ISER Working Paper Series

Birth Weight and the Dynamics of Early Cognitive and Behavioural Development

Emilia Del Bono
Institute for Social & Economic Research
University of Essex

John Ermisch
Institute for Social & Economic Research
University of Essex

No. 2009-16
May 2009

INSTITUTE FOR SOCIAL
& ECONOMIC RESEARCH

www.iser.essex.ac.uk

Non-technical summary

Recent medical advances have meant that an increasing number of children born at low birth weight (<2,500 grams) and extremely low birth weight (<1,500 grams) are surviving. At the same time, a growing body of epidemiological research has highlighted a strong association between low birth weight and infant mortality, various morbidities (asthma, high blood pressure, lung disease, etc.), cognitive and emotional impairment in childhood, and depression and anxiety in adulthood. Not surprisingly, this body of evidence has had a very strong impact on health policies in many developed and developing countries.

Economic research has largely supported the importance of birth weight as a policy target. Using datasets much larger than those available to epidemiologists, and different empirical strategies, most economic studies have found that low birth weight children tend to have lower educational attainment, poorer self-reported health status, higher disability rates, and reduced employment and earnings as compared to their heavier counterparts. A closer look at these findings however reveals several inconsistencies. In some cases the estimated effects are rather large, in others very small. Several papers find a significant impact of birth weight on adult outcomes, but apparently no big effects during the perinatal period or in early childhood.

The aim of our study is to provide new evidence on the effects of birth weight on a variety of child outcomes. In order to do so, we use data from the UK Millennium Cohort Study. Our focus is on the early stages of child cognitive and behavioural development, as we believe that if birth weight has large long-term effects on education, income, and earnings, we should observe at least some impact on earlier outcomes.

Our empirical strategy differs from what has been done in the previous literature in several aspects. Whereas most existing studies rely on within-mother variation in birth weight (using twin or sibling-pairs) in order to control for unobserved characteristics of the mothers, we exploit the availability of multiple outcomes observed for the same individual at the same point in time. Secondly, we adopt a production function approach, which theoretically underpins some of the assumptions required by our empirical setting. Finally, we explore the role of dynamics, i.e. we analyse whether the effect of birth weight on later outcomes acts independently of previous achievements.

Our results show that birth weight has a significant effect on male cognitive development at age 3 and on female cognitive and behavioural outcomes at age 3. The magnitude of these effects is however very small. We estimate that a 200-250g increase in birth weight - the most realistic policy target - would increase age 3 cognitive scores by at most 0.03-0.04 standard deviations for boys and at most 0.04-0.05 standard deviations for girls. We also find that birth weight has no significant impact on outcomes measured at age 5 other than through previous outcomes, and in the latter case the dynamics are such that the overall impact of birth weight fades out over time. These results call into question the effectiveness of birth weight as a policy target and highlight the fact that much remains to be understood about the factors which impact on early child development.

Birth Weight and the Dynamics of Early Cognitive and Behavioural Development

Emilia Del Bono*
Institute for Social & Economic Research
University of Essex
and IZA, Bonn

John Ermisch
Institute for Social & Economic Research
University of Essex

In this paper we explore the impact of birth weight on children's cognitive and behavioural outcomes using data from the UK Millennium Cohort Study. In order to deal with the endogeneity of birth weight we use an estimator based on the eliminant method. When coupled with ordinary least squares, this estimator allows us to bound the effects of birth weight. The results show that birth weight has significant but very small effects on male cognitive development at age 3 and on female cognitive and behavioural outcomes at age 3. We also find that birth weight affects age 5 outcomes only through previous achievements, and that the overall impact fades out over time. These findings call into question the effectiveness of birth weight as a policy target.

Keywords: birth weight, production function, child development
JEL classification: I12, I21, J13, J24

Financial support from this project was provided by the UK Economic and Social Research Council (RES-000-22-2044). We would like to thank the MCS team at the Centre for Longitudinal Study, Institute of Education, University of London for their assistance with the data. Any mistakes are the authors' sole responsibility.

*Corresponding author at: Institute for Social & Economic Research, University of Essex, Colchester CO4 3SQ, UK. Tel.: +44 01206 873569; fax: +44 01206 873151.
Email addresses: edelbono@essex.ac.uk (E. Del Bono); ermij@essex.ac.uk (J. Ermisch)

1. Introduction

Recent medical advances have meant that an increasing number of children born at low birth weight (<2,500 grams) and extremely low birth weight (<1,500 grams) are surviving. At the same time, a growing body of epidemiological research has highlighted a strong association between low birth weight and infant mortality, various morbidities (asthma, high blood pressure, lung disease, etc.), cognitive and emotional impairment in childhood, and depression and anxiety in adulthood. These findings are not confined to highly selected groups of low birth weight infants, but can be generalized across the entire birth weight distribution (Kelly et al. 2001).

Not surprisingly, this body of evidence has had a very strong impact on health policies around the globe. Reducing the incidence of low birth weight is a stated goal of several social programs targeting infant health in the United States (Medicaid), the UK (Tackling Health Inequalities: A Programme for Action), and many other countries (World Bank Integrated Nutrition Project). Large sums of money are spent on research focusing on the prevention of low birth weight births and on smoking cessation programs, as maternal smoking has been identified as the most significant modifiable risk factor for the occurrence of low weight births in developed countries (Kramer 1987).

Economic research has largely supported the importance of birth weight as a policy target. Using datasets much larger than those available to epidemiologists, and identification strategies based on twin-pair differences or the occurrence of quasi-natural experiments, most economic studies have found that low birth weight children tend to have lower educational attainment, poorer self-reported health status, higher disability rates, and reduced employment and earnings as compared to their heavier counterparts (Behrman and Rosenzweig 2004; Case et al. 2005; Almond 2006; Oreopoulos et al. 2006; Black et al. 2007).

A closer look at these findings however reveals several inconsistencies. In some cases twin studies show effects which are larger than the corresponding cross sectional effects (Behrman and Rosenzweig 2004). Others find the opposite (Almond et al. 2005; Royer 2009). It seems that the magnitude of the estimates is very unstable across cohorts, and that the effects of birth weight are largest on long-term rather than short-term outcomes (Black et al. 2007). There is also no consensus on whether the returns to birth weight are largest at the lower tail (Behrman and Rosenzweig 2004) or in the normal range (Royer 2009) of the birth weight distribution.

The aim of this study is to provide new evidence on the effects of birth weight on a variety of child outcomes using a new UK dataset. In doing so, we offer several contributions to the previous literature.

First and foremost, we focus on the effect of birth weight on the early stages of child development. It is difficult to believe that birth weight will have large long-term effects on education, income, and earnings if we cannot observe any impact on ability measures collected at an early age. Also, we consider here behavioural as well as cognitive measures of development, as there is increasing recognition that both these domains contribute to adult socio-economic outcomes (Heckman et al. 2006).

Another important element of departure from the previous literature is our identification strategy. Whereas most studies on the effect of birth weight make use of within-mother variations in birth weight (using twin or sibling-pairs), our empirical strategy exploits the availability of multiple outcomes observed for the same individual at the same point in time in order to net out the effect of unobservables.[1] When coupled with ordinary least squares, this approach allows us to bound the effect of birth weight on children's outcomes and make informative comparisons with previous results.

Thirdly, we adopt a production function approach. All empirical studies in which birth weight is an "input" in the production of later outcomes are embedded within a reduced-form framework. By contrast, there is a large literature which analyses the production technology of infant health, where birth weight is essentially seen as an "outcome" (Corman et al. 1987; Grossman and Joyce 1990; Rosenzweig and Wolpin 1991, 1995; Del Bono et al. 2008). We will show here that by following a production function approach it is possible to provide a sound theoretical underpinning for some of the assumptions required by our empirical framework.

Finally, as our data provide information on multiple outcomes at different points in time, we explore the role of dynamics. In other words, we analyse whether the effects of birth weight act independently of previous achievements or not. As one can imagine, this is a question of very high policy relevance. If birth weight influences all stages of development irrespective of previous outcomes, then it would be very difficult for any policy to "undo" the damage caused at birth. If, on the other hand, the effects of birth weight on future outcomes mostly depend on the way this variable affects very early aspects of development, policy interventions targeted at low birth weight children which are implemented very soon after birth might be most effective in reducing socio-economic inequalities.

Our findings are relevant for the future development of health research and for the definition of effective policy targets. Although we find that there is a positive and significant association between birth weight and age 3 cognitive outcomes for boys, and age 3 cognitive and non-cognitive outcomes for girls, the magnitude of the effects is very small. We also find that birth weight has no significant impact on outcomes measured at age 5 other than through previous outcomes, and that the dynamic effects are such that the overall impact of birth weight fades out over time. These findings are in line with recent research by Almond et al. (2005), who report very small effects of birth weight on infant mortality, and Royer (2009), who finds negligible impacts on a range of short and long-run outcomes.

The paper is organized as follows. Section 2 describes our dataset and our measures of child development in some detail. Section 3 presents our methodological framework and explains our identification strategy. The main results are shown and commented upon in Section 4.

[1] For another recent application of this method see also Conti and Pudney (2007).

The last section concludes, summarizing our main findings and discussing their policy implications.

2. Data

The data used in this analysis come from the Millennium Cohort Study (MCS), a large prospective study of infants born in 2000-2002 in the United Kingdom. The first wave of data collection took place when the infants were around 9 months old and includes data on 18,818 infants in 18,552 families. The sampling design allowed for over-representation of areas with high levels of childhood deprivation and high proportions of ethnic minorities. Infants born on eligible dates in eligible areas were selected from the Child Benefit Register.[2]

At the first interview the mothers (usually the main respondent) were asked detailed information about pregnancy, birth, infant health, infant development, and several questions about their mental health, health behaviour and their social and economic circumstances. The second and third interviews took place when the children were about 3 and 5 years old, respectively. At this time, the interviewer administered a physical and a cognitive assessment of the child, while the mother was asked to report about the socio-economic circumstances of the family as well as the child's health and emotional development. [3]

The sample used in this analysis excludes families with multiple births, families where the main respondent was not the natural mother, and non-white families. We also exclude observations with missing information on any of the outcomes, including measures of cognitive and non-cognitive development at ages 3 and 5 as well as birth weight and gestation. The analytical sample therefore consists of 7,657 children, of which 3,826 are boys and 3,831 are girls.

Data on birth weight, gestation, and other indicators of the health of the child at birth – such as the type of delivery, and whether the newborn had breathing problems at birth – were derived from the first wave. For the purposes of this analysis birth weight is expressed in kilograms, and gestation is represented by a dummy variable assuming value 1 if the birth was preterm (<37 weeks). Means and standard deviation of weight at birth are presented in table 1 for boys and girls separately. As we can see from the table, the mean birth weight for boys is about 3.5kg for boys and 3.4kg for girls. We also see that only about 6 percent of boys and 4 percent of girls are born preterm.

The main behavioural assessments at ages 3 and 5 are based on the *Strengths and Difficulties Questionnaire* (SDQ), which is part of a self-completion module filled out by the main carer (in our case this is always the mother). The SDQ contains 25 items, the responses to which are: 'not true', 'somewhat true' and 'certainly true'. These responses generate scores on five separate items, including: conduct problems, emotional problems, hyperactivity problems,

[2] The Child Benefit is a universal benefit in the UK. The benefit accrues to parents of children aged under 16 (or under 20 and in relevant education or training) who are resident in the UK.

[3] The data for the first three waves of the Millennium Cohort Study are publicly available through the UK Data Archive: SN 4683 (First Survey); SN 5350 (Second Survey); SN 5795 (Third Survey). For more information, please see http://www.data-archive.ac.uk/findingData/mcsTitles.asp.

problems with peers and prosocial behaviour (Goodman 1997). We consider only the first 4 subscales, and derive a total score at ages 3 and 5.[4] All subscales have been recoded so that a higher value represents better socio-emotional skills or behaviour. Another measure of emotional development taken at ages 3 and 5 is the child *Self-Regulation* score, as measured by selected items from the Child Social Behaviour Questionnaire used on the EPPE and EPPNI projects for children aged 7 and 10 (Sammons et al. 2004; Melhuish et al. 2004). Higher values on this score indicate a higher ability of the child to perform tasks independently, to concentrate and to control his/her emotional responses. Finally, the self-completion questionnaire at age 3 reported several items from the *Child-Parent Relationship Scale* (*Pianta*). The items involve the respondent's feelings and beliefs about her relationship with her child, and about the child's behaviour towards her. A higher score reflects a more positive relationship (Pianta 1992). All these indicators were transformed into z-scores, using the sample mean and the standard deviation of the raw variables. Means and standard deviations of the raw scores are presented in table 1, and show that girls usually score slightly higher on all these measures.

At age 3 two measures of cognitive ability were obtained. The *British Ability Scales Naming Vocabulary* (BAS Naming Vocabulary) assesses the spoken vocabulary of young children. The *School Readiness Composite* (SR Composite) of the Revised Bracken Basic Concept Scale is used to assess basic concept development. This indicator is thought to be directly related to early childhood education and to predict readiness for more formal education. At age 5, the list of cognitive measures include: the *British Ability Scales Naming Vocabulary*, the *British Ability Scales Picture Similarities* (BAS Picture Similarities), and the *British Ability Scales Pattern Construction* (BAS Pattern Construction). All these assessments were administered by the interviewer, who demonstrated how to perform the test before the child attempted to do so. Our analysis uses the age-adjusted BAS ability scores, which reflect the raw score and the difficulty of the items administered; and the School Readiness Composite Standard Score, which is derived from the total number of correct answer to the six sub-tests (colours, letters, numbers/counting, sizes, comparisons, shapes) of the Braken School Readiness Assessment.[5] These scores were then converted into z-scores using the sample means and standard deviations. Descriptive statistics of the non-standardized values of these variables are presented in table 1. In all cases, increases in the scale signify higher cognitive ability.

The main parental inputs considered in this analysis include measures of maternal health and health behaviour. In particular, we look at the impact of breastfeeding (incidence and exclusivity), maternal smoking, maternal drinking, and maternal nutrition (proxied by her body mass index). We consider the number of immunizations the child has had as another proxy of maternal health behaviour, and distinguish children who had all immunization recommended for their age from those who had only some but not all the recommended doses of vaccines.[6] Maternal health can also be thought of in terms of her mental well-being. This is measured at 9 months by a selection of items from the *Condon Maternal Attachment Questionnaire* (Condon and Corkindale 1998), and a modified version of the *Malaise Inventory Scale* (Rutter et al. 1970).

[4] Pro-social behaviour exhibited hardly any variation according to the characteristics of the child or her parents.
[5] See Hansen (2008) for more details on these measures and their derivation.
[6] We do not consider the MMR vaccine because of the controversy surrounding its potential side-effects.

At ages 3 and 5, maternal mental health is captured instead by the *Kessler Scale* (Kessler 6-item), a screening device frequently used to diagnose mental illness (Kessler et al. 2003).

3. Production function approach

Most approaches to the identification of the effect of birth weight on future outcomes are based on reduced-form specifications, whereby outcomes at various points in time are regressed on birth weight (or a function of it, when non-linearities are taken into account) and a vector of family and local area characteristics. Given that birth weight is endogenous, as it is influenced by maternal characteristics which are also likely to affect later outcomes, identification of its impact is achieved by means of twin comparisons (Behrman and Rosenzweig 2004; Almond et al. 2005; Oreopoulos et al. 2006; Black et al. 2007; Royer 2009), or comparisons across cohorts differentially exposed to a particular historical event (see Lumey and Stein 1997 for an analysis of the impact of the 1944 Dutch famine, and Almond 2006 for a study of the 1918 influenza epidemic).

However, while reduced form models are appealing and provide key information on the likely importance of birth weight on a variety of outcomes, the estimates may be unstable over time. In contrast, estimates based on structural models, when appropriately specified, offer a sounder basis for policy advice (Rosenzweig and Schultz 1988).

For this reason, we adopt here a different approach, and specify the relationship between birth weight and a range of child outcomes within a production function framework. Two elements are key for the successful implementation of such an empirical strategy. The first is the availability of a rich set of parental inputs that can be thought to have an impact on the outcomes of interest. Secondly, as the process of child development takes place over time, it is important to measure outcomes at different ages and take into account dynamic effects, whereby previous outcomes become inputs in the production of future outcomes. This is in line with Cunha and Heckman (2007), who investigate the dynamic process governing child development and distinguish "critical" and "sensitive" periods for parental and government interventions, and recent work by Todd and Wolpin (2003, 2007).

To date, the UK Millennium Cohort Study offers data on child outcomes at four points in time: at birth, at 9 months, at 3 years and at 5 years. In the analysis which follows we will look at the impact of outcomes at birth - the child's birth weight in particular - on different aspects of development at ages 3 and 5. This is because measures of child development at ages 3 and 5 are more directly comparable over time, and can be easily classified as belonging to the cognitive (BAS various modules, and SR Composite) or socio-emotional sphere (SDQ, Self-regulation, Pianta). By contrast, outcomes at 9 months, which include results from the *Denver Development Test* and the *Carey Infant Temperament Scale*, have no immediate counterpart at ages 3 and 5, and combine elements of physical, emotional and cognitive development which at this early stage are very hard to separate.

As we consider here many different outcomes at various points in time, it is useful to introduce some specific notation. For each outcome measure at age 5, m_i^5, let the following be its production function:

(1) $$m_i^5 = \sum_k \beta_{ik}^5 x_k^5 + \sum_j \gamma_{ij}^5 m_j^3 + \sum_l \delta_{il}^5 m_l^0 + \alpha_i^5 \theta + u_i^5, \text{ for } i=1, \ldots, n^5; k=1,\ldots,K^5,$$

where x_i^5 are inputs measured at age 5 (which might be indicative of inputs used between ages 3 and 5), m_j^3 are outcomes measured at age 3, and m_l^0 are outcomes measured at birth, which include but are not limited to birth weight. The next term in (1) includes θ, which represents a child-specific factor influencing all outcomes measured at 5 years, while u_i^5 are zero mean residual influences including measurement error. The parameters to be estimated are β_{ik}^5, γ_{ij}^5, α_i^5 and δ_{il}^5.

Analogously, the production function at age 3 can be specified as follows:

(2) $$m_i^3 = \sum_k \beta_{ik}^3 x_k^3 + \sum_l \delta_{il}^3 m_l^0 + \alpha_i^3 \theta + u_i^3, \qquad \text{for } i=1, \ldots, n^3; k=1,\ldots,K^3.$$

One could write a similar equation for the outcomes at birth, m_l^0, as these are a function of parental investments made during pregnancy or even before. However, since we are interested here in the effect of birth weight on child development, we consider the vector of outcomes measured at birth, m_l^0, simply as inputs in the production function of the outcomes measured at ages 3 and 5.

The inputs, x_k^a (for $a=3,5$), include mother's age and its square, parental investments such as breastfeeding, maternal smoking, maternal drinking, maternal nutrition (proxied by the mother's body mass index), immunizations, and maternal mental health. The birth outcomes, m_l^0, are the child's birth weight, whether the child was born premature (before 37 weeks of gestation), the type of delivery (caesarian, instrumental or natural), and whether the child had breathing problems at birth.[7]

Of course, estimating (1) or (2) by OLS is likely to lead to biased estimates of the parameters of interest as many of the inputs we consider are possibly correlated with unobserved parental characteristics or preferences which also affect outcomes. Also, no matter how comprehensive our list of inputs, it is possible that there are omitted terms in equations (1) and (2). All of these considerations give rise to serious concerns about the exogeneity of the input measures; that is there are good reasons to suspect that the x_k^a and m_l^0 are correlated with θ or with u_i^a, or both.

In order to deal with this problem we need to impose further assumptions onto the model in (1) and (2). In particular, we will assume that $E(u_i^a x_k^a)=E(u_i^a m_l^0)=0=E(u_i^a \theta)$; $E(u_i^a u_{i'}^a)=0$ for $i \neq i'$

[7] Other controls include whether the child was the first child, as that usually implies a lower birth weight, and the logarithm of age (in days) of the child at the time of the interview. These should not be thought of as inputs, but simply as possible confounders of the effects of birth weight, the main variable of interest.

and $E(u_i^a u_i^a) = \sigma_{ia}^2$. Also, we will assume that $E(u_i^3 u_i^5)=0$, but we will allow correlation between the inputs, x_k^a and m_l^0, and the common factor, θ.

In particular, we assume the following structure:

(3) $$x_k^a = \lambda_k^a \theta + \nu_k^a, \qquad \text{for } k=1, \ldots, K^a,$$

(4) $$m_l^0 = \pi_l^0 \theta + \xi_l^0, \qquad \text{for } l=1, \ldots, n^0,$$

where $E(\nu_k^a \theta)=E(\xi_l^0 \theta)=0=E(\nu_k^a u_i^a)=E(\xi_l^0 u_i^a)$, $\text{cov}(\nu_{k'}^a \nu_k^a)=\sigma_{\nu_k \nu_{k'}}^a$ (for $a=3,5$), and $\text{cov}(\xi_{l'}^0 \xi_l^0)=\sigma_{\xi_l \xi_{l'}}^a$.

The systems in (1)-(4) assume that there is a common factor θ for each child, and there are multiple outputs from the same set of inputs at age a. The effect of each input can vary among outputs (β_{ik}^a and δ_{il}^a vary with i and a), and the effect of the common factor θ can vary among outputs (α_{ik}^a varies with i and a). Similarly, the dynamic impact of earlier endogenous outputs varies across future outputs and over time (γ_{ij}^a varies with i and a). The choice of input levels is correlated with the common factor θ and the correlation can vary with the inputs, i.e. λ_k^a varies with k and a, while π_l^0 varies with l and a. Thus, the endogeneity of all inputs (x_k^a and m_l^0) only operates through θ.

Our main interest is in the parameters representing the effect of the birth outcomes, δ_{il}^a, and in particular birth weight, but we will also estimate the effects of parental inputs between birth and age a, β_{ik}^a, and the effects of previous endogenous outcomes, γ_{ij}^a. The OLS estimates of these parameters will be biased upwards since we expect that: $E(x_k^a \theta)=\lambda_k^a \text{var}(\theta) \geq 0$; $E(m_l^0 \theta)=\pi_l^0 \text{var}(\theta) \geq 0$; and $E(m_i^3 \theta)=\{\sum_k \beta_{ik}^3 E(x_k^3 \theta)+\sum_l \delta_{jl}^3 E(m_l^0 \theta)+\alpha_i^3 \text{var}(\theta)\} \geq 0$.[8] Thus, OLS should provide estimates of β_{ik}^a, δ_{il}^a, and γ_{ij}^a which are *not smaller* than their true values.

The other estimator we use is an *eliminant estimator*, similar to the one introduced in Pudney (1982) to estimate the parameters of a class of models involving latent variables and recently applied to the study of the dynamics of cognitive development by Conti and Pudney (2007).[9] The idea here is remove the common factor by a simple transformation of the model before estimation, and then estimate the new system using 3SLS.

For $i=1$ and $a=3$, we can write $\theta=(m_1^3-\sum_k \beta_{1k}^3 x_k^3 - \sum_l \delta_{1l}^3 m_l^0 - u_1^3)/\alpha_1^3$; substituting for θ into the system (1) yields:

---

[8] All outcomes are re-scaled so that the higher is the value the more favourable is the outcome, so that we expect that $\alpha_i^a \geq 0$ for $i=1, \ldots n^a$ (a=3,5).

[9] Jakubson (1991) also used an 'eliminant procedure' similar to the one here to test for the validity of the conventional fixed effect panel data model against an alternative in which parameters and factor loadings on the fixed effect can vary over time.

$$(5)\quad m_i^5 = \sum\nolimits_k \beta_{ik}^5 x_k^5 - \frac{\alpha_i^5}{\alpha_1^3}\sum\nolimits_k \beta_{1k}^3 x_k^3 + \frac{\alpha_i^5}{\alpha_1^3} m_1^3 + \sum\nolimits_j \gamma_{ij}^5 m_j^3 + \sum\nolimits_l \left(\delta_{il}^5 - \frac{\alpha_i^5}{\alpha_1^3}\delta_{1l}^3\right) m_l^0 + \left(u_i^5 - \frac{\alpha_i^5}{\alpha_1^3} u_1^3\right).$$

Note that m_1^3 is endogenous in (5) as it is correlated with the error term through u_1^3, and the error terms in the system of equations represented by (5) are correlated among each other because of u_1^3. On the other hand, x_k^5, x_k^3, m_j^3, and m_l^0 are exogenous in (5), as the source of endogeneity runs only through the common factor θ and this has now been substituted out.

Estimation of the system in (5) through 3SLS would provide consistent estimates of β_{ik}^5, and γ_{ij}^5 as well as a consistent estimate of the parameter (δ_{il}^5-$\alpha_i^5\delta_{1l}^3/\alpha_1^3$).

Substituting θ into the rest of the system (2) yields:

$$(6)\qquad m_i^3 = \sum\nolimits_k \left(\beta_{ik}^3 - \frac{\alpha_i^3}{\alpha_1^3}\beta_{1k}^3\right) x_k^3 + \frac{\alpha_i^3}{\alpha_1^3} m_1^3 + \sum\nolimits_l \left(\delta_{il}^3 - \frac{\alpha_i^3}{\alpha_1^3}\delta_{1l}^3\right) m_l^0 + \left(u_i^3 - \frac{\alpha_i^3}{\alpha_1^3} u_1^3\right).$$

Estimating the system (6) through 3SLS would provide consistent estimates of (β_{ik}^3-$\alpha_i^3\beta_{1k}^3/\alpha_1^3$) and ($\delta_{il}^3$-$\alpha_i^3\delta_{1l}^3/\alpha_1^3$).

Since x_k^a and m_l^0 are all inputs in the production function of m_i^3 and m_i^5, we expect $\beta_{1k}^3 \geq 0$ and $\delta_{1l}^3 \geq 0$, so that 3SLS estimates of (β_{ik}^3-$\alpha_i^3\beta_{1k}^3/\alpha_1^3$), ($\delta_{il}^3$-$\alpha_i^3\delta_{1l}^3/\alpha_1^3$), and ($\delta_{il}^5$-$\alpha_i^5\delta_{1l}^3/\alpha_1^3$) should be *not larger* than the structural parameters β_{ik}^3, δ_{il}^3, and δ_{il}^5, respectively. This means that, apart from sampling error, the parameters representing the effect of birth outcomes δ_{il}^3 and δ_{il}^5 should lie between the OLS and the 3SLS estimates of (1) and (2). Clearly, the choice of which outcome represents m_1^3 matters, and it might change the estimates of what we consider as the lower bounds of the parameters of interest. We will return to this issue below.

4. Parameter estimates

The analysis is performed separately for boys and girls, as girls tend to be lighter at birth but usually score better on several measures of development (see table 1). We consider here a total of five outcomes pertaining to the behavioural domain: the SDQ score at ages 3 and 5, the Self-Regulation score at ages 3 and 5, and the Pianta score at age 3; and five outcomes pertaining to the cognitive domain: the SR Composite score at age 3, the BAS Naming Vocabulary at ages 3 and 5, the BAS Picture Similarities at age 5, and the BAS Pattern Construction at age 5.[10]

---

[10] Although in principle we could have considered other domains, such as those related to the physical health of the child, our identification strategy is based on the existence of a single common factor which affects all the outcomes of interest. It is clear that this assumption would become weaker if we were to expand the range of possible outcomes.

The transformation of the systems (1) and (2) into the systems (5) and (6) is performed by choosing an outcome, m_1^3, in order to net out the influence of the unobserved common factor. However, m_1^3 is endogenous in the systems (5) and (6). This implies that we need some instruments. Any variable z which satisfies the moment condition $E(zu_i^a)=0$ (for all *i* and *a*) is a valid instrument for m_1^3, that is the instrumental variables may be correlated with θ, and indeed we hope they are.

So, the instruments we propose are variables which reflect the family socio-economic position such as: mother's education, family income, tenure of the household's accommodation, combined employment and partnership status of the mother, and number of siblings. These variables are not only well placed to capture family-specific unobserved elements which have an impact on various aspects of child development, but are also valid restrictions within our theoretical framework. Indeed, according to the production function approach adopted here, family socio-economic circumstances are not inputs, but simply constraints that affect the way in which the inputs are chosen. To the extent that our model includes all the relevant inputs, these variables are therefore rightly excluded from our specification.[11]

Given these considerations, the choice of which outcome should represent m_1^3 is based on the following consideration. In table 2 we regress each outcome on the vector of inputs (including birth outcomes) and the vector of instruments. We do so for all outcomes, ages, and specifications of the model. As we can see, the vector of instruments exhibits the highest explanatory power for the SR Composite score at age 3. For this outcome the F-statistics on the instruments is 12.96 in the basic model for boys, and 18.57 in the basic model for girls. When we take into account maternal health and health behaviours, these statistics fall to 7.85 and 11.35 for boys and girls, respectively, but remain consistently higher than what is observed for other outcomes.

Impact of birth weight at age 3

We now analyse the parameters obtained from the estimation of the systems (1) and (2) through OLS, and those obtained from the estimation of the systems in (5) and (6) through 3SLS. Table 3 to 7 present each set of parameter estimates for the outcomes at ages 3 and 5. Each system is estimated separately by gender of the child, and each table presents two specifications: a basic model where no inputs other than the birth outcomes and some basic variables such as age of the mother are included (Panel A), and a model which considers also the mother's health and health behaviours (Panel B). As we will see, the results are broadly consistent across these different specifications.

We start by considering the effect of birth weight on measures of cognitive and behavioural outcomes for the subsample of boys in the basic model (Panel A of table 3). As we can see, the OLS estimates show that birth weight has no significant effect on non-cognitive measures,

---

[11] If there were relevant omitted inputs correlated with the included ones in our specification, these exclusion restrictions would be more difficult to justify. Notice however that any time-invariant omitted input is captured by the common factor, and the problem would arise only if we were omitting relevant time-varying inputs.

but has a positive impact on the main measure of cognitive ability at age 3, the BAS Naming Vocabulary score. The 3SLS estimates represent a lower bound of the structural parameters, but reflect the same picture as the OLS estimates, showing a significant positive impact of birth weight only for the cognitive measure. The 3SLS coefficient is clearly below the OLS coefficient, but despite the relatively large sample size and despite the fact that the endogenous outcome (SR Composite) is capturing a significant amount of variation, the difference is not significant. Similar results can be seen when considering the second specification of the model (Panel B of table 3), which includes maternal health and health behaviours.

Results for the subsample of girls are presented in table 4. Here the OLS regressions show that the impact of birth weight is more wide-ranging, touching upon all aspects of behaviour as well as cognitive outcomes. Estimates from the 3SLS model are more conservative, but still show significant effects of birth weight on two non-cognitive outcomes (the SDQ and Self-Regulation) as well as the cognitive measure. Once again, OLS and 3SLS are not significantly far apart, although there is clear evidence that the endogenous outcome is significant in all specifications. The input-augmented specification in the bottom panel shows very similar qualitative results, except that in this case the 3SLS estimates show no significant effect of birth weight on Self-Regulation.

To sum up, these regressions show significant effects of birth weight on measures of cognitive outcomes for boys, and on measures of behavioural and cognitive outcomes for girls at age 3. The effects are quite small in magnitude, however. Taking for example the OLS estimates in the basic model for boys, we see that a 1kg increase in birth weight improves the score on the BAS Naming Vocabulary scale by about a sixth of a standard deviation (0.156). For girls, such an increase in birth weight would improve the scores in the SDQ and in the BAS Naming Vocabulary by about a fifth of a standard deviation (0.212 and 0.218, respectively). Any reasonable policy intervention aimed at modifying birth weight (such as smoking cessation programs) can achieve an increase in birth weight of 200-250g at best, so that a policy targeting birth weight as a way to improve children's emotional and cognitive development would achieve modest results according to these parameter estimates.[12]

Although we are mainly interested in the effect of birth weight, it is worth considering briefly our results with respect to the other inputs in the model, as they have policy relevance in their own right. Let us take for example the effects of breastfeeding. A growing number of epidemiological studies have shown a positive association between breastfeeding and measures of child and adult IQ (Anderson et al. 1999; Caspi et al. 2007; Kelly and Watts 2005; Mortensen et al. 2002; Oddy 2006), although recent evidence exploiting changes in maternity leave legislation or studying sibling pairs has questioned the existence of a causal link (Baker and Milligan 2008; Der et al. 2006). As shown in tables 3 and 4, while the OLS estimates show positive effects of breastfeeding over the entire range of outcomes considered, and in

[12] Boys [girls] in the lower income quartile score 0.33 [0.30] points of a standard deviation below the mean in the BAS Naming Vocabulary at age 3. According to our OLS estimates, increasing their birth weight by 200-250g would lead to an improvement in their cognitive scores of only 0.03-0.04 [0.04-0.05] points of a standard deviation.

particular on the SDQ and the BAS Naming Vocabulary, the 3SLS estimates are always insignificantly different from zero. This result casts significant doubts on the importance of breastfeeding as a means to improve children outcomes.

We also consider other maternal health behaviours, such as smoking, drinking and nutritional intake (here proxied by maternal body mass index or BMI). As we can see from looking at the OLS estimates, maternal smoking is found to have a negative effect on almost all measures of child cognitive and non-cognitive behaviour for boys as well as girls. The 3SLS are more conservative, however, and show negative effects of mother smoking on mother-child interaction scores (Pianta) and cognitive ability for boys, and on measures of non-cognitive behaviour for girls. Maternal drinking, maternal BMI and immunization practices show basically no significant effects on the outcomes considered. The main exception is represented by a negative effect of maternal drinking on the mother-child interaction for girls.

Mother's mental health - either measured at 9 months through the Postnatal Attachment score and the Malaise Inventory score, or measured at age 3 through the 6-item Kessler score - shows a significant impact only for outcomes in the socio-emotional domain. It is possible that this signifies a true causal relationship, but it could also indicate the extent to which the reporting of the child's behavioural outcomes is affected by the mother's own mental well-being (cognitive outcomes are assessed by the interviewer).

The last rows of table 3 and 4 report the coefficient on the endogenous outcome variable, the SR Composite. The estimates show that this is significantly related to all the outcomes analysed, and therefore can be taken as a good representation of the common unobserved factor.

Impact of birth weight at age 5

We now turn to analyse the results at age 5 which are presented in tables 5 and 6. Because of the dynamic structure of the model, our specifications now include past outcomes and, for the transformed system in (5) also past inputs. As shown in the previous section, by estimating (5) we can identify the exact structural parameters on the age 5 inputs, but only lower bounds on the parameters related to the birth outcomes. So, as in the age 3 models, we expect that – sampling error aside – the 3SLS estimate of birth weight lies below the OLS estimate.

The first aspect to comment upon is the role of the dynamic elements. As we can see from the top panel in table 5, for most behavioural measures outcomes at age 3 have a significant effect on outcomes at age 5. The correlations are not as high as we expected, however, so that for example a SDQ score one standard deviation higher at age 3 translates into a SDQ score less than half a standard deviation higher at age 5. The dynamic effects for the other outcomes are even smaller. Consider, however, that there is no counterpart at age 3 of the BAS Picture Similarities or BAS Pattern Construction scores measured at age 5, so it is not surprising to find little evidence of dynamic effects in these cases. Also notice that the coefficient on the SR Composite is almost always significant, and this shows that even though it refers to an age 3 outcomes it is still a satisfactory measure for the common unobserved factor at age 5.

As for birth weight, when looking at boys (table 5) we find a significant impact of this variable only on measures of cognitive ability represented by the BAS Pattern Construction score. This is so before and after including maternal health and health behaviours. By contrast, for girls (table 6) the results of the basic model show some small and significant effects of birth weight on measures of non-cognitive ability (SDQ) and on different aspects of cognitive behaviour (BAS Picture Similarities and BAS Pattern Constructions). Once we control for maternal health and health behaviour, however, the only significant effects are confined to measures of cognitive ability.

Once again, the effects of birth weight are not large. According to the OLS estimates of the basic model, a 1kg increase in birth weight would shift the cognitive score by at most 0.142 points of a standard deviation for boys, and 0.110 points for girls. There are also virtually no differences between the OLS and the 3SLS estimates of this effect. This is explained by the fact that the OLS estimates are significantly reduced when controlling for outcomes at age 3, and that birth weight shows an insignificant impact on age 5 outcomes over and above its impact on age 3 outcomes. This means that once dynamic effects are included, the OLS estimates of the effects of birth weight on age 5 outcomes are affected by very small (if any) bias.[13]

Other effects worth reporting are seen with respect to maternal smoking, which is found to decrease measures of cognitive ability for boys but not for girls. Having less than the recommended number of immunizations is also found to decrease the SDQ for boys, but not for girls. Maternal drinking (moderate), by contrast, is found to be of no significance or even to have a positive impact on cognitive and non-cognitive scores for girls. A low maternal BMI has a somewhat negative effect on measures of non-cognitive behaviour for boys, and a rather puzzling positive effect for girls. As we saw for the age 3 results, measures of maternal mental health are found to affect mainly self-reported measures of child development.

All in all these results show a consistent pattern throughout different specifications and subsamples. The overall picture is that of small but statistically significant effects of birth weight on cognitive outcomes at age 3 for boys, and on cognitive and behavioural outcomes at age 3 for girls. We also find that birth weight has very little or no impact on outcomes measured at age 5 other than through dynamic effects, i.e. previous outcomes. A positive and significant estimate on the birth weight coefficient for BAS Pattern Construction scores at age 5 for boys and girls, and the tiny effect observed on the BAS Picture Similarity scores for girls do not contradict this interpretation of the evidence, as there are no age 3 equivalents of these measures of cognitive ability and therefore a limited possibility to control for dynamic effects in these cases. Finally, as the dynamic parameter estimates are relatively small in magnitude we can expect that the overall effects of birth weight fade out rapidly.

[13] As a check, we re-estimated the effect of birth weight on age 5 outcomes by setting the dynamic effects to zero (i.e. excluding the age 3 outcomes from the specification). In this case we observed larger differences between the OLS and 3SLS coefficients, and this was mainly due to larger OLS estimates of the impact of birth weight.

Robustness checks

As discussed above, one of the main assumptions of our approach is that all the relevant inputs in the production function of cognitive and non-cognitive development are captured by the vector of observed inputs analysed here. In particular, we focused on inputs representing maternal health and health behaviour as our main variable of interest is the birth weight of the child. To the extent that other potentially relevant and time-varying inputs are omitted from our specification and that these are correlated with the included inputs, however, our approach is invalid and our instrumentation strategy would be flawed.

We therefore take advantage of the wealth of information offered by the MCS and include among our inputs other aspects of parental behaviour which could be thought to affect children's early development. We first consider parental educational activities and parenting styles. As in Ermisch (2008), these activities are captured by four main variables: (i) whether the mother reads to the child every day; (ii) whether someone in the family takes the child to the library at least once a month; (iii) the principal component obtained by considering variables describing other educational activities such as teaching the alphabet, numbers, songs, and drawing at age 3, and teaching reading, writing, counting, songs, and drawing at age 5; (iv) the principal component obtained by considering the presence of rules about bed time, meal time and TV time in the family.

Table 7 reports our main results - those obtained in the lower panels of tables 3 and 4 (Panel A) - and confronts them with the results obtained when we include the parental inputs described above (Panel B). As we can see, the estimates of the effect of birth weight are virtually identical, although there is evidence that these additional inputs are capturing something which was previously working through the common factor as the coefficient on the SR Composite is now slightly lower.[14]

Another aspect which is potentially important and which is excluded in our analysis is the consideration of child-care arrangements. Arguably, childcare reflects parental income and should not be considered an input in the production function framework adopted here. Nevertheless it is worthwhile to see whether considering variation in childcare arrangements has an impact on our main results. In Panel C of table 7 we include different dummy variables which capture whether the child is looked after by the grandparents, is in a nursery, or is looked after by friends or neighbours at age 3 (being looked after by the mother is the omitted category). As we can see, the effect of birth weight is very robust to the inclusion of these controls. Interestingly, there is a much smaller effect of these variables on the coefficients of the endogenous outcome, whose coefficient changes only very slightly with respect to the main specification.

Another possible concern with our analysis is that in order to model dynamic effects we need information at different points in time. As long as there is a problem of attrition, considering only individuals who are present at the age 3 and age 5 surveys may introduce bias in the estimated coefficients. Unfortunately, apart from modeling the process of attrition, there is

[14] Complete results for the model estimated at age 3 and for similar specifications estimated at age 5 (not shown) are available from the authors.

not much which can be done in this case. The only check we can perform is to take all the individuals present at the first and second surveys (i.e. irrespective of whether they will also respond to the age 5 questionnaire) and run again our regressions at age 3. The results of this exercise are shown in Panel D of table 7. It is clear that there are some small differences in the OLS coefficients, but the 3SLS estimates of the parameters are very close to those in Panel A. So, as far as we can see, attrition should not be a major issue in our case.

Comparing our results to the existing literature

Although there are comprehensive studies on the impact of birth weight on various outcomes using a variety of identification strategies, this study is unique along several dimensions. It focuses on early measures of children's cognitive and behavioural development, whereas most of the existing research analyses early health outcomes, such as perinatal mortality, or later adult outcomes, such as wages and education. It is also the first such study to adopt a production function approach, whereas most of the existing evidence is based on reduced-form models. Its identification strategy rests on the availability of multiple measures of child outcomes, whereas others have used twin-pair differences or the occurrence of quasi-natural experiments. Finally, it explicitly recognizes the role of dynamics, which are a key aspect in modeling child development, but which have been so far ignored by the literature on birth weight because of its reduced-form nature.

All these considerations make it hard to assess the results we find in this paper in relation to those usually presented in the literature. In an effort to provide some indication of how our results compare to what is found by others we perform a twin-pair analysis of the kind usually seen in the literature on this subject. The MCS dataset contains a very small sample of twins, however, so this exercise and the results we derive here should be taken with a degree of caution.

As our sample of twins is tiny, we consider only outcomes at the first interview (at 9 months of age) in order to minimize problems of attrition which in this case would be very severe. Therefore the outcomes we analyse in this section are different from those we reported on previously. In particular, we consider: (i) the cumulative score on the *Denver Development Test*, (ii) the cumulative score of the *Carey Infant Temperament Scale*, and (iii) whether the child was placed in a neonatal care unit. The *Denver Development Test* is used to assess social and communication skills, as well as fine and gross motor coordination typical for a 9 months old child (Frankenburg et al. 1967). The *Carey Infant Temperament Scale* is a widely used measure to assess infant temperament (Carey and McDevitt 1977; Carey and McDevitt 1995). In the MSC questionnaire three dimensions of temperament were assessed, including mood, receptivity to novelty, and regularity (Pickett et al. 2008).

Table 8 reports pooled OLS and fixed-effects (FE) regressions on our sample of twins for each output measure available at 9 months. The regressions are performed on the entire sample of twins for which information on any of the scores is available, and include non-whites in order to keep the number of observations as high as possible. The pooled OLS regressions control for maternal age and its square, a dummy for preterm, a dummy for firstborn, a dummy for ethnic group white, the log in days at the time of the first interview, and a male dummy. The FE regressions include only a male dummy, as this is the only characteristic that varies across twins.

The first row in table 8 reports our main results. Focusing first on the *Denver Development Test*, we find evidence of a positive effect in the OLS regression, of about 0.62 points of a standard deviation. The FE estimates are smaller though, implying an effect of only 0.12 points of a standard deviation. The FE parameter estimate is not significantly different from zero - and this is not surprising given the small sample - but its magnitude is comparable to that found in Royer (2009) in her analysis of twin-pairs from the Early Childhood Longitudinal Study (her FE point estimate is 0.15).

No effect of birth weight is found in either the OLS or the FE regressions with respect to our second outcome, the *Carey Infant Temperament Scale*. By contrast, when looking at neonatal care we find a significant OLS estimate indicating that a 1kg increase in birth weight reduces the use of intensive care facilities by about 21 percent. Once again, the FE estimate is smaller, showing an overall effect of about 7.4 percent. Although this effect is not significantly different from zero, the magnitude is once again remarkably similar to the 0.04 coefficient found in Royer (2009).

As twins are usually smaller than singletons, any analysis on twin-pairs is difficult to generalize to the overall population. We therefore check the robustness of our results excluding twin-pairs where one of the children weighted less than 1500 grams at birth. This way we intend to exclude variations in birth weight at the very low tail of the distribution. As we can see from the bottom panel of table 7, the results we obtain, particularly after FE estimation, are very close to those on the whole sample of twins. The effect of birth weight on the *Denver Development Test* are about 0.15 points of a standard deviation, while the effect on neonatal care incidence does not change at all.

This simple exercise reveals that the results we obtain from our MCS dataset are comparable with those found for the US. Moreover, and most importantly, the twin comparisons confirm what we find in our main analysis, i.e. that the magnitude of the effect of birth weight on measures of child development is likely to be small. This general result echoes that found in Royer (2009) using birth register data for the U.S. on a variety of short and long term outcomes, and calls for a cautionary note on policies aimed at addressing inequalities over the life-cycle through pre-birth interventions.

5. Conclusions

A consistent association between weight at birth and long-term adult outcomes, such as completed schooling, earnings and income has been established across many different countries using a variety of datasets and identification strategies. This justifies the fact that birth weight has become a direct target of health and social policy. However, large differences in the magnitudes of the estimated long-term effects and the puzzling finding of no or negligible impacts on short-term health outcomes call into question the robustness of these associations and the policy relevance of this indicator of infant health.

In this paper we aim to shed new light on the relationship between birth weight and various measures of early cognitive and behavioural development using recent data from the UK Millennium Cohort Study. In contrast with the recent economic literature on the subject which has focused on within-mother variation in birth weight, our identification is based on within-individual variation in outcomes. The estimator we use addresses the problem of endogeneity by exploiting the availability of different output measures for the same individual at each point in time to net out the effect of unobservables. Considering these estimates together with ordinary least squares estimates, we obtain bounds on the effect of birth weight. Our empirical strategy is embedded in a production function framework, which helps to clarify and justify the assumptions required by our empirical setting. We take explicitly into account the presence of dynamic effects, whereby previous achievements appear as "inputs" in the production function of current outcomes.

Our results show that birth weight has a significant effect on male cognitive development at age 3 and on female cognitive and behavioural outcomes at age 3. The magnitude of these effects is however very small. We estimate that a 200-250g increase in birth weight - the most realistic policy target - would increase age 3 cognitive scores by at most 0.03-0.04 standard deviations for boys and at most 0.04-0.05 standard deviations for girls. We also find that birth weight has no significant impact on outcomes measured at age 5 other than through previous outcomes, and in the latter case the dynamics are such that the overall impact of birth weight fades out over time. These results call into question the effectiveness of birth weight as a policy target and highlight the fact that much remains to be understood about the factors which impact on early child development. The availability of new and very rich data sources, such as the UK Millennium Cohort Study, as well as the adoption of a more structural approach as advocated in Cunha and Heckman (2007) are the main tools which will allow us to achieve this ultimate goal.

References

Almond, D. Is the 1918 Influenza Pandemic Over? Long-Term Effects of In Utero Influenza Exposure in the Post-1940 US Population. Journal of Political Economy 2006; 114(4); 672-712.

Almond, D., Chay, K., Lee, D. The Costs of Low Birth Weight. Quarterly Journal of Economics 2005; 102(3); 1031-83.

Anderson, J., Johnstone, B., Remley, D. Breast-feeding and cognitive development: A meta-analysis. American Journal of Clinical Nutrition 1999; 70; 525-35.

Baker, M., Milligan, K.. Maternal employment, breastfeeding and health: Evidence from maternity leave mandates. Journal of Health Economics 2008; 27; 871-87.

Behrman, J., Rosenzweig, M. Returns to Birthweight. Review of Economics and Statistics 2004; 86(2); 586-601.

Black, S., Devereux, P., Salvanes, K. From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes. Quarterly Journal of Economics 2007; 122(1); 409-39.

Carey, W., McDevitt, S. Infant Temperament Questionnaire. Department of Educational Psychology, Temple University: Philadelphia; 1977.

Carey, W., McDevitt, S. Revised Infant Temperament Questionnaire. Scottsdale, AZ, Behavioural-Developmental Initiatives; 1995.

Case, A., Fertig, A., Paxson, C. The Lasting Impact of Childhood Health and Circumstance. Journal of Health Economics 2005; 24 ; 365-89.

Caspi, A., Williams B., Kim-Cohen, J., Craig, I., Milne, B., Poulton, R., Schalkwyk, L., Taylor, A., Werts, H., Moffit, T. Moderation of breastfeeding effects on the IQ by genetic variation in fatty acid metabolism. Proceedings of the National Academy of Sciences; 5 November 2007.

Condon, J., Corkindale, C. The assessment of parent-to-infant attachment, development of a self-report questionnaire instrument. Journal of Reproductive and Infant Psychology 1998; 16; 57-76.

Corman, H., Joyce, T., Grossman, M.. Birth Outcome Production Functions in the US. Journal of Human Resources 1987; 22; 339-360.

Conti, G., Pudney, S. The Dynamics of Cognitive Development, paper presented at La Pietra-Mondragone Workshop, Florence, 2007.

Cunha, F., Heckman, J. The Technology of Skill Formation. American Economic Review 2007; 97(Papers and Proceedings); 31-47.

Del Bono, E., Ermisch, J., Francesconi, M. Intrafamily Resource Allocations, A Dynamic Model of Birthweight. ISER Working Paper 2008-27; University of Essex: Colchester.

Der, G., Batty, G., Deary, I. Effect of breast feeding on intelligence in children, prospective study, sibling pairs analysis, and meta-analysis. British Medical Journal 2006; 333(7575); 945-48.

Ermisch, J. Origins of Social Immobility and Inequality, Parenting and Early Child Development. National Institute Economic Review 2008; 205; 62-71.

Frankenburg, W., Dodds, J., Denver, B. Developmental Screening Test. Journal of Paediatrics 1967; 71; 181-191.

Goodman, R. The Strengths and Difficulties Questionnaire, A Research Note. Journal of Child Psychology and Psychiatry 1997; 38; 581-586.

Grossman, M., Joyce, T. Unobservables, Pregnancy Resolutions, and Birth Weight Production Functions. Journal of Political Economy 1990; 98; 983-1007.

Hansen, K. Millennium Cohort Study First, Second and Third Surveys, A Guide to the Datasets. Institute of Education: London; 2008.

Heckman, J., Stixrud, J., Sergio, U. Effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. Journal of Labor Economics 2006; 24(3); 411-82.

Jakubson, G. (1991) Estimation and test of the union wage effect using panel data, *Review of Economic Studies* 58:971-991.

Kelly, Y., Nazroo, J., McMunn, A., Boreham, R., Marmot, M. Birthweight and behavioural problems in children, a modifiable effect? International Journal of Epidemiology 2001; 30; 88-94.

Kelly, Y., Watts, R. Breastfeeding initiation and exclusive duration at 6 months by social class, Results from the Millennium Cohort Study. Public Health Nutrition 2005; 8; 417-21.

Kessler, R., Barker, P., Colpe, L., Epstein, J., Gfroerer, J., Hiripi, E., Howes, M., Normand, S., Manderscheid, R., Walters, E., Zaslavsky, A. Screening for serious mental illness in the general population. Archives of General Psychiatry 2003; 60; 184-189.

Kramer, M. Intrauterine Growth and Gestational Duration Determinants. Pediatrics 1987; 80; 502-11.

Lumey, L., Stein, A. In Utero Exposure to Famine and Subsequent Fertility, The Dutch Famine Birth Cohort Study. American Journal of Public Health 1997; 87(12); 1962-66.

Melhuish, E., Hanna, K., Quinn, L., Sylva, K., Siraj-Blatchford, I., Sammons, P., Taggart, B. The Effective Pre-school Provision in Northern Ireland Project. Technical Paper 11: Pre-school Experience and Social/Behavioural Development at the End of Year 3 of Primary School. Stranmillis University Press: Belfast; 2004.

Mortensen, E., Michaelsen, K., Sanders, S., Reinisch, J. The association between duration of breastfeeding and adult intelligence. Journal of the American Medical Association 2002; 287(18); 2365-71.

Oddy, W. Fatty acid nutrition, immune and mental health development from infancy through childhood. In: Huang. J.D. (ed.), Frontiers in Nutrition Research. Nova Science Publishers: New York; 2006.

Oreopoulos, P., Stabile, M., Walld, R., Roos, L. Short, Medium, and Long Term Consequences of Poor Infant Health, An Analysis Using Siblings and Twins. National Bureau of Economic Research Working Paper n. 11998-2006.

Pianta, R. Child-parent relationship scale. Unpublished measure, University of Virginia; 1992.

Pickett, K., Wood, C., Adamson, J., DeSouza, L., Wakschlag, L. Meaningful differences in maternal smoking behaviour during pregnancy, implications for infant behavioural vulnerability. Journal of Epidemiology and Community Health 2008; 62; 318-324.

Pudney, S. Estimating Latent Variable System when Specification is Uncertain, Generalized Component Analysis and the Eliminant Method. Journal of the American Statistical Association 1982; 77(380); 883-89.

Rosenzweig, M., Schulz, T. The stability of household production technology, a replication. Journal of Human Resources 1988; 23; 535-549.

Rosenzweig, M., Wolpin, K. Inequality at Birth, The Scope for Policy Intervention. Journal of Econometrics 1991; 50; 205-28.

Rosenzweig, M., Wolpin, K. Sisters, Siblings, and Mothers, The Effect of Teen-age Childbearing on Birth Outcomes in a Dynamic Family Context. Econometrica 1995; 63; 303-26.

Royer, H. Separated at Girth, US Twin Estimates of the Effects of Birth Weight. American Economic Journal: Applied Economics 2009; 1(1); 49-85.

Rutter, M., Tizard, J., Whitmore, K. Education, Health and Behaviour. London, Longmans; 1970.

Sammons, P., Sylva, K., Melhuish, E., Siraj-Blatchford, I., Taggart, B, Elliott, K., Marsh, A. The Effective Provision of Pre-school Education (EPPE) Project. Technical Paper 11: The continuing effect of pre-school education at age 7 years. Institute of Education: London; 2004.

Todd, P.E. and Wolpin, K.I. On the Specification and Estimation of the Production Function for Cognitive Achievement, Economic Journal 2003; 113, F3-F33.

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

Table 1
Descriptive Statistics

Birth and age 9 months			*Age 3*			*Age 5*		
	Boys	*Girls*		*Boys*	*Girls*		*Boys*	*Girls*
Birth weight (kg)	3.486	3.366	SDQ	30.651	31.501	SDQ	32.823	33.759
	(0.562)	(0.544)		(4.967)	(4.833)		(4.780)	(4.395)
Pretem	0.055	0.050	Self-Regulation	21.743	22.344	Self-Regulation	35.798	36.865
Firstborn	0.442	0.419		(2.734)	(2.628)		(4.036)	(3.722)
Breathing problems at birth	0.051	0.038	Pianta	34.345	34.885	BAS Nam. Voc.	110.938	111.545
Delivery: normal	0.660	0.692		(6.889)	(6.649)		(13.832)	(13.185)
Delivery: pl. caesarian	0.088	0.096	BAS Nam. Voc.	74.859	79.036	BAS Pic. Sim.	82.408	83.581
Delivery: em. caesarian	0.137	0.110		(15.668)	(14.667)		(11.142)	(11.126)
Delivery: assisted	0.115	0.102	SR Composite	104.269	108.083	BAS Pat. Con.	88.352	90.807
Breastfeeding: none	0.300	0.308		(15.822)	(14.317)		(18.753)	(17.227)
Breastfeeding: 0-2m	0.306	0.285	Age in days	1141	1142	Age in days	1903	1903
Breastfeeding: 2-4m	0.090	0.087		(67)	(70)		(90)	(90)
Breastfeeding: >4 part.	0.158	0.144	Mother: no smoke	0.704	0.709	Mother: no smoke	0.719	0.727
Breastfeeding: >4 excl.	0.146	0.176	Mother: smokes <10	0.165	0.167	Mother: smokes <10	0.155	0.156
Mother's age	29.398	29.511	Mother: smokes >10	0.131	0.124	Mother: smokes >10	0.127	0.116
	(5.684)	(5.692)	Mother: no drink	0.514	0.506	Mother: no drink	0.514	0.488
Mother's Postnatal	50.658	50.953	Mother: drink mod.	0.414	0.429	Mother: drink mod.	0.414	0.438
Attachment Score	(4.662)	(4.619)	Mother: drink freq.	0.071	0.065	Mother: drink freq.	0.072	0.075
Mother's Malaise	1.562	1.547	BMI: underweight	0.020	0.025	BMI: underweight	0.016	0.020
Inventory Score	1.665	1.685	BMI: normal	0.494	0.494	BMI: normal	0.481	0.487
Income: <10.4K	0.185	0.183	BMI: overweight	0.188	0.197	BMI: overweight	0.187	0.189
Income: 10.4-20.8K	0.327	0.322	BMI: obese	0.126	0.122	BMI: obese	0.129	0.130
Income: 20.8-31.2K	0.235	0.244	BMI: pregnant	0.075	0.070	BMI: pregnant	0.055	0.045
Income: >31.2K	0.253	0.251	BMI: missing	0.097	0.092	BMI: missing	0.131	0.130
Mother: no qualifications	0.105	0.105	Immunizations: all	0.947	0.952	Immunizations: all	0.920	0.924
Mother: less than O levels	0.110	0.108	Immunizations: some	0.032	0.026	Immunizations: some	0.056	0.055
Mother: O levels	0.367	0.372	Immunizations: miss.	0.020	0.022	Immunizations: miss.	0.024	0.021
Mother: A levels	0.110	0.111	Mother's Kessler	3.136	3.096	Mother's Kessler	2.978	2.979
Mother: diplomas in HE	0.113	0.103	6-item	(3.596)	(3.504)	6-item	(3.611)	(3.628)
Mother: degree or higher	0.196	0.200	Number of siblings	1.103	1.114	Number of siblings	1.293	1.294
Tenure: owner	0.692	0.688		(0.985)	(0.984)		(0.982)	(0.965)
Tenure: rent privately	0.072	0.071						
Tenure: rent from council	0.148	0.141						
Tenure: rent from h. ass.	0.045	0.056						
Tenure: other	0.044	0.044						
Both partners work	0.508	0.516						
Main work, partner no work	0.020	0.022						
Main no work, partner work	0.300	0.285						
Both no work	0.045	0.049						
No partner	0.127	0.128						

Notes: The number of observations is 3,826 for boys and 3,831 for girls. Means and standard deviations of the total scores on the various tests and measures of behavioural development presented here. Standard deviations shown in parenthesis only for continuous variables.

Table 2
F-test on exclusion restrictions

	Basic specification	*Maternal health and health behaviour*
Boys		
SDQ (age 3)	10.34	6.92
Self-Regulation (age 3)	5.01	3.95
Pianta (age 3)	3.64	2.42
BAS Nam. Voc. (age 3)	7.68	5.15
SR Composite (age 3)	12.96	7.85
SDQ (age 5)	8.07	4.53
Self-Regulation (age 5)	4.55	2.95
BAS Nam. Voc. (age 5)	9.28	7.89
BAS Pic. Sim. (age 5)	4.29	4.00
BAS Pat. Con. (age 5)	4.38	3.90
Girls		
SDQ (age 3)	11.14	7.15
Self-Regulation (age 3)	6.25	4.59
Pianta (age 3)	5.05	3.24
BAS Nam. Voc. (age 3)	6.86	4.94
SR Composite (age 3)	18.57	11.35
SDQ (age 5)	9.07	5.49
Self-Regulation (age 5)	5.89	3.67
BAS Nam. Voc. (age 5)	9.53	8.10
BAS Pic. Sim. (age 5)	5.19	4.97
BAS Pat. Con. (age 5)	5.48	4.13

Notes: The number of observations is 3,826 for boys and 3,831 for girls. F-test on exclusion restrictions (28 degrees of freedom) shown. The specifications include all the explanatory variables shown in tables 3-6, with the exception that the regressions with outcomes at age 5 do not include outcomes at age 3 among the explanatory variables. The excluded variables are: squared root of the number of siblings, mother's education (5 dummies), family income (3 dummies), tenure of the household accommodation (4 dummies), region (11 dummies), and combined employment and partnership status of the mother (4 dummies).

Table 3
Impact of birth weight on behavioural and cognitive scores, boys, age 3

	SDQ		*Self-Regulation*		*Pianta*		*BAS Nam. Voc.*	
	OLS	3SLS	OLS	3SLS	OLS	3SLS	OLS	3SLS
Panel A: Basic specification								
Birth weight	0.057	-0.001	0.051	0.016	0.052	0.026	0.156**	0.105**
	(0.033)	(0.036)	(0.033)	(0.033)	(0.034)	(0.033)	(0.031)	(0.028)
SR Composite		0.753**		0.451**		0.339**		0.661**
		(0.064)		(0.059)		(0.059)		(0.050)
Panel B: Maternal health and health behaviour								
Birth weight	0.008	-0.024	0.026	0.006	0.006	-0.008	0.121**	0.090**
	(0.031)	(0.033)	(0.032)	(0.032)	(0.032)	(0.030)	(0.031)	(0.028)
Breastfeeding: 0-2m	0.012	-0.056	-0.024	-0.067	0.043	0.015	0.105**	0.039
	(0.040)	(0.043)	(0.041)	(0.042)	(0.040)	(0.039)	(0.040)	(0.036)
Breastfeeding: 2-4m	0.133*	-0.005	-0.032	-0.120	-0.004	-0.062	0.130*	-0.004
	(0.055)	(0.064)	(0.060)	(0.063)	(0.057)	(0.059)	(0.060)	(0.054)
Breastfeeding: >4m part.	0.233**	0.053	0.128*	0.012	0.086	0.010	0.191**	0.015
	(0.047)	(0.057)	(0.051)	(0.055)	(0.047)	(0.052)	(0.050)	(0.048)
Breastfeeding: >4m excl.	0.234**	0.038	0.105*	-0.020	0.085	0.003	0.250**	0.060
	(0.048)	(0.059)	(0.053)	(0.057)	(0.049)	(0.054)	(0.052)	(0.050)
Smokes: <10 cigs.	-0.093*	0.005	-0.082	-0.020	-0.118**	-0.077	-0.123**	-0.028
	(0.043)	(0.048)	(0.044)	(0.046)	(0.043)	(0.043)	(0.044)	(0.040)
Smokes: >10 cigs.	-0.227**	-0.101	-0.160**	-0.080	-0.196**	-0.143**	-0.231**	-0.109*
	(0.051)	(0.054)	(0.051)	(0.053)	(0.052)	(0.049)	(0.050)	(0.046)
Drinks: moderately	0.015	-0.016	0.013	-0.007	0.008	-0.005	-0.016	-0.046
	(0.031)	(0.035)	(0.033)	(0.034)	(0.031)	(0.032)	(0.032)	(0.029)
Drinks: frequently	0.095	-0.010	0.031	-0.036	0.066	0.022	0.038	-0.065
	(0.053)	(0.068)	(0.065)	(0.066)	(0.059)	(0.062)	(0.063)	(0.057)
BMI: underweight	0.027	0.111	0.131	0.184	0.040	0.075	-0.110	-0.029
	(0.117)	(0.117)	(0.111)	(0.114)	(0.127)	(0.107)	(0.113)	(0.099)
BMI: overweight	-0.079*	0.033	-0.076	-0.004	-0.054	-0.007	-0.153**	-0.045
	(0.040)	(0.046)	(0.043)	(0.044)	(0.039)	(0.042)	(0.041)	(0.038)
BMI: obese	-0.150**	-0.032	-0.171**	-0.096	-0.055	-0.006	-0.076	0.038
	(0.047)	(0.054)	(0.050)	(0.052)	(0.048)	(0.049)	(0.048)	(0.045)
Immunizations: some	-0.103	-0.130	-0.094	-0.111	-0.052	-0.063	0.021	-0.005
	(0.099)	(0.091)	(0.090)	(0.089)	(0.099)	(0.083)	(0.084)	(0.077)
Postnatal Attachment	0.018**	0.023**	0.022**	0.025**	0.041**	0.043**	-0.005	0.000
	(0.003)	(0.004)	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
Malaise Inventory	-0.062**	-0.064**	-0.019	-0.020	-0.043**	-0.044**	-0.002	-0.004
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.010)	(0.011)	(0.009)
Kessler 6-item	-0.060**	-0.053**	-0.044**	-0.039**	-0.075**	-0.072**	-0.002	0.005
	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)	(0.005)	(0.005)	(0.004)
SR Composite		0.651**		0.416**		0.273**		0.632**
		(0.073)		(0.071)		(0.067)		(0.062)

Notes: Coefficients from OLS and 3SLS reported. Number of children is 3,826. The specifications also include: maternal age and its square, a dummy for preterm, a dummy for firstborn, a dummy for breathing problems at birth, dummies for type of delivery (planned caesarean, emergency caesarean, assisted), the log of age in days at the time of the second interview, a dummy if the mother is pregnant, a dummy if BMI is missing, and a dummy if immunization information is missing. *Symbols:* **= significant at the 1% level; *=significant at the 5% level.

Table 4

Impact of birth weight on behavioural and cognitive scores, girls, age 3

	SDQ		*Self-Regulation*		*Pianta*		*BAS Nam. Voc.*	
	OLS	3SLS	OLS	3SLS	OLS	3SLS	OLS	3SLS
Panel A: Basic specification:								
Birth weight	0.212**	0.131**	0.128**	0.076*	0.086**	0.052	0.218**	0.153**
	(0.030)	(0.035)	(0.032)	(0.033)	(0.031)	(0.033)	(0.033)	(0.028)
SR Composite		0.694**		0.447**		0.293**		0.552**
		(0.051)		(0.049)		(0.049)		(0.041)
Panel B: Maternal health and health behaviour								
Birth weight	0.153**	0.110**	0.076*	0.051	0.035	0.022	0.189**	0.148**
	(0.029)	(0.032)	(0.032)	(0.033)	(0.030)	(0.030)	(0.034)	(0.029)
Breastfeeding: 0-2m	0.127**	0.024	0.071	0.008	0.039	0.007	0.114**	0.011
	(0.041)	(0.043)	(0.042)	(0.043)	(0.041)	(0.040)	(0.040)	(0.038)
Breastfeeding: 2-4m	0.155**	-0.056	0.064	-0.063	-0.048	-0.114	0.098	-0.111
	(0.054)	(0.064)	(0.058)	(0.065)	(0.056)	(0.060)	(0.062)	(0.057)
Breastfeeding: >4m part.	0.239**	0.038	0.157**	0.036	0.089	0.026	0.262**	0.063
	(0.047)	(0.056)	(0.051)	(0.056)	(0.048)	(0.053)	(0.050)	(0.049)
Breastfeeding: >4m excl.	0.273**	0.087	0.208**	0.095	0.119*	0.061	0.160**	-0.024
	(0.045)	(0.054)	(0.050)	(0.054)	(0.046)	(0.050)	(0.049)	(0.047)
Smokes: <10 cigs.	-0.057	0.001	-0.127**	-0.092*	-0.031	-0.013	-0.061	-0.004
	(0.044)	(0.045)	(0.046)	(0.045)	(0.043)	(0.042)	(0.043)	(0.039)
Smokes: >10 cigs.	-0.301**	-0.136*	-0.263**	-0.163**	-0.248**	-0.197**	-0.190**	-0.027
	(0.053)	(0.054)	(0.050)	(0.054)	(0.051)	(0.050)	(0.050)	(0.047)
Drinks: moderately	0.047	-0.003	-0.003	-0.033	-0.014	-0.030	0.028	-0.021
	(0.031)	(0.034)	(0.033)	(0.034)	(0.032)	(0.032)	(0.032)	(0.030)
Drinks: frequently	0.103	0.033	0.060	0.018	-0.149*	-0.171**	0.011	-0.057
	(0.056)	(0.067)	(0.069)	(0.068)	(0.062)	(0.063)	(0.063)	(0.059)
BMI: underweight	0.055	0.041	-0.044	-0.053	-0.002	-0.007	0.053	0.039
	(0.103)	(0.100)	(0.102)	(0.101)	(0.095)	(0.094)	(0.112)	(0.088)
BMI: overweight	-0.036	-0.002	0.015	0.036	-0.007	0.004	-0.005	0.029
	(0.038)	(0.042)	(0.040)	(0.042)	(0.039)	(0.039)	(0.040)	(0.037)
BMI: obese	-0.102*	0.020	-0.048	0.026	-0.017	0.021	-0.057	0.063
	(0.052)	(0.053)	(0.051)	(0.053)	(0.050)	(0.049)	(0.050)	(0.046)
Immunization: some	-0.081	-0.042	0.129	0.153	0.022	0.034	0.114	0.152
	(0.100)	(0.098)	(0.099)	(0.099)	(0.094)	(0.092)	(0.109)	(0.086)
Postnatal Attachment	0.020**	0.025**	0.022**	0.025**	0.039**	0.040**	0.001	0.005
	(0.003)	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.004)	(0.003)
Malaise Inventory	-0.064**	-0.062**	-0.037**	-0.036**	-0.057**	-0.056**	0.010	0.011
	(0.011)	(0.011)	(0.010)	(0.011)	(0.011)	(0.010)	(0.011)	(0.009)
Kessler 6-item	-0.058**	-0.043**	-0.029**	-0.020**	-0.067**	-0.062**	-0.019**	-0.005
	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
SR Composite		0.568**		0.344**		0.178**		0.562**
		(0.060)		(0.061)		(0.056)		(0.053)

Notes: Coefficients from OLS and 3SLS reported. Number of children is 3,831. The specifications also include: maternal age and its square, a dummy for preterm, a dummy for firstborn, a dummy for breathing problems at birth, dummies for type of delivery (planned caesarean, emergency caesarean, assisted), the log of age in days at the time of the second interview, a dummy if the mother is pregnant, a dummy if BMI is missing, and a dummy if immunization information is missing. *Symbols:* **= significant at the 1% level; *=significant at the 5% level.

Table 5
Impact of birth weight on behavioural and cognitive scores, boys, age 5

	SDQ		*Self-Regulation*		*BAS Nam. Voc.*		*BAS Pic. Sim.*		*BAS Pat. Con.*	
	OLS	3SLS	OLS	3SLS	OLS	3SLS	OLS	3SLS	OLS	3SLS
Panel A: Basic specification										
Birth weight	0.045	0.044	0.029	0.029	0.041	0.040	0.059	0.059	0.142**	0.142**
	(0.026)	(0.027)	(0.026)	(0.027)	(0.028)	(0.029)	(0.031)	(0.032)	(0.031)	(0.031)
SDQ (age 3)	0.478**	0.447**	0.265**	0.251**	0.034	-0.011	0.069**	0.059*	0.067**	0.035
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.021)	(0.022)	(0.024)	(0.023)	(0.023)
Self-Regulation (age 3)	0.065**	0.038*	0.285**	0.273**	0.024	-0.016	0.017	0.008	0.060**	0.031
	(0.017)	(0.018)	(0.017)	(0.018)	(0.018)	(0.019)	(0.020)	(0.021)	(0.020)	(0.020)
Pianta (age 3)	0.117**	0.123**	0.117**	0.120**	-0.015	-0.006	-0.010	-0.008	-0.011	-0.004
	(0.018)	(0.017)	(0.018)	(0.017)	(0.018)	(0.018)	(0.021)	(0.020)	(0.022)	(0.020)
BAS Nam. Voc. (age 3)	0.039**	-0.137**	0.045**	-0.033	0.403**	0.148**	0.139**	0.085*	0.180**	-0.002
	(0.013)	(0.035)	(0.014)	(0.035)	(0.015)	(0.037)	(0.017)	(0.041)	(0.016)	(0.040)
SR Composite (age 3)		0.386**		0.171*		0.561**		0.118		0.401**
		(0.072)		(0.071)		(0.076)		(0.084)		(0.082)
Panel B: Maternal health and health behaviour										
Birth weight	0.019	0.018	0.013	0.012	0.035	0.032	0.058	0.053	0.131**	0.132**
	(0.026)	(0.026)	(0.027)	(0.027)	(0.028)	(0.028)	(0.031)	(0.033)	(0.031)	(0.031)
Smokes: <10 cigs.	-0.086*	-0.041	0.028	0.068	-0.060	-0.028	-0.017	0.038	-0.081	-0.124*
	(0.038)	(0.047)	(0.039)	(0.049)	(0.039)	(0.052)	(0.049)	(0.060)	(0.048)	(0.057)
Smokes: >10 cigs.	-0.176**	-0.071	-0.136**	-0.074	-0.112*	-0.027	-0.034	0.023	-0.103*	-0.144*
	(0.044)	(0.060)	(0.045)	(0.062)	(0.046)	(0.066)	(0.052)	(0.075)	(0.051)	(0.073)
Drinks: moderately	0.032	0.022	-0.038	-0.030	0.008	0.005	-0.008	-0.032	0.029	0.047
	(0.026)	(0.032)	(0.028)	(0.033)	(0.030)	(0.035)	(0.034)	(0.040)	(0.033)	(0.039)
Drinks: frequently	-0.001	-0.043	-0.062	-0.040	0.138*	0.006	0.022	-0.020	-0.017	-0.078
	(0.052)	(0.062)	(0.055)	(0.064)	(0.056)	(0.068)	(0.060)	(0.077)	(0.066)	(0.075)
BMI: underweight	-0.159	-0.360**	0.004	-0.152	-0.005	-0.072	0.229*	0.157	-0.003	-0.198
	(0.110)	(0.116)	(0.101)	(0.120)	(0.113)	(0.127)	(0.107)	(0.145)	(0.121)	(0.140)
BMI: overweight	-0.090**	-0.020	-0.014	0.029	-0.050	-0.020	-0.028	0.008	0.004	-0.013
	(0.035)	(0.040)	(0.036)	(0.042)	(0.039)	(0.044)	(0.045)	(0.050)	(0.042)	(0.048)
BMI: obese	-0.104**	-0.011	-0.032	0.015	-0.067	-0.054	-0.022	-0.005	-0.032	-0.083
	(0.040)	(0.056)	(0.043)	(0.058)	(0.046)	(0.061)	(0.050)	(0.070)	(0.048)	(0.067)
Immunization: some	-0.116*	-0.137*	-0.056	-0.069	-0.037	-0.060	-0.085	-0.102	-0.063	-0.087
	(0.055)	(0.056)	(0.059)	(0.059)	(0.060)	(0.062)	(0.077)	(0.071)	(0.069)	(0.068)
Kessler 6-item	-0.044**	-0.038**	-0.027**	-0.027**	0.002	0.004	-0.003	-0.005	-0.001	-0.000
	(0.005)	(0.004)	(0.004)	(0.005)	(0.004)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
SDQ (age 3)	0.452**	0.427**	0.253**	0.246**	0.028	-0.010	0.064**	0.062**	0.061**	0.035
	(0.020)	(0.019)	(0.020)	(0.020)	(0.020)	(0.021)	(0.022)	(0.024)	(0.023)	(0.023)
Self-Regulation (age 3)	0.062**	0.043*	0.283**	0.268**	0.024	-0.005	0.016	0.021	0.059**	0.035
	(0.016)	(0.017)	(0.017)	(0.018)	(0.018)	(0.019)	(0.020)	(0.021)	(0.020)	(0.020)
Pianta (age 3)	0.081**	0.072**	0.096**	0.085**	-0.014	0.002	-0.011	-0.002	-0.013	-0.001
	(0.018)	(0.017)	(0.018)	(0.018)	(0.018)	(0.019)	(0.021)	(0.021)	(0.023)	(0.020)
BAS Nam. Voc. (age 3)	0.034*	-0.082*	0.045**	-0.027	0.397**	0.193**	0.137**	0.144**	0.177**	0.018
	(0.013)	(0.037)	(0.014)	(0.039)	(0.015)	(0.041)	(0.017)	(0.047)	(0.017)	(0.045)
SR Composite (age 3)		0.266**		0.163*		0.450**		-0.030		0.358**
		(0.079)		(0.083)		(0.087)		(0.100)		(0.096)

Notes: Number of children is 3,826. The specifications also include: maternal age and its square, a dummy for preterm, a dummy for firstborn, a dummy for breathing problems at birth, dummies for type of delivery (planned caesarean, emergency caesarean, assisted), the log of age in days at the time of the second interview, a dummy if the mother is pregnant, a dummy if BMI is missing, and a dummy if immunization information is missing. The 3SLS specifications also include: mother's smoking, drinking and BMI at age 3, immunizations received by age 3, maternal Postnatal Attachment score at 9 months, maternal Malaise Inventory score at 9 months, and maternal 6-item Kessler score at age 3. *Symbols:* **= significant at the 1% level; *=significant at the 5% level.

Table 6
Impact of birth weight on behavioural and cognitive scores, girls, age 5

	SDQ		*Self-Regulation*		*BAS Nam. Voc.*		*BAS Pic. Sim.*		*BAS Pat. Con.*	
	OLS	3SLS	OLS	3SLS	OLS	3SLS	OLS	3SLS	OLS	3SLS
Panel A: Basic specification										
Birth weight	0.055*	0.058*	0.044	0.046	-0.001	0.002	0.078*	0.078*	0.110**	0.113**
	(0.026)	(0.027)	(0.028)	(0.027)	(0.028)	(0.028)	(0.033)	(0.032)	(0.032)	(0.032)
SDQ (age 3)	0.476**	0.429**	0.242**	0.219**	0.037	-0.017	0.061**	0.044	0.043*	-0.017
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.021)	(0.023)	(0.023)	(0.022)	(0.024)
Self-Regulation (age 3)	0.095**	0.086**	0.287**	0.283**	0.048**	0.037*	0.021	0.018	0.077**	0.065**
	(0.016)	(0.016)	(0.017)	(0.017)	(0.018)	(0.017)	(0.020)	(0.019)	(0.020)	(0.020)
Pianta (age 3)	0.116**	0.116**	0.152**	0.152**	-0.009	-0.009	0.004	0.004	-0.008	-0.008
	(0.017)	(0.016)	(0.017)	(0.017)	(0.018)	(0.017)	(0.020)	(0.019)	(0.020)	(0.020)
BAS Nam. Voc. (age 3)	0.043**	-0.121**	0.040**	-0.039	0.383**	0.196**	0.139**	0.082*	0.158**	-0.053
	(0.013)	(0.028)	(0.014)	(0.028)	(0.016)	(0.029)	(0.017)	(0.033)	(0.017)	(0.034)
SR Composite (age 3)		0.380**		0.183**		0.430**		0.130		0.487**
		(0.056)		(0.057)		(0.059)		(0.067)		(0.068)
Panel B: Maternal health and health behaviour										
Birth weight	0.045	0.043	0.040	0.033	-0.015	-0.012	0.077*	0.074*	0.098**	0.105**
	(0.026)	(0.026)	(0.028)	(0.027)	(0.029)	(0.029)	(0.034)	(0.033)	(0.032)	(0.033)
Smokes: <10 cigs.	-0.078*	-0.092	-0.035	0.022	-0.062	0.024	-0.002	0.086	-0.043	0.055
	(0.037)	(0.048)	(0.038)	(0.050)	(0.042)	(0.053)	(0.047)	(0.060)	(0.048)	(0.060)
Smokes: >10 cigs.	-0.102*	-0.100	-0.043	0.049	-0.107*	0.001	-0.018	0.105	-0.197**	-0.120
	(0.049)	(0.061)	(0.048)	(0.063)	(0.047)	(0.067)	(0.054)	(0.076)	(0.054)	(0.076)
Drinks: moderately	0.106**	0.090**	0.033	0.009	0.040	0.023	0.022	0.015	0.089**	0.090*
	(0.026)	(0.034)	(0.028)	(0.035)	(0.030)	(0.037)	(0.034)	(0.042)	(0.033)	(0.042)
Drinks: frequently	0.050	0.001	-0.056	-0.100	0.070	-0.008	0.081	0.073	0.059	0.036
	(0.048)	(0.063)	(0.053)	(0.065)	(0.057)	(0.068)	(0.056)	(0.078)	(0.057)	(0.078)
BMI: underweight	0.084	0.268*	0.119	0.194	0.064	0.075	0.212	0.376**	-0.019	0.016
	(0.103)	(0.110)	(0.099)	(0.114)	(0.094)	(0.120)	(0.132)	(0.136)	(0.099)	(0.136)
BMI: overweight	-0.052	-0.008	-0.033	-0.018	0.020	0.091*	0.022	0.007	-0.112*	-0.025
	(0.033)	(0.042)	(0.036)	(0.043)	(0.038)	(0.045)	(0.043)	(0.052)	(0.044)	(0.052)
BMI: obese	-0.061	0.002	-0.007	0.008	0.028	0.089	0.035	0.059	-0.089	0.008
	(0.039)	(0.055)	(0.041)	(0.057)	(0.047)	(0.060)	(0.050)	(0.068)	(0.050)	(0.068)
Immunization: some	-0.051	-0.061	-0.004	-0.011	0.084	0.074	-0.123	-0.128	-0.092	-0.108
	(0.060)	(0.057)	(0.059)	(0.059)	(0.061)	(0.062)	(0.069)	(0.070)	(0.067)	(0.070)
Kessler 6-item	-0.046**	-0.045**	-0.033**	-0.033**	-0.014**	-0.012*	-0.011*	-0.018**	-0.011*	-0.010
	(0.004)	(0.004)	(0.005)	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
SDQ (age 3)	0.438**	0.410**	0.219**	0.207**	0.021	-0.015	0.050*	0.050*	0.024	-0.017
	(0.020)	(0.019)	(0.020)	(0.020)	(0.020)	(0.021)	(0.023)	(0.024)	(0.022)	(0.024)
Self-Regulation (age 3)	0.097**	0.088**	0.289**	0.283**	0.048**	0.039*	0.023	0.023	0.077**	0.064**
	(0.015)	(0.016)	(0.017)	(0.016)	(0.018)	(0.017)	(0.020)	(0.020)	(0.021)	(0.020)
Pianta (age 3)	0.088**	0.091**	0.130**	0.125**	-0.017	-0.013	-0.001	0.016	-0.013	-0.007
	(0.016)	(0.016)	(0.017)	(0.017)	(0.018)	(0.018)	(0.020)	(0.020)	(0.020)	(0.020)
BAS Nam. Voc. (age 3)	0.043**	-0.087**	0.042**	-0.025	0.382**	0.233**	0.138**	0.129**	0.153**	-0.027
	(0.013)	(0.031)	(0.013)	(0.032)	(0.016)	(0.034)	(0.018)	(0.038)	(0.017)	(0.038)
SR Composite (age 3)		0.313**		0.162*		0.347**		0.012		0.426**
		(0.066)		(0.069)		(0.072)		(0.082)		(0.082)

Notes: Number of children is 3,831. The specifications also include: maternal age and its square, a dummy for preterm, a dummy for firstborn, a dummy for breathing problems at birth, dummies for type of delivery (planned caesarean, emergency caesarean, assisted), the log of age in days at the time of the second interview, a dummy if the mother is pregnant, a dummy if BMI is missing, and a dummy if immunization information is missing. The 3SLS specifications also include: mother's smoking, drinking and BMI at age 3, immunizations received by age 3, maternal Postnatal Attachment score at 9 months, maternal Malaise Inventory score at 9 months, and maternal 6-item Kessler score at age 3. *Symbols:* **= significant at the 1% level; *=significant at the 5% level.

Table 7
Robustness checks, age 3

	SDQ		*Self-Regulation*		*Pianta*		*BAS Nam. Voc.*	
	OLS	3SLS	OLS	3SLS	OLS	3SLS	OLS	3SLS
				Boys				
Panel A: main results								
Birth weight	0.008	-0.024	0.026	0.006	0.006	-0.008	0.121**	0.090**
	(0.031)	(0.033)	(0.032)	(0.032)	(0.032)	(0.030)	(0.031)	(0.028)
SR Composite (age 3)		0.651**		0.416**		0.273**		0.632**
		(0.073)		(0.071)		(0.067)		(0.062)
Panel B: parental educational activities								
Birth weight	0.001	-0.030	0.025	0.005	0.005	-0.006	0.124**	0.090**
	(0.031)	(0.032)	(0.032)	(0.032)	(0.032)	(0.030)	(0.031)	(0.028)
SR Composite (age 3)		0.558**		0.358**		0.207**		0.609**
		(0.078)		(0.078)		(0.072)		(0.068)
Panel C: child care								
Birth weight	0.005	-0.024	0.024	0.005	0.004	-0.008	0.119**	0.090**
	(0.031)	(0.033)	(0.032)	(0.032)	(0.032)	(0.030)	(0.031)	(0.028)
SR Composite (age 3)		0.621**		0.394**		0.254**		0.625**
		(0.075)		(0.073)		(0.069)		(0.064)
Panel D: all families present at age 3								
Birth weight	0.020	-0.017	0.018	-0.007	0.002	-0.013	0.130**	0.093**
	(0.029)	(0.031)	(0.030)	(0.031)	(0.030)	(0.028)	(0.029)	(0.026)
SR Composite (age 3)		0.660**		0.441**		0.268**		0.666**
		(0.070)		(0.068)		(0.064)		(0.059)
				Girls				
Panel A: main results								
Birth weight	0.153**	0.110**	0.076*	0.051	0.035	0.022	0.189**	0.148**
	(0.029)	(0.032)	(0.032)	(0.033)	(0.030)	(0.030)	(0.034)	(0.029)
SR Composite (age 3)		0.568**		0.344**		0.178**		0.562**
		(0.060)		(0.061)		(0.056)		(0.053)
Panel B: parental educational activities								
Birth weight	0.141**	0.107**	0.074*	0.055	0.027	0.017	0.181**	0.145**
	(0.029)	(0.032)	(0.032)	(0.032)	(0.030)	(0.030)	(0.034)	(0.029)
SR Composite (age 3)		0.503**		0.271**		0.136*		0.528**
		(0.066)		(0.067)		(0.063)		(0.059)
Panel C: child care								
Birth weight	0.150**	0.111**	0.075*	0.051	0.033	0.021	0.188**	0.148**
	(0.029)	(0.032)	(0.032)	(0.033)	(0.030)	(0.030)	(0.034)	(0.028)
SR Composite (age 3)		0.559**		0.346**		0.171**		0.580**
		(0.063)		(0.063)		(0.059)		(0.055)
Panel D: all families present at age 3								
Birth weight	0.140**	0.107**	0.065*	0.041	0.026	0.014	0.170**	0.130**
	(0.027)	(0.030)	(0.030)	(0.031)	(0.028)	(0.028)	(0.032)	(0.027)
SR Composite (age 3)		0.518**		0.361**		0.188**		0.617**
		(0.052)		(0.054)		(0.050)		(0.048)

Notes: Number of children is 3,826 for boys and 3,831 for girls in Panels A and C; 3,773 for boys and 3,800 for girls in Panel B; 4,353 for boys and 4,378 for girls in Panel D. All specifications include all the variables shown in Panel B of Tables 3 and 4, as well as: maternal age and its square, a dummy for preterm, a dummy for firstborn, a dummy for breathing problems at birth, dummies for type of delivery (planned caesarean, emergency caesarean, assisted), the log of age in days at the time of the second interview, a dummy if the mother is pregnant, a dummy if BMI is missing, and a dummy if immunization information is missing. *Symbols:* **= significant at the 1% level; *=significant at the 5% level.

Table 8
OLS and FE estimates on a sample of twins

	Denver Development Test Score		*Carey Infant Temperament Scale*		*Neonatal care*	
	OLS	FE	OLS	FE	OLS	FE
All twin-pairs						
Birth weight	0.622**	0.117	-0.195	-0.020	-0.210**	-0.074
	(0.130)	(0.167)	(0.125)	(0.177)	(0.049)	(0.070)
Observations	486	486	458	458	486	486
All twin-pairs > 1500 grams						
Birth weight	0.367*	0.153	-0.182	0.069	-0.164**	-0.074
	(0.150)	(0.173)	(0.163)	(0.191)	(0.059)	(0.077)
Observations	430	430	406	406	430	430

Notes: Coefficients from OLS and FE regression reported. For the pooled OLS, standard errors are clustered at the family level. The measures of the Denver Development Test and the Carey Infant Temperament Scale are standardized using the entire sample (i.e. including singletons). The OLS model also includes: maternal age and its square, a dummy for preterm, a dummy for firstborn, a dummy for ethnic group white, the log of age in days at the time of the first interview, and a male dummy. *Symbols:* **= significant at the 1% level; *=significant at the 5% level; $^{+}$=significant at the 10% level.