(0.239)	(0.102)	
Age of mother at birth/10 squared	-0.074*	-0.041***	
	(0.040)	(0.014)	
Parental socioeconomic status			
- managerial/professional	0.206***	0.097***	0.359***
	(0.055)	(0.031)	(0.060)
- intermediate	0.062	0.087***	0.256***
	(0.050)	(0.029)	(0.070)
- small employer/self employed	0.174***	0.031**	0.101
	(0.064)	(0.014)	(0.086)
- lower supervisors/technical	0.070	0.019	0.028
	(0.055)	(0.012)	(0.077)
Mother NVQ 5 or higher			0.028
			(0.112)
Number of siblings			-0.020
			(0.044)
Single parent household			0.104
			(0.175)

*** 1%, ** 5%, * 10% level of significance

Measurement equations

The measures available in the first wave are somewhat limited compared to the rest of the periods. The first column of Tables A4 and A5 and Table A6 in the appendix report the parameter estimates of the measurement equations at 9 months for the cognitive, non-cognitive and parental investment latent variables respectively. The signs are as expected. Higher levels of cognitive development lead to lower probabilities of gross and fine motor function and communicative gestures delays. Similarly, higher levels of non-cognitive development decrease the probabilities of showing low positive mood, distress to novelty and irregularity. In terms of parental investment, the probabilities of strongly agreeing with the importance of stimulating, talking and cuddling the baby and regular sleeping and feeding habits increase for higher levels of the underlying latent variable. The largest loading is on the importance of talking to the baby and the lowest is on the importance of regular sleeping and feeding times.

Columns 2 to 4 of Tables A4 and A5 in the appendix report the estimated coefficients of the measurement equations of cognitive and non-cognitive development variables respectively at ages 3, 5 and 7. The factor loadings are significant and have the expected signs. In addition, there are a number of points around the effects of the covariates worth mentioning. It is found that for the same level of development, children who speak a second language at home tend to score lower in cognitive tests. The effect, however, is only significant for one of the measures of cognitive development by the time children are 7 years of age. We also find that mothers tend to report higher levels of hyperactivity and conduct problems for boys and more emotional problems for girls (at 5 and 7 years) for the same level of latent non-cognitive development.

When the children are 3 and 5 years old, the survey contains many more usable measures of parental investment and we conduct an exploratory factor analysis to investigate the possibility that more than one latent variable underlies parental investment. Unlike most studies which make the assumption of only one latent parental investment, we find evidence of two separate latent variables underlying parental investment with changes in composition over time. Table A7 in the appendix presents a summary of the results. In both waves we find that all the statistics and indicators point towards two latent factors in parental investment[2].

Table 2 presents the parameter estimates $\mathbf{D}_t^I$ of the measurement equations (6) for the parental investment latent variables at ages 3 and 5 after splitting the measures according to the results of the factor analysis. Note that no additional covariates are included in the parental investment measurement equations. The first factor is a more general factor which includes activities that parents and children carry out together as well as TV watching and bedtime routines. The second factor is more focused on helping children with pre-school and school related matters. At the age of 3, there are no cross loadings across the factors, that is, only

[2] In both cases there are two eigenvalues above 1, the Root Mean Square Error of Approximation is significantly under 0.05, both CFI and TLI are above the usual rule of thumb of 0.95 and the Standardised Root Mean Square Residual is also at an acceptable value.

one of the factors appears in each of the measurement equations. However, by the age of 5, there are cross loadings in two measures; the frequency the child paints/draws at home and the frequency the child is helped with reading. Nevertheless, the second factor remains clearly related to school activities.

Table 2: Estimated parameters of the parental investment measurement equations.

	Age of child - 3 years		Age of child - 5 years	
Measure	$\lambda_1^{(1)}$	$\lambda_1^{(2)}$	$\lambda_2^{(1)}$	$\lambda_2^{(2)}$
Frequency mother reads to the child	1		1	
	(-)		(-)	
Frequency father reads to the child	0.544***		0.173***	
	(0.052)		(0.020)	
Frequency child taken to the library	0.392***		0.164***	
	(0.041)		(0.022)	
Frequency child paints/draws at home		1	0.217***	1
		(-)	(0.025)	(-)
Frequency child helped with alphabet		2.357***		
		(0.198)		
Frequency child helped with reading			0.535***	2.165***
			(0.067)	(0.157)
Frequency child helped with writing				4.258***
				(0.301)
Frequency child helped with counting/maths		4.206***		2.882***
		(0.633)		(0.170)
Frequency regular bedtime	0.409***		0.174***	
	(0.043)		(0.023)	
Frequency watching TV	-0.175***		-0.085***	
	(0.027)		(0.014)	

*** 1%, ** 5%, * 10% level of significance

It is of note that at the age of 3 years, the two largest coefficients (factor loadings) in the more general factor are in the equations for the measures of the frequency with which the mother and the father read to the child. By the age of 5, the frequency the mother reads to the child still has the largest factor loading but the coefficient in the measure of the frequency the father reads to the child is now much smaller in relative terms. At this age, the frequency a child is helped with reading becomes the second largest. When looking at the second parental latent variable we find that the largest factor loading at age 3 is in the measure of help with counting being overtaken by help with writing by the age of 5 following the patterns of learning over time.

Latent parental investment equations

Table 3 depicts the estimated coefficients $\mathbf{C}_t^I$ of the parental investment equations (2) at ages 3 and 5. The coefficients are in line with those of the initial period. Mother's with NVQ level 5 and higher tend to have a higher average parental investment in the child. The presence of other siblings and the absence of a partner in the household are associated with lower average parental investment. Households with high socioeconomic background have higher average parental investment in the child. It is of note however, that the socioeconomic status dummy variables are not significant in the factors related directly to school activities. In fact, none of these variables are significant in the equation for the latent parental investment of school activities at age 3 but by the age of 5, the time when children have started the first compulsory year at school, mothers education, the number of other siblings and having had a specially arranged meeting with the teachers become significant. Even though the parental socioeconomic variables are not significant, it is worth noting that the mean parental investment appears to go down as we move up the socioeconomic groups. This could be the result of the ability of parents of higher socioeconomic groups to choose better quality schools for their children which might in turn lead them a lower parental investment and an increased reliance on the school for academic achievements.

Table 3: Estimated parameters of the parental investment equations.
(standard errors in brackets)

	Age of child - 3 years		Age of child - 5 years	
	$\lambda_1^{(1)}$	$\lambda_1^{(2)}$	$\lambda_2^{(1)}$	$\lambda_2^{(2)}$
Mother NVQ 5 or higher	0.369***	0.019	0.264**	0.047*
	(0.097)	(0.029)	(0.109)	(0.027)
Number of siblings	-0.057	-0.014	-0.102**	-0.019**
	(0.047)	(0.015)	(0.048)	(0.010)
Single parent household	-0.074	-0.002	-0.307***	-0.013
	(0.080)	(0.022)	(0.085)	(0.019)
Parental socioeconomic status				
- managerial/professional	0.611***	0.013	0.450***	-0.022
	(0.066)	(0.017)	(0.083)	(0.017)
- intermediate	0.319***	0.002	0.230***	-0.014
	(0.062)	(0.020)	(0.079)	(0.018)
- small employer/self employed	0.286***	0.013	0.149	-0.006
	(0.074)	(0.023)	(0.095)	(0.020)
- lower supervisors/technical	0.151***	0.025	0.103	0.027
	(0.058)	(0.019)	(0.089)	(0.019)
Specially arranged meeting			0.025	0.027*
			(0.064)	(0.015)

*** 1%, ** 5%, * 10% level of significance

Dynamic latent variable equations

Table 4 presents the parameter estimates $\mathbf{A}_t$, $\mathbf{B}_t$ and $\mathbf{C}_t$ of the dynamic equation (1). There is a significant autoregressive effect in both cognitive and non-cognitive development; higher levels of cognitive (non-cognitive) development foster higher levels of development in future periods. There is also evidence of cross equation dependence. Cognitive development increases non-cognitive development in the next period but this effect is only statistically significant when the child is 3 years of age. The level of non-cognitive development has a significant effect on future levels of cognitive development at age 3, it turns insignificant at age 5 but becomes significant again at the age of 7 (p-value of 0.053).

Parental investment is also a significant determinant of future cognitive and non-cognitive development. The first and more general parental investment factor is significant throughout. The second parental investment factor related directly to helping children with school related chores is found to be significant for cognitive development when children start school but then by the time children are 7 it loses its significance. This might be capturing the effort parents make when children start school to ensure a good start. Once children are settled into a school which by and large has been chosen by the parents, they might delegate the role to the school. This is an important issue worth of further investigation which requires linking a rich set of school level data to the MCS.

It is of note that the number of months at school when children start compulsory education has a significant effect on the level of cognitive development.

Table 4: Parameter estimates of the developmental dynamic equations

	Age of child - 3 years		Age of child - 5 years		Age of child - 7 years	
Covariates	θ_1^C	θ_1^N	θ_2^C	θ_2^N	θ_3^C	θ_3^N
θ_{t-1}^C	0.802***	0.142***	0.686***	0.001	0.901***	0.005
	(0.168)	(0.052)	(0.033)	(0.005)	(0.042)	(0.006)
θ_{t-1}^N	5.763***	3.302***	0.043	0.470***	0.101*	0.988***
	(1.977)	(1.109)	(0.038)	(0.019)	(0.052)	(0.033)
$\lambda_{t-1}^{(1)}$	0.073***	0.020	0.444***	0.063***	0.136***	0.020***
	(0.026)	(0.013)	(0.049)	(0.011)	(0.025)	(0.006)
$\lambda_{t-1}^{(2)}$			0.595***	-0.028	0.049	0.020
			(0.103)	(0.026)	(0.076)	(0.023)
Months at school			0.030***			
			(0.010)			

*** 1%, ** 5%, * 10% level of significance

Table A8 in the appendix shows the estimated correlation matrix of the latent variables in the model. These correlations amongst the latent variables are due to the dynamics as well as the common influences of the covariates. The correlations between the initial level of cognitive (non-cognitive) development and the levels of development at other time periods are always positive which reflects the raw correlations we see in the data in which children with a high level of cognitive (non-cognitive) development are more likely to show higher levels of cognitive (non-cognitive) development as they grow up. These correlations with the initial levels tend to decrease slowly over time; the correlations of the initial latent cognitive (non-cognitive) development with the latent cognitive (non-cognitive) development at 3, 5 and 7 years are 0.333, 0.261 and 0.231 (0.453, 0.388 and 0.355) respectively. However, the correlations between latent development in subsequent periods increases substantially; the estimated correlation between cognitive (non-cognitive) development between 3 and 5 is 0.745 (0.804) and increases to 0.865 (0.960) between the ages of 5 and 7. These numbers give an indication of the complex nature of development and show that the starting developmental position has an important influence on the developmental path but other continuing influences such as the parenting activities become very important influences as time goes by. This is in line with findings in the literature and has been termed dynamic complementarity (see for example Cuhna and Heckman 2007)

4.2 Factor scores

The model estimated in the last section shows that the developmental position of a child in each period is a function of the past developmental position and is subject to influences coming directly from the socioeconomic environment as well as parental influences. It is difficult to interpret directly the magnitude of the coefficients and their effect on the level of development since the latent developmental variables that we are interested in are not observed. However, we can assign values (factor scores) to these latent variables based on the estimated model and use them as a simple illustration of their evolution. We use the expected a posteriori method to calculate the factor scores with 10,000 Monte Carlo simulations. Table 5 shows the developmental positions of three children in our sample by calculating the proportion of children with an estimated factor score below that attributed to the child. All three children have the same level of cognitive development in the first survey and are at the very bottom of the distribution. In terms of non-cognitive development there are already substantial differences in the position of the children in the distribution of non-cognitive development; the first child is at the bottom of the distribution, the second one is towards the top of the distribution and the third one is in between. The background socioeconomic characteristics of these children (Table 6) and the implied levels of parental investment (Table 5) are very different. Child (a) is in a less favourable position than the other two children; born in a lower socioeconomic household to a younger mother who already had 4 other children and will have one more baby before the child is 3 years old, the mother's education is below NVQ level 5, the child was not breastfed as a baby and by the time the child reaches 7 years of age the father is no longer present in the household. In addition, the implied levels of parental investment are also lower, with large differences observed when compared to Child (b). By the time the children are 7 years old there exist large differences in their developmental positions. Child (a) has the lowest level of both cognitive and non-cognitive development out of the three children. A sizeable gap has opened between Child (b) and the other two in both cognitive and non-cognitive development and for these three children in particular, these large gaps are already present by the time the children are 3 years old. This example highlights the richness and complexities in the dynamics of the interactions between cognitive and non-cognitive development and parental investment and other background characteristics.

Table 5: Proportion of children in the survey with a smaller factor score

		Child (a)	Child (b)	Child (c)
Age of child	θ_0^C	0.02	0.02	0.02
9 months	θ_0^N	0.07	0.88	0.40
	λ_0	0.23	0.29	0.78
Age of child	θ_1^C	0.02	0.67	0.17
3 years	θ_1^N	0.04	0.70	0.13
	$\lambda_1^{(1)}$	0.01	0.64	0.23
	$\lambda_1^{(2)}$	0.00	0.52	0.13
Age of child	θ_2^C	0.01	0.74	0.13
5 years	θ_2^N	0.02	0.82	0.09
	$\lambda_2^{(1)}$	0.01	0.58	0.08
	$\lambda_2^{(2)}$	0.31	0.86	0.09
Age of child	θ_3^C	0.01	0.73	0.18
7 years	θ_3^N	0.02	0.85	0.09

Table 6: Background socioeconomic characteristics of selected children in the sample

	Child (a)	Child (b)	Child (c)
Birthweight (Kg)	2.81	3.63	2.89
Gender	girl	boy	boy
Ethnicity	white	white	white
Child's age - 1st wave (months)	9	9	9
Child's age - 4st wave (months)	86	83	84
Breatfeeding duration (months)	0	0	1
Mother's age at birth	28	33	42
Mother's education (all waves)	NVQ 4 or lower	NVQ 4 or lower	NVQ 5 or higher
Socioeconomic group	amall employer/self-employed	managerial/professional	managerial/professional
Language spoken at home	english	english	english
Household composition (1st wave)	Two parents	Two parents	Two parents
Household composition (2nd wave)	Two parents	Two parents	Two parents
Household composition (3rd wave)	Single parent	Two parents	Two parents
Number of siblings (1st wave)	4	2	0
Number of siblings (2nd wave)	5	3	0
Number of siblings (3rd wave)	5	3	0
Specially arranged meeting with teachers (3rd wave)	Yes	No	No

5 Discussion

It is widely recognised that the first few years of a child's life are crucial in shaping their future. Recent research has shown that the levels of cognitive and non-cognitive development influence medium term achievements and decisions such as schooling and also long term outcomes such as employment and wages, smoking and participation in illegal activities. Research has also shown that gaps in cognitive and non-cognitive abilities form very early in life and are highly persistent over time. Children from disadvantaged backgrounds have on average lower levels of cognitive and non-cognitive development. For this reason the document setting out the Government's policy framework (DfE 2011) places great emphasis on the development of children during the foundation years and the number of policies and interventions aimed at reducing inequalities in the formative years has risen. However, despite all the efforts, the "Two Years On" data released in 2012 by the UCL Institute of Health Equity showed that 40% of children in the UK do not achieve a good level of development at the age of 5.

Children's development is a complex dynamic process influenced by the socioeconomic environment children grow up in, parenting behaviour and what has been termed the home learning environment. It is necessary to better understand the complexity of these dynamic relationships if we are to design effective policy interventions to reduce inequalities in cognitive and non-cognitive development before large gaps form.

The present paper uses longitudinal data from the Millennium Cohort Study to estimate a dynamic factor model of child development in the UK using the framework of Cuhna and Heckman (2008). The data covers the early years of a child's life starting from birth up to the age of 7, the period where interventions to alleviate disadvantages are likely to have the largest effect and is the first comprehensive study of development in these early years in the UK. In line with similar research using data from other countries or data from different developmental ages, we find significant evidence of a self-productive effect in both cognitive and non-cognitive development; higher levels of cognitive (non-cognitive) development today foster higher levels of cognitive (non-cognitive) levels in the future. We also find evidence of cross-dependence between different abilities. In particular, non-cognitive development increases cognitive development in the following period, however cognitive development only appears to influence significantly non-cognitive development in the pre-school years. We also find that parental investment is another significant influence in children's developmental trajectories similar to other research in the area. However, the present paper departs from the current literature in one important respect as we investigate the possibility of more than one factor behind our measures of parental input. We find evidence of two different factors which evolve over time. One factor is more general and covers a range of activities that parents carry out with their children as well as usual routines and practices. This factor has a significant effect in both cognitive and non-cognitive development throughout all these early years of development. The second parental investment factor is related to helping children with school matters. This latent parental input affects only cognitive development at the age of 5, the first year of compulsory education in the UK. One possible reason for this pattern could be that once parents have made a decision about the choice of school for their children, they leave the school to push children forward in academic matters. The other possible reason is that the timing of this latent variable is different to the first with an almost simultaneous effect rather than a lagged one as postulated in our model.

As a final point in the paper we illustrate the complex dynamics and interactions between the developmental variables, the parental inputs and other socioeconomic variables by showing the relative developmental positions of three children in the sample over time. These three children had the same level of cognitive development at 9 months but different levels of non-cognitive development. By the age of 7 a large gap has opened up in the level of development (both cognitive and non-cognitive) between one of the children and the other two. The large gap is partly explained by the child having a significantly higher level of non-cognitive development at 9 months and the much higher levels of parental investment at the ages of 3 and 5.

Given the importance of parental investment in the children's developmental trajectories throughout the first few years, efforts should concentrate on designing policies to help parents improve the home learning environment. However, these policies will not be successful if they only have a temporary effect in parental investment as other influences such as other socioeconomic circumstances might eventually outweigh the effect. Policies designed to increase parental investment at different stages of a child's life complemented by policies to tackle the source of initial inequalities will have a much higher likelihood of reducing the gaps in cognitive and non-cognitive development.

References

[1] Allen, G. M. (2011). Early intervention: the next steps. An Independent Report to Her Majesty's Government. London: HM Government.

[2] Atzaba-Poria, N., Pike, A., and Barrett, M. (2004) Internalising and externalising problems in middle childhood: A study of Indian (ethnic minority) and English (ethnic majority) children living in Britain. International Journal of Behavioral Development, 28, 449-460, doi:10.1080/01650250444000171

[3] Becker, G. and Lewis, H. G. (1973). On the Interaction between the Quantity and Quality of Children, Journal of Political Economy 81, S279–S288.

[4] Bergeman, C. S., and Plomin, R. (1988). Parental mediators of the genetic relationship between home environment and infant mental development. British Journal of Developmental Psychology, 6, 11-19.

[5] Blanden, J., Gregg, P. and Macmillan, L. (2007) Accounting for intergenerational income persistence: non-cognitive skills, ability and education, Economic Journal, 117, C43-C60.

[6] Bracken, B.A. (2002) Bracken School Readiness Assessment: Administration Manual, San Antonio, Texas: Psychological Corporation.

[7] Bradley, R. H., & Corwyn, R. F. (2006). The family environment. In L. Balter & C. S. Tamis-LeMonda (Eds.), Child psychology: A handbook of contemporary issues (2nd ed., pp. 493–520). New York, NY: Psychology Press.

[8] Bradley, R. H., Caldwell, B. M., Rock, S. L., Barnard, K. E., Gray, C., Hammond, M. A., Mitchell, S., Siegel, L., Ramey, C. T., Gottfried, A. W. & Johnson, D. L. (1989). Home environment and cognitive development in the first 3 years of life: a collaborative study involving six sites and three ethnic groups in North America. Developmental Psychology. 25, 217–235.

[9] Byford, M., Kun, D., and Richards, M. (2012). Parenting practices and intergenerational associations in cognitive ability. International Journal of Epidemiology 41(1): 263-272.

[10] Carey, W.B. (1972) Measuring infant temperament. The Journal of Pediatrics 81: 414.

[11] Carey, W.B. and McDevitt, S.C. (1978) Revision of the Infant Temperament Questionnaire. Pediatrics 61: 735–9.

[12] Carneiro, P., Heckman, J.J. and Masterov, D.V. (2005). Labor Market Discrimination and Racial Differences in Premarket Factors. Journal of Law and Economics 48(1): 1-39.

[13] Cunha, F. and Heckman, J.J. (2007). The technology of Skill Formation. American Economic Review, 97(2):31-47.

[14] Cunha, F., Heckman, J.J., Lochner, L. and Masterov, D. (2006). Interpreting the Evidence on Life Cycle Skill Formation in Handbook of the Economics of Education, edited by E. Hanushek and F. Welch, 697-812, North Holland: Amsterdam.

[15] Cunha, F. and Heckman, J.J. (2008). Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation. Journal of Human Resources 43(4): 738-782.

[16] Cunha F., Heckman J.J. and Schennach S.M. (2010). Estimating the technology of cognitive and noncognitive skill formation. Econometrica 78(3): 883–931.

[17] Currie, J. 2011. Inequality at Birth: Some Causes and Consequences. American Economic Review 101(3): 1-22.

[18] Department for Education, Department of Health (2011) Supporting Families in the Foundation Years. London: Department for Education, Department of Health.

[19] Elliott C D, Smith P and McCulloch K (1996) British Ability Scales Second Edition (BAS II): Administration and Scoring Manual. London: NFER-Nelson.

[20] Elliott C D, Smith P and McCulloch K (1997) British Ability Scales Second Edition (BAS II): Technical Manual. London: NFER-Nelson.

[21] Ermisch, J. (2008) Origins of social immobility and inequality: parenting and early child development. National Institute Economic Review, 205: 62 – 71

[22] Feinstein, L. (2000) The relative economic importance of academic, psychological, and behavioural attributes developed in childhood. CEP Discussion Paper.

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

[24] Frankenburg, W.K. and Dodds, J. (1967). The Denver Developmental Screening Test. The Journal of Pediatrics 71 (2): 181–191.

[25] Frankenburg, W.K., Dodds, J., Archer, P., Shapiro, H., and Bresnick, B. (1992). The Denver II: a major revision and restandardization of the Denver Developmental Screening Test. Pediatrics 89 (1):91-7.

[26] Frankenburg, W. K. (2002) Developmental Surveillance and Screening of Infants and Young Children. Pediatrics 109 (1): 144-145.

[27] Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A Research Note. Journal of Child Psychology and Psychiatry 38:581-586.

[28] Goodman A, Goodman R (2011) Population mean scores predict child mental disorder rates: validating SDQ prevalence estimators in Britain. Journal of Child Psychology and Psychiatry, 52, 100-8.

[29] Grossman, M. (1972). On the Concept of Health Capital and the Demand for Health, Journal of Political Economy 80, 223–255.

[30] Hackett, L. and Hackett, R. (1993) Parental ideas of normal and deviant child behaviour. A comparison of two ethnic groups. The British Journal of Psychiatry, 162: 353-357, doi: 10.1192/bjp.162.3.353

[31] Hansen, K. (ed.) (2012) Millennium Cohort Study First, Second, Third and Fourth Surveys: A Guide to the Datasets (Sixth Edition), Centre for Longitudinal Studies, Institute of Education, London, July.

[32] Heckman, J.J. 2006 Skill Formation and the Economics of Investing in Disadvantaged Children, Science 312: 1900–12.

[33] Heckman, J.J., Stixrud, J. and Urzua, S. (2006) The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. Journal of Labor Economics 24(3): 411-482.

[34] Hobcraft, J.N. and 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

[35] Kelly,Y., Sacker,A., Del Bono, E., Francesconi, M. and Marmot, M. (2011)What role for the home learning environment and parenting in reducing the socio-economic gradient in child development? Findings from the Millennium Cohort Study, Archives of Disease in Childhood, 96: 832-837

[36] Kiernan, K.E. and Huerta, M.C. (2008). Economic deprivation, maternal depression, parenting and children's cognitive and emotional development in early childhood, British Journal of Sociology 59(4): 783-806.

[37] Kiernan, K.E and Mensah, F.K. (2009) Poverty, maternal depression, family status and children's cognitive and behavioural development in early childhood: A longitudinal study, Journal of Social Policy, 38: 569-588.

[38] Kiernan, K.E. and Mensah, F.K. (2011) Poverty, family resources and children's early educational attainment: the mediating role of parenting, British Educational Research Journal, 37(2): 317-336.

[39] Knudsen, E. I., J. J. Heckman, J. Cameron, and J. P. Shonkoff (2006, July). Economic, neurobiological, and behavioral perspectives on building America's future workforce. Pro-ceedings of the National Academy of Sciences 103 (27), 10155-10162.

[40] Lugo-Gil, J. and Tamis-LeMonda, C. S. (2008), Family Resources and Parenting Quality: Links to Children's Cognitive Development Across the First 3 Years. Child Development, 79: 1065–1085. doi: 10.1111/j.1467-8624.2008.01176.x

[41] Marmot, M. (2010). Fair society, healthy lives: The Marmot Review. Strategic review of health inequalities in England post-2010. London: University College London.

[42] Marmot Review Team (2011). Interim Second Report on Social Determinants of Health and the Health Divide in the WHO European Region. Copenhagen, Denmark: WHO Euro.

[43] Meltzer, H., Gatward, R., Goodman, R., and Ford, F. (2000) Mental health of children and adolescents in Great Britain. London: The Stationery Office

[44] Miner, Jennifer L.; Clarke-Stewart, K. Alison (2008) Trajectories of externalizing behavior from age 2 to age 9: Relations with gender, temperament, ethnicity, parenting, and rater. Developmental Psychology, Vol 44(3), May 2008, 771-786.

[45] Muthén, L.K. and Muthén, B.O. (2010). Mplus User's Guide. Sixth Edition. Los Angeles, CA: Muthén & Muthén.

[46] National Institute of Child Health and Human Development Early Child Care Research Network 2005 Duration and Developmental Timing of Poverty and Children's Cognitive and Social Development From Birth Through Third Grade, Child Development 76(4): 795–810.

[47] Neal, D. and Johnson, W. (1996). The Role of Premarket Factors in Black-White Wage Differences. Journal of Political Economy, 104: 869-895.

[48] Plewis, I. (ed.) (2007) The Millennium Cohort Study: Technical Report on Sampling (Fourth Edition), Centre for Longitudinal Studies, Institute of Education, London, July.

[49] Shonkoff, J.P. and Phillips, D.A. 2000 From Neurons to Neighborhoods: The Science of Early Child Development, Washington, D.C.: National Academy Press.

[50] Schoon, I., Cheng, H. and Jones, E. (2010) Resilience in children's development in Kirstine Hansen, Heather Joshi and Shirley Dex (eds) Children of the 21st century: The first five years. Bristol: Policy Press.

[51] Schoon, I., Jones, E., Cheng, E. and Maughan, B. (2011) Family hardship, family instability, and cognitive development Journal of Epidemiology and Community Health, Online First, April 2011.

[52] Todd, P.E. and Wolpin, K.I. (2007) The Production of Cognitive Achievement in Children: Home, School, and Racial Test Score Gaps. Journal of Human Capital 1(1): 91-136

[53] UCL Institute of Health Equity (2012) 'Two Years On' Data press release 15 February 2012 (http://www.instituteofhealthequity.org/media/press-releases/two-years-on-data)

[54] Zwirs, Barbara W. C.; Burger, Huibert; Buitelaar, Jan K.; et al. (2006) Ethnic differences in parental detection of externalizing disorders. European Child and Adolescent Psychiatry,15: 7, 418-426 DOI: 10.1007/s00787-006-0550-7

APPENDIX

Table A1: Unweighted sample descriptive statistics of cognitive and non-cognitive measurement variables

	Age of child											
	9 months			3 years			5 years			7 years		
	Mean	Std dev	n	Mean	Std dev	n	Mean	Std dev	n	Mean	Std dev	n
Cognitive measures												
Gross motor function delay(1)	0.106	0.308	9087									
Fine motor function delay(1)	0.125	0.329	9054									
Communicative gestures delay(1)	0.388	0.487	9077									
BSR composite standard score				53.642	28.336	8313						
BAS Naming Vocabulary				53.401	28.162	8693	53.033	28.218	9068			
BAS Picture Similarity							51.814	28.442	9063			
BAS Pattern Construction							52.948	28.464	9042	52.391	28.432	8963
BAS Word Reading										51.909	28.283	8876
Numerical&Analytical Skills										52.199	28.220	8989
Non-cognitive measures												
Low positive mood(1)	0.488	0.500	8330									
Distress to novelty(1)	0.445	0.497	6015									
Irregularity(1)	0.431	0.495	8644									
Hyperactivity Scale(2)				4.086	1.698	8745	4.004	1.495	8914	4.155	1.495	8910
Emotional Symptoms Scale(2)				1.263	1.388	8810	1.285	1.500	8949	1.440	1.690	8913
Conduct Problems Scale(2)				3.307	1.612	8823	2.473	1.063	8955	2.462	1.052	8928
Peer Problems Scale(2)				4.345	1.325	8757	4.645	1.127	8938	4.682	1.138	8912

(1)Dummy variable

(2)Higher values indicate worse behavioural problems

Table A2: Unweighted sample descriptive statistics of parental investment measurement variables.

	Age of child								
	9 months			3 years			5 years		
Measure	Mean	Std dev	n	Mean	Std dev	n	Mean	Std dev	n
Importance of stimulating baby[1]	1.386	0.601	8904						
Importance of talking to baby[1]	1.176	0.427	8906						
importance of cuddling baby[1]	1.172	0.438	8906						
Importance of regular sleeping/eating for baby[1]	1.588	0.742	8904						
Frequency mother reads to the child[2]				3.324	1.002	9089	3.265	0.920	9089
Frequency father reads to the child[2]				2.091	1.370	7874	2.010	1.312	8303
Frequency child taken to the library[3]				0.495	0.747	9089	0.441	0.652	9089
Frequency child paints/draws at home[2]				3.174	0.918	9089	1.990	0.994	9089
Frequency child helped with alphabet[2]				2.138	1.384	9089			
Frequency child helped with reading[2]							3.423	0.850	9089
Frequency child helped with writing[2]							2.729	1.173	9088
Frequency child helped with counting/maths[2]				3.165	1.045	9089	2.848	1.116	9088
Frequency regular bedtime[4]				2.151	0.896	9089	2.490	0.778	9089
Frequency watching TV[5]				1.928	0.649	9089	1.917	0.631	9087

[1] Coded as (1) strongly agree, (2) agree, (3) neither agree nor disagree, (4) disagree, (5) strongly disagree

[2] Coded as (0) not at all, (1) once/twice/less a month, (2) once/twice a week, (3) several times a week, (4) every day

[3] Coded as (0) never/special occasion, (1) at least once a month, (2) once a week or more

[4] Coded as (0) never or almost never, (1) sometimes, (2) usually, (3) always

[5] Coded as (0) none, (1) up to one hour, (2) between 1 and 3 hours, (3) more than 3 hours

Table A3: Unweighted descriptive statistics of sample covariates

	Mean	Std dev
Birthweight (Kg)	3.47	0.46
Male	0.50	0.50
White	0.88	0.33
Child's age - 1st wave (months)	9.19	0.50
Child's age - 2nd wave (months)	37.54	2.33
Child's age - 3rd wave (months)	62.64	2.91
Child's age - 4th wave (months)	86.78	2.97
Mother's age at birth/10	29.23	5.61
Parental socioeconomic status		
- managerial/professional	0.48	0.5
- intermediate	0.14	0.35
- small employer/self employed	0.07	0.25
- lower supervisors/technical	0.09	0.29
- semi-routine and routine (baseline)		
Mother NVQ 5 or higher - 1st wave	0.36	0.48
Mother NVQ 5 or higher - 2nd wave	0.36	0.48
Mother NVQ 5 or higher - 3rd wave	0.39	0.49
Single parent household - 1st wave	0.11	0.31
Single parent household - 2nd wave	0.13	0.34
Single parent household - 3rd wave	0.15	0.36
Length of breastfeeding (months)	2.88	3.42
Number of siblings - 1st wave	0.92	1.01
Number of siblings - 2nd wave	1.20	1.03
Number of siblings - 3rd wave	1.38	1.02
Number of months at school- 3rd wave	6.98	2.97
Specially arranged meeting with teachers - 3rd wave	0.16	0.36
English and other language spoken at home	0.08	0.28
No English spoken at home - other language	0.03	0.16
Child a little tired during the assessment according to interviewer	0.21	0.41
Child very tired during the assessment according to interviewer	0.01	0.11

Table A4: Parameter estimates of the measurement equations for the cognitive latent variable

	Covariates (latent variables)			
Measure	θ_0^C	θ_1^C	θ_2^C	θ_3^C
Gross motor function delay[(1)]	-1			
	(-)			
Fine motor function delay[(1)]	-1.029***			
	(0.214)			
Communicative gestures delay[(1)]	-0.752***			
	(0.152)			
BSR composite standard score		1.257***		
		(0.047)		
BAS Naming Vocabulary		1	1	
		(-)	(-)	
BAS Picture Similarity			0.782***	
			(0.035)	
BAS Pattern Construction			0.959***	0.954***
			(0.044)	(0.045)
BAS Word Reading				1
				(-)
Numerical&Analytical Skills				1.177***
				(0.045)

	Covariates (observed variables)					
	Child's age	White	English and other	No English	a little tired	Very tired
Gross motor function delay[(1)]	5.061***					
	(0.498)					
Fine motor function delay[(1)]	-0.829					
	(0.514)					
Communicative gestures delay[(1)]	-3.364***					
	(0.349)					
BSR composite standard score	0.912***	0.434***	-0.601***	-1.610***		
	(0.130)	(0.140)	(0.157)	(0.228)		
BAS Naming Vocabulary	0.687***	0.736***	-1.442***	-2.677***		
(age 3)	(0.155)	(0.125)	(0.142)	(0.292)		
BAS Naming Vocabulary	-0.073	0.557***	-1.587***	-2.336***	-0.260***	-0.501*
(age 5)	(0.210)	(0.122)	(0.159)	(0.260)	(0.080)	(0.272)
BAS Picture Similarity	0.126	0.115	-0.110	-0.469**	-0.314***	-0.538*
	(0.250)	(0.134)	(0.153)	(0.235)	(0.087)	(0.304)
BAS Pattern Construction	-1.534***	0.215	-0.541***	-0.613**	-0.595***	-1.801***
(age 5)	(0.251)	(0.146)	(0.166)	(0.258)	(0.077)	(0.356)
BAS Pattern Construction	0.241	0.800***	-0.101	-0.118		
(age 7)	(0.208)	(0.146)	(0.162)	(0.267)		
BAS Word Reading	0.053	-0.477***	0.192	0.196		
	(0.216)	(0.155)	(0.159)	(0.246)		
Numerical&Analytical Skills	1.821***	-0.021	-0.413**	-0.527**		
	(0.243)	(0.151)	(0.167)	(0.268)		

[(1)]Dummy variable

*** 1%, ** 5%, * 10% level of significance

Table A5: Parameter estimates of the measurement equations for the non-cognitive latent variable

	Covariates (latent variables)			
Measure	θ_0^{NC}	θ_1^{NC}	θ_2^{NC}	θ_3^{NC}
Low positive mood[(1)]	-1			
	(-)			
Distress to novelty[(1)]	-2.693***			
	(0.924)			
Irregularity[(1)]	-3.103***			
	(0.987)			
Hyperactivity Scale[(2)]		-0.959***	-1.302***	-1.268***
		(0.040)	(0.048)	(0.056)
Emotion Symptoms Scale[(2)]		-0.943***	-1.841***	-2.125***
		(0.035)	(0.059)	(0.065)
Conduct Problems[(2)]		-1	-1	-1
		(-)	(-)	(-)
Peer Problems[(2)]		-0.508***	-0.910***	-0.942***
		(0.025)	(0.037)	(0.036)

	Covariates (observed variables)	
	White	Male
Hyperactivity Scale[(2)]	-0.071	0.187***
(age 3)	(0.083)	(0.044)
Emotion Symptoms Scale[(2)]	-0.178***	0.002
(age 3)	(0.059)	(0.033)
Conduct Problems[(2)]	-0.063	0.089**
(age 3)	(0.089)	(0.038)
Peer Problems[(2)]	-0.248***	0.002
(age 3)	(0.052)	(0.034)
Hyperactivity Scale[(2)]	-0.189***	0.092***
(age 5)	(0.071)	(0.034)
Emotion Symptoms Scale[(2)]	-0.187***	-0.110***
(age 5)	(0.071)	(0.037)
Conduct Problems[(2)]	0.009	0.076***
(age 5)	(0.860)	(0.024)
Peer Problems[(2)]	-0.131**	0.037
(age 5)	(0.054)	(0.027)
Hyperactivity Scale[(2)]	-0.158**	0.200***
(age 7)	(0.070)	(0.037)
Emotion Symptoms Scale[(2)]	-0.070	-0.147**
(age 7)	(0.084)	(0.040)
Conduct Problems[(2)]	0.058	0.082***
(age 7)	(0.050)	(0.025)
Peer Problems[(2)]	-0.155***	0.048*
(age 7)	(0.048)	(0.027)

[(1)]Dummy variable

[(2)]Higher values indicate worse behavioural problems

*** 1%, ** 5%, * 10% level of significance

Table A6: Parameter estimates of the measurement equations of the parental investment latent variable at 9 months.

Measure	λ_0
Importance of stimulating the baby[(1)]	-1
	(-)
Importance of talking to the baby[(1)]	-2.926***
	(0.424)
Importance of cuddling the baby[(1)]	-1.043***
	(0.040)
Importance of regular sleeping/feeding times[(1)]	-0.422***
	(0.021)

[(1)]Coded as discrete values ranging from 0 = strongly agree to 4 = strongly disagree
*** 1%, ** 5%, * 10% level of significance

Table A7: Summary results of the Exploratory Factor Analysis for parental investment

	Children's age 3 years			
Eigenvalues	2.106	1.484	0.924	
	RMSEA [p<0.05]	CFI	TLI	SRMR
One factor	0.084[0.000]	0.646	0.504	0.090
Two factors	0.024[1.000]	0.981	0.960	0.019
Three factors	0.003[1.000]	1.000	1.000	0.005
	Children's age 5 years			
Eigenvalues	2.617	1.266	0.981	
	RMSEA [p<0.05]	CFI	TLI	SRMR
One factor	0.067[0.000]	0.886	0.847	0.064
Two factors	0.034[1.000]	0.980	0.961	0.027
Three factors	0.019[1.000]	0.996	0.987	0.014

Note: RMSEA: Root Mean Square Error of Approximation
SRMR: Standardised Root Mean Square Residual

Table A8: Correlation matrix of latent variables

	θ_0^C	θ_0^N	λ_0	θ_1^C	θ_1^N	$\lambda_1^{(1)}$	$\lambda_1^{(2)}$	θ_2^C	θ_2^N	$\lambda_2^{(1)}$	$\lambda_2^{(2)}$	θ_3^C	θ_3^N
θ_0^C	1												
θ_0^N	0.141	1											
λ_0	0.022	0.065	1										
θ_1^C	0.333	0.453	0.089	1									
θ_1^N	0.166	0.493	0.065	0.252	1								
$\lambda_1^{(1)}$	0.071	0.188	0.043	0.099	0.098	1							
$\lambda_1^{(2)}$	0.006	0.012	0.004	0.007	0.006	0.016	1						
θ_2^C	0.261	0.388	0.077	0.745	0.230	0.364	0.133	1					
θ_2^N	0.142	0.417	0.057	0.216	0.804	0.212	-0.012	0.231	1				
$\lambda_2^{(1)}$	0.038	0.102	0.023	0.054	0.053	0.078	0.009	0.063	0.053	1			
$\lambda_2^{(2)}$	0.002	0.001	0.001	0.001	0.001	0.013	0.004	0.005	0.002	0.010	1		
θ_3^C	0.231	0.355	0.070	0.647	0.225	0.325	0.114	0.865	0.232	0.180	0.016	1	
θ_3^N	0.142	0.409	0.057	0.221	0.773	0.212	-0.009	0.239	0.960	0.114	0.017	0.246	1