

The Effect of Breastfeeding on Children's Cognitive and Noncognitive Development

Cristina Borra

University of Seville

Maria Iacovou

University of Essex

Almudena Sevilla

*University of Oxford
and IZA*

Discussion Paper No. 6697
June 2012

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

The Effect of Breastfeeding on Children's Cognitive and Noncognitive Development

This paper uses propensity score matching methods to investigate the relationship between breastfeeding and children's cognitive and noncognitive development. We find that breastfeeding for four weeks is positively and statistically significantly associated with higher cognitive test scores, by around one tenth of a standard deviation. The association between breastfeeding and noncognitive development is weaker, and is restricted to children of less educated mothers. We conclude that interventions which increase breastfeeding rates would improve not only children's health, but also their cognitive skills, and possibly also their noncognitive development.

JEL Classification: I10, J0

Keywords: breastfeeding

Corresponding author:

Almudena Sevilla
School of Business and Management
Queen Mary, University of London
Francis Bancroft Building
Mile End Road
London E1 4NS
United Kingdom
E-mail: a.sevilla@qmul.ac.uk

The Effect of Breastfeeding on Children's Cognitive and Noncognitive Development*

The most valuable of all capital is that invested in human beings; and of that capital the most precious part is the result of the care and influence of the mother.

Alfred Marshall (1890), Paragraph VI.IV.11.

1 Introduction

This paper examines the relationship between breastfeeding and children's later cognitive and noncognitive outcomes. This is a topic of considerable importance for policy in the UK: the World Health Organization recommends breastfeeding exclusively for six months and alongside solid foods for two years, but in the UK, barely one in three infants is exclusively breastfed during the first four months of life. Given the increasing recognition of the importance of very early interventions in children's development and later outcomes; and given the huge social gradient in breastfeeding rates, with the most privileged mothers currently being many times more likely to breastfeed than the least privileged mothers, breastfeeding may well be a significant route for the intergenerational transmission of human capital. Recent research shows a significant impact of behavioural and psycho-social outcomes on earnings and education (Duncan and Dunifon 1998; Heckman et al. 2006; Mueller and Plug 2006). Differences in children's cognitive development emerge at early ages (Illsey 2002; Feinstein 2003; Cunha et al. 2010), and the importance of timely parental investments (pre-natal as well as post-natal) is increasingly recognized as a major factor in fostering child development (Carneiro and Heckman 2003; Del Bono et al. 2008). A fuller understanding of the relationship between breastfeeding and various aspects of child development is therefore crucial for an understanding of the intergenerational transmission of inequality, and for policy-making aimed at reducing inequality.

There is a well-established association between breastfeeding and a range of positive health outcomes in children, such as a lower incidence of asthma and middle ear and urinary tract infections (Dyson et al., 2006). A smaller body of research also shows breastfeeding to be related to better gross motor development (Sacker et al., 2006), and improved cognitive ability (Anderson et al., 1999). Other potential effects of breastfeeding, such as cognitive and noncognitive outcomes of the type investigated here, are much less well

* This paper has benefited from comments provided by participants at the British Society of Population Studies and at the 5th conference of Epidemiological Longitudinal Studies in Europe; from our colleagues at ISER, particularly Emilia del Bono and Birgitta Rabe. This work was funded by the ESRC under grant RES-062-23-1693. We are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. The UK Medical Research Council (Grant Ref: 74882); the Wellcome Trust (Grant Ref: 076467) and the University of Bristol provide core support for ALSPAC. This publication is the work of the authors and they will serve as guarantors for the contents of this paper.

researched. Despite the growing literature in this area, scientists' understanding of the mechanisms behind these observed relationships remains incomplete. One theory is that, several components of breast milk (notably long-chain polyunsaturated fatty acids, which accumulate in the brain and retina) may affect cognitive development via their effects on neural development. (Innis, 2004; Petryk et al., 2007). In the case of the non-cognitive outcomes discussed in this paper, the effect of sensory stimuli on the development of the nervous system may be crucial. It has been suggested that the skin-to-skin contact between mother and infant inherent in breastfeeding forms an important stimulus to the child (Britton et al., 2006). Breastfeeding also releases hormones in the mother, which are believed to be related to nurturing behaviour (Petryk et al., 2007). Recently, research in the emerging field of epigenetics shows that breastfeeding induces genetic pathways related to responses to stress (Weaver et al., 2004) or intestinal health (Chapkin et al., 2010) which are quite different from those in non-breastfed individuals.

One problem which arises in considering the relationship between breastfeeding and later outcomes is the difficulty of identifying whether the observed relationships are causal, as opposed to arising because breastfeeding is more likely to be practiced by mothers whose characteristics (higher social class, higher IQ, higher levels of human capital, higher motivation etc) favour more positive outcomes. For ethical and practical reasons, the implementation of randomized trials is usually not an option in this area of research (although experiments which randomize the provision of facilities to promote breastfeeding are feasible, as discussed by Kramer et al. 2001 and Kramer et al. 2008). It is only relatively recently that researchers have made systematic attempts to address these issues of causality. The most widely practiced method in the literature is to include a vast array of control variables in the outcome regressions (Rothstein 2011, Heikkila et al 2011, Quigley et al. 2011, Belfield and Kelly 2010, Denny and Doyle 2010, Gibson-Davis and Brooks-Gunn 2006). Some other papers study sibling pairs to control for unobserved family characteristics (Rothstein 2011, Belfield and Kelly 2010, Rees and Sabia 2009, Der et al 2006, Evenhouse and Reilly 2005). Another approach is to use an instrumental-variables technique. Rothstein (2011) uses State breastfeeding rates and laws about breastfeeding in public as instruments; Del Bono and Rabe (2011) use whether the hospital where the child was born participated in a breastfeeding promotion program; Denny and Doyle (2010) use whether the birth was via c-section; and Belfield and Kelly (2010) use caesarean birth, mother's smoking and alcohol consumption, and county-level variables regarding health care and social assistance.

In this paper we follow a fourth approach, which involves controlling for selection on observables with propensity score matching techniques. To our knowledge only four very recent papers use this method in an attempt to identify causal effects of breastfeeding on children outcomes (Rothstein 2011, Jiang et al 2011, McCrory and Layte 2011, Belfield and Kelly 2010). PSM involves "twinning" each breastfed baby with one or more babies who were not breastfed, but who in all other observable respects are similar to the breastfed baby. In this way, we effectively simulate an experiment by creating matched "treatment" and "control" samples,

composed respectively of women who do and who do not breastfeed, but who are identical in every other observable respect (Rosenbaum and Rubin 1983a). If matching is perfect, differences in mean outcomes between both samples may be used as estimates of the causal effect of breastfeeding. One advantage of PSM over regression analysis is that PSM is non-parametric. Matching neither imposes functional form restrictions such as linearity on the outcome equations nor assumes a homogeneous treatment effect across the population. Both assumptions are usually unjustified either by economic theory or by the data (Zhao 2008).

The data we use come from the Avon Longitudinal Study of Parents and Children (ALSPAC). ALSPAC contains a particularly rich set of variables on infant feeding, from which we may establish the duration of breastfeeding and of exclusive breastfeeding; the stage at which formula milk, animal milks and solid foods were introduced; and the types of supplementary foods which were introduced. However, the feature of ALSPAC which makes this data particularly attractive for the implementation of PSM is the fact that parents were interviewed several times prior to the birth of their children, and data collected on their attitudes to breastfeeding and whether they themselves had been breastfed as babies. The data are discussed in Section 2. The data are thus particularly well suited for PSM, which rests on the assumption that the important factors affecting an individual's choice to breastfeed are observable, and that there is no significant selection on unobservables¹. The very rich data we use mean that we are able to implement matching to a very high standard, and thus, that the risk of bias caused by unobservable heterogeneity is low.

We find that breastfeeding is strongly related to children's cognitive outcomes. This is statistically significant, and stands at around one twentieth of a standard deviation at age five, and on average a little over one tenth thereafter. It persists over time, and is common across all the cognitive measures we examine. PSM estimates tend to be slightly lower than OLS estimates controlling for the same set of background variables. This suggests that not taking into account the selection effect of breastfeeding may overestimate the actual returns to breastfeeding, but in this case only very slightly. We find little evidence of a significant relationship between breastfeeding and noncognitive outcomes. If anything, breastfeeding appears to be associated with an increase in behavioural problems for very small children; to the extent that there is a positive relationship, it is found later in children's lives, and is found predominantly in the children of lower-educated mothers.

When considering breastfeeding duration with generalized propensity score (GPS) and OLS methods, we find important nonlinearities, in relation to both cognitive and noncognitive outcomes. For cognitive outcomes, we observe a marked dose-response relationship, with a strong association between breastfeeding and improved cognitive outcomes for the first few months of breastfeeding, becoming flatter at longer

¹ Thus, PSM differs importantly from Instrumental Variable methods, which rely on the assumption that unobservables are important and difficult to control for regardless of how precise the information on the individual and her environment is. IV is often difficult to implement in the study of breastfeeding, due to a lack of suitable instruments. Nevertheless, selection on observables is a strong assumption and we discuss its plausibility at length below.

durations. For noncognitive indicators, by contrast, we find that the first few months of breastfeeding are not strongly associated with improved outcomes, but that an association emerges at durations longer than two or three months. Overall, we find a very robust positive relationship between breastfeeding and cognitive outcomes, which is consistent for different definitions of breastfeeding; in the noncognitive case, the relationship is less robust, and observable only for children of lower educated mothers and for older children breastfed for at least two months.

This paper makes at least two significant contributions to a growing literature on breastfeeding. First, in contrast to other papers studying the relationships between breastfeeding and cognitive and noncognitive abilities, which are usually based on a single breastfeeding variable, we use the rich information in the ALSPAC data to construct several measures of breastfeeding: initiation, duration of breastfeeding, exclusive breastfeeding, formula-feeding at birth, and the introduction of supplementary foods. We are thus able to investigate whether exclusivity or duration of breastfeeding, or both, are decisive in children's development. To our knowledge, none of the previous studies has analyzed the introduction of unsuitable foods or used the duration of breastfeeding as a continuous measure, allowing for nonlinearities. Our second contribution is that this is the first study to analyse a wide set of both cognitive and noncognitive outcomes over the long term (up to 14 years of age). We use a longitudinal data set where both cognitive and noncognitive outcomes are observed repeatedly. Only Del Bono and Rabe (2011) and Belfield and Kelly (2010) consider noncognitive outcomes together with cognitive measures, but with a much smaller age range, 3 to 7 year olds, in the first case, and 1 to 4, in the second. The only other study following children for a comparably long period of time is that of Oddy et al. (2010), who analyze 2 to 14 year-olds, but they consider a more limited number of outcomes related to mental health.

This paper is organized as follows. The data are discussed in Section 2. Baseline results and heterogeneity are considered in Section 3. Section 4 examines possible deviations from the Conditional Independence Assumption (CIA), Section 5 explores the sensitivity of results to different breastfeeding measures, and Section 6 concludes.

2 Data and descriptive statistics

The Avon Longitudinal Survey of Parents and Children (ALSPAC) is a longitudinal study of around 12,000 children born in the Avon area in the early 1990s (Golding et al. 2001; Gregg et al. 2005). Mothers were recruited into the sample at the point at which they first reported their pregnancy to their doctors. Data were collected at four points during pregnancy and at several points following birth – from both parents, from the child him/herself, and from the child's teacher and school. Topic areas covered include physical and mental health, socioeconomic status, and child development; school-level data and children's test results are available via merged records. Ethical approval for this study was obtained from the ALSPAC Law and Ethics Committee and the Local Research Ethics Committees.

2.1 *Sample*

For our analysis, we consider a sample of children in the “core sample” of ALSPAC. This sample consists of 14,541 pregnancies that resulted in 14,676 known foetuses of which 14,062 were live births and 13,988 were alive at one year.² The number of children for which the mother enrolled in the ALSPAC study and had either returned at least one questionnaire or attended a “Children in Focus” clinic by 19/07/99, and returned at least one post-birth questionnaire is 12,268. We employ a maximizing strategy with respect to sample size which implies using as many observations as possible for each outcome measure potentially affected by breastfeeding. Sample sizes thus vary depending on the outcome measure actually used. Table A.1 in the Appendix shows the effect of the different sample selections on sample size and the distribution of selected variables. Overall the different samples retain the main sample representativeness. We acknowledge, nonetheless, a slight decrease in the proportion of breastfed children, together with a decrease in the proportion of high educated mothers, in the sample used in the Entry Level cognitive assessment, and a slight increase in the proportion of breastfed children, together with a minor increase in the proportion of high-educated mothers, in the sample used in the 42-months noncognitive evaluation. In comparative terms, the representativeness of these samples that correspond to young children (five- and three-and-a-half-years-old, respectively) is not as satisfactory as the accuracy of the samples that consider older children.

2.2 *Main variables*

The cognitive outcome variables used in this paper are Standard Attainment Test (SATs) scores³ at ages 7, 11 and 14⁴, and the results of school entry tests at age 5. All scores are standardized to have mean zero and standard deviation 1; thus, all the results we present may be interpreted as proportions of a standard deviation. In the UK, children generally start school in the September following their fourth birthday, and move up a school year every September. During the years when the ALSPAC cohort were in school, pupils’ progress was monitored by compulsory SATs tests, administered at the end of each of the “Key Stages” in schools (See Table 1). When the ALSPAC cohort entered school there were no national compulsory entry assessments. However, the four Local Education Authorities covering the former Avon area (Bristol, South Gloucestershire, Bath & North East Somerset and North Somerset) all used the same entry assessment scheme, which was used by 80% of the local state schools.

² For reasons of confidentiality data on the 13 triplet and quadruplet children were not available for analysis.

³ Scores were matched to ALSPAC data from the National Pupil Database (NPD), a central repository for pupil level educational data established in 2002, which assigns every pupil in England a Unique Pupil Number (UPN).

⁴ SATs are also taken at age 16. However, we do not consider these in this paper, since not all the children in our sample took the same national test at this point.

Table 1: Structure of UK schools and testing procedures

Class	Age	Key Stage	Tests
--	3 - 4		--
Reception	4 - 5	Foundation	Local entry assessments
Year 1	5 - 6		--
Year 2	6 - 7	1	National tests and tasks in English and Maths
Year 3	7 - 8		--
Year 4	8 - 9		--
Year 5	9 - 10		--
Year 6	10 - 11	2	National tests in English, Maths and Science
Year 7	11 - 12		--
Year 8	12 - 13		--
Year 9	13 - 14	3	National tests in English, Maths and Science
Year 10	14 - 15		Some children take GCSEs
Year 11	15 - 16	4	Most children take GCSEs, GNVQs etc.

Source: Department for Children, Schools and Families (DCSF) at <http://www.dfes.gov.uk/>

The entry-level assessment covers eight areas, of which four (language, reading, writing and mathematics) relate to cognitive development and four to non-cognitive development⁵. The KS1 battery consists of tasks in reading, reading comprehension, writing, spelling and mathematics. We report results for reading, writing and mathematics at entry level and KS1; at Key Stages 2 and 3, tests are administered in English, Maths and Science, and we report results in all of these areas. Scores are recorded on a scale of 2 to 7 (entry level); 0 to 5 (KS1); and on more detailed ranges of marks at KS2 and KS3.

The first set of non-cognitive outcome variables is derived from parental assessments at 42 months. When the child was about three and a half years old, parents were asked to assess their child on different dimensions following the Revised Rutter Parent Scale for Preschool Children, which is an extension of the Rutter behaviour scale (Elander and Rutter 1996). Parents were given descriptions of children, and were asked to tick the box that best described their child. For example, to the statement “Tends to do things on his own, rather solitary” parents were given the choice of answering “yes certainly”, “yes sometimes”, and “no”. There were a total of 43 such questions; responses were aggregated by the ALSPAC team to create scores in five different domains: emotional difficulties, conduct difficulties, hyperactivity, behaviour difficulties, and prosocial behaviour (Golding et al. 2003).

Another set of non-cognitive outcome variables come from the teacher version of the Strengths and Difficulties Questionnaire (SDQ) (Goodman, 1997), which the children’s class teachers were asked to complete when children were in Years 3 and 6 at school. Teachers were asked to think about the child’s behaviour over the past six months, and to tick the response which most closely corresponded to their impression of the child. For example, in answer to the question “Is [*the child*] considerate of other people’s feelings?” teachers could answer “not true”, “somewhat true”, “certainly true”. Scores on the five SDQ domains (emotional symptoms,

⁵ We do not include results from these non-cognitive tests in our analysis, since assessment in these areas was voluntary and only a small fraction of the sample (25%) completed these tests.

conduct problems, hyperactivity, peer problems, and prosocial behaviour) were constructed as weighted sums of these variables (ALSPAC Study Team 2008a, 2008b); the scores we use are constructed on completed cases only.⁶ For all noncognitive scores except those measuring prosocial behaviour, higher scores indicate a greater incidence of problems. As with the cognitive outcomes, all scores are standardized to have mean zero and standard deviation 1; thus, all the results we present may be interpreted as proportions of a standard deviation.

Most of the analysis in this paper is based on whether a mother is breastfeeding (at all) when her baby is four weeks old. This has a number of advantages. First, it splits the sample approximately in half, and provides reasonably large samples both of relatively privileged women who are not breastfeeding at four weeks, and of relatively less privileged women who are still breastfeeding at four weeks. It also largely circumvents the issue of solid food, as only a tiny percentage of babies are given solid food before four weeks old. Nevertheless we include a robustness check section (Section 6) in which other breastfeeding measures are analyzed.

As a first approximation to our research question, Table 2 summarizes all outcome variables by breastfeeding status. Breastfeeding for at least four weeks is associated with more favourable outcomes on both cognitive and noncognitive measures.

⁶ Only about 2% of the observations had some missing information in any of these scores.

Table 2 Descriptive statistics: outcome variables.

			Not breastfed at 4 weeks		Breastfed at 4 weeks	
Cognitive outcomes	Age 5	Reading	-0.122	0.980	0.220	0.967
		Writing	-0.077	0.998	0.166	0.962
		Maths	-0.108	0.994	0.218	0.960
	Age 7	Reading	-0.050	1.003	0.176	0.944
		Writing	-0.139	0.967	0.242	0.970
		Maths	-0.133	0.999	0.232	0.935
	Age 11	English	-0.187	1.004	0.253	0.917
		Maths	-0.181	1.012	0.248	0.918
		Science	-0.201	1.016	0.285	0.875
	Age 14	English	-0.194	0.971	0.272	0.962
		Maths	-0.164	0.979	0.222	0.971
		Science	-0.099	1.016	0.155	0.935
Non-cognitive outcomes	Parental assessment 42 months	Emotional Difficulties	0.052	1.025	-0.048	0.973
		Conduct Problems	-0.008	1.014	-0.001	0.987
		Hyperactivity	0.056	1.008	-0.047	0.993
		Prosocial	0.011	1.009	-0.007	0.991
	Teacher assessment Year 3	Emotional Symptoms	0.026	1.015	-0.062	0.953
		Conduct Problems	0.068	1.079	-0.119	0.832
		Hyperactivity	0.081	1.030	-0.135	0.927
		Peer problems	0.031	1.022	-0.064	0.955
		Prosocial	-0.034	1.007	0.070	0.977
	Teacher assessment Year 6	Emotional Symptoms	0.028	1.025	-0.071	0.927
		Conduct Problems	0.042	1.062	-0.117	0.828
		Hyperactivity	0.060	1.025	-0.134	0.921
		Peer problems	0.001	1.001	-0.035	0.978
Prosocial		-0.001	0.992	0.048	0.982	

Notes:

The table shows means (standard deviations) by breastfeeding status at 4 weeks. Source: ALSPAC Data, core sample.

2.3 Control variables

A major advantage of ALSPAC is that it contains an extremely rich set of background variables, which allows the implementation of PSM to a very high standard. Table 3 presents the mean and standard deviations for our control variables by breastfeeding status (see Table A.2 in Appendix A for a precise definition of all control variables). Consistent with available literature (see for instance Imbens and Wooldridge 2009), only variables that may simultaneously influence the breastfeeding decision and the outcome variable are considered. In general, mothers who do not breastfeed when their babies are 4 weeks old have generally worse

pregnancy outcomes, worse health, belong to lower socio-economic groups and are more likely to be in a non-marital relationship. Mothers (and fathers) of children who are breastfed have higher educational attainments. In particular, parents of breastfed babies are more likely to have a university degree or A-levels than parents of babies who were not breastfed at 4 weeks. So it is not clear that the association between breastfeeding and cognitive and noncognitive outcomes of Table 2 is entirely causal. If children from more advantaged households are also more likely to be breastfed (as shown by the positive correlation between maternal education and breastfeeding incidence in Table 3) and also perform better at these tests, either because of higher innate ability or because they are better trained for them, then the association between breastfeeding and cognitive and noncognitive outcomes shown in Table 2 might well be spurious.

Table 3: Descriptive statistics. Control variables

		Not breastfed at 4 weeks		Breastfed at 4 weeks	
Pregnancy and birth	Mother works at 18 weeks	0.572	(0.007)	0.662	(0.007)
	Gestation in weeks	39.321	(0.185)	39.517	(0.173)
	Mother's age at birth	27.021	(0.159)	29.454	(0.149)
	C-section	0.118	(0.004)	0.095	(0.004)
	Smoked previously	0.405	(0.007)	0.25	(0.006)
	Cigarettes at 32 w	3.205	(0.078)	1.269	(0.07)
	Previous alcohol consumption	0.903	(0.005)	0.939	(0.005)
	Alcoholic drinks at 8w	1.701	(0.07)	1.659	(0.065)
Child characteristics at birth	Female	0.477	(0.007)	0.491	(0.007)
	Twin	0.034	(0.002)	0.015	(0.002)
	Birth weight	3366.421	(17.006)	3443.19	(15.866)
	Head circumference	34.72	(0.161)	34.827	(0.15)
	Crown-heel length	50.488	(0.235)	50.76	(0.219)
Mother's health in pregnancy	Mother bad health	1.82	(0.012)	1.716	(0.011)
	Locus/control score	4.87	(0.037)	3.814	(0.035)
	CC score at 32w	16.019	(0.132)	14.455	(0.119)
	EPND score at 32w	7.522	(0.077)	6.53	(0.07)
Socio-economic variables	Owner occupier	0.674	(0.006)	0.827	(0.006)
	Private rented	0.069	(0.004)	0.066	(0.003)
	Number of rooms	1.452	(0.014)	1.717	(0.013)
	Neighbourhood qual.	7.98	(0.049)	8.293	(0.045)
Demographic variables	White mother	0.91	(0.004)	0.941	(0.004)
	Mother cohabiting	0.239	(0.006)	0.167	(0.005)
	Mother single	0.039	(0.002)	0.018	(0.002)
Father's education	Degree	0.089	(0.006)	0.294	(0.005)
	A-level	0.282	(0.007)	0.301	(0.006)
	O-level	0.261	(0.007)	0.217	(0.006)
	Vocational	0.119	(0.004)	0.074	(0.004)
	CSE	0.221	(0.005)	0.107	(0.005)
Mother's education	Degree	0.049	(0.005)	0.213	(0.004)
	A-level	0.168	(0.006)	0.304	(0.005)
	O-level	0.418	(0.007)	0.335	(0.006)
	Vocational	0.143	(0.004)	0.071	(0.004)
	CSE	0.222	(0.005)	0.078	(0.005)
Breastfeeding attitudes & intentions	Mother was breastfed	0.462	(0.008)	0.669	(0.007)
	Father was breastfed	0.59	(0.01)	0.718	(0.009)
	Mother breastfeeding attitudes	14.318	(0.07)	17.57	(0.063)
	Father breastfeeding attitudes	14.387	(0.081)	16.445	(0.072)

Notes:

The table shows means (standard deviations) by breastfeeding status at 4 weeks. Source: ALSPAC Data, core sample.
Source: ALSPAC Data, core sample.

3 Breastfeeding and child's cognitive and non-cognitive development

This section presents estimates of the relationship between breastfeeding and cognitive and non-cognitive test scores using OLS and PSM estimation methods. We use the latter method to estimate two parameters of interest: the average treatment effect on the treated (ATT) and the average treatment effect on the untreated (ATU). If we consider D the treatment indicator; Y_1 the outcome if treated (breastfed); and Y_0 the outcome if untreated (not breastfed), the ATT measures the difference between the average outcome measure for babies who were breastfed, and the average outcome measure for the same group under the hypothetical scenario that they had not been breastfed, $ATT = E(Y_1|D=1) - E(Y_0|D=1)$, and the ATU estimates the difference between the average outcome measure in the hypothetical case that they were breastfed, and the actual case in which they were not breastfed, $ATU = E(Y_1|D=0) - E(Y_0|D=0)$.⁷ Compared to OLS, the advantage of using propensity score matching is that it is possible to determine how well the treatment and control groups overlap, that is, to assess the quality of the matching; given that this method is non-parametric estimates are therefore less sensitive to the choice of functional form in the model (Rosenbaum and Rubin 1983a; Dehejia and Wahba 2002).

Table 4 presents estimates of the relationship between breastfeeding and cognitive test scores using OLS and PSM estimation methods. The first specification includes very few additional controls – only the child's sex and his or her age at the time of taking the test. The second specification adds in a range of variables indicating parental education, while the third specification controls for the full set of variables described in Section 2. Specifications (4) and (5) are the ATT and the ATU from the PSM regressions where all the controls used in Specification (3) are included.

⁷ Under the assumption of homogeneous treatment effects ATT and ATU should give the same results. However, as Heckman et al. (1997) have shown, treatment effects are rarely homogeneous. We therefore report both sets of estimates.

Table 4: Breastfeeding and Cognitive outcomes.

		(1)		(2)		(3)		(4)		(5)	
		OLS: Specific. 1		OLS: Specific.2		OLS: Specific. 3		PSM: ATT		PSM: ATU	
Entry level	Reading	0.328 (0.021)	***	0.148 (0.022)	***	0.079 (0.023)	***	0.095 (0.028)	***	0.082 (0.028)	**
	Writing	0.226 (0.021)	***	0.085 (0.022)	***	0.037 (0.024)		0.040 (0.028)		0.012 (0.031)	
	Maths	0.307 (0.021)	***	0.122 (0.021)	***	0.049 (0.023)	*	0.052 (0.025)	*	0.040 (0.030)	
Key Stage1	Reading	0.221 (0.025)	***	0.108 (0.025)	***	0.060 (0.027)	*	0.062 (0.030)	*	0.114 (0.033)	***
	Writing	0.363 (0.019)	***	0.147 (0.020)	***	0.078 (0.021)	***	0.076 (0.027)	**	0.097 (0.022)	***
	Maths	0.357 (0.019)	***	0.157 (0.020)	***	0.100 (0.021)	***	0.098 (0.031)	**	0.112 (0.027)	***
Key Stage 2	English	0.420 (0.019)	***	0.149 (0.019)	***	0.098 (0.020)	***	0.068 (0.025)	**	0.122 (0.022)	***
	Maths	0.410 (0.019)	***	0.151 (0.019)	***	0.111 (0.021)	***	0.102 (0.023)	***	0.127 (0.026)	***
	Science	0.464 (0.019)	***	0.200 (0.019)	***	0.146 (0.020)	***	0.119 (0.020)	***	0.161 (0.023)	***
Key Stage3	English	0.450 (0.020)	***	0.174 (0.020)	***	0.118 (0.021)	***	0.115 (0.028)	***	0.116 (0.027)	***
	Maths	0.368 (0.021)	***	0.158 (0.021)	***	0.118 (0.023)	***	0.110 (0.030)	***	0.137 (0.028)	***
	Science	0.245 (0.021)	***	0.125 (0.022)	***	0.106 (0.024)	***	0.085 (0.030)	**	0.115 (0.026)	***

Notes:

Columns 1, 2, and 3 show OLS regression coefficients of the different test scores on breastfeeding for at least 4 weeks. Specification 1 controls for the child's age and sex. Specification 2 also controls for maternal education. Specification 3 additionally includes the baby's birth weight, crown-heel length, and head circumference; gestation and mode of delivery (vaginal or caesarean section); whether the child is twin; the mother's age at birth; the mother's race and marital status; the education levels of both parents, housing tenure, the size of the home, neighbourhood characteristics, whether the mother had been in care as a child, whether she had divorced parents, whether her carer died prematurely; the mother's health, whether mother had smoked during pregnancy, the mother's mental health; the mother's labour market participation; and whether the mother and father had been breastfed themselves as babies and on their attitudes towards breastfeeding, measured prenatally. Standard errors in parentheses.

Columns 4 and 5 show ATT- and ATU- PSM estimates of the effect of breastfeeding for at least 4 weeks including all the above variables as controls. We use the Epanechnikov kernel algorithm with 0.05 bandwidths, imposing the common support condition with the psmatch2 Stata command (Leuven and Sianesi 2003). PSM standard errors in parentheses computed by bootstrapping with 100 repetitions.

Significance denoted by asterisks: * = 5%, ** = 1%, *** = 0.1%

Source: ALSPAC core sample.

Estimates in the first specification (column 1 in Table 4) indicate that breastfeeding is associated with test scores higher by about one third of a standard deviation, with coefficients ranging from 0.221 (Reading, KS1) to 0.464 (Science, KS2). As soon as parental education is controlled for (OLS: Specification 2) these estimates drop sharply, to approximately one third of their original size. Estimates now range from 0.085

(Writing, entry level) to 0.200 (Science, KS2). As we would expect, controlling for additional factors reduces the estimates still further – but in fact, not by very much. A case in point is the coefficient on English at KS3, which falls from 0.450 to 0.174 between Specifications 1 and 2, but only to 0.118 in Specification 3. This suggests that a very large amount of the heterogeneity between breastfeeding and non-breastfeeding families is captured by parental education, with all the other controls together (including attitudes to breastfeeding and intention to breastfeed) capturing only a small proportion after education is controlled for. The estimates in the third specification are of the order of 10 per cent of a standard deviation; most of the estimates remain statistically significant, although some estimates in the early years are not.

Columns 4 and 5 present estimates of the ATT and ATU obtained via PSM techniques when all control variables are included. We use the Epanechnikov kernel algorithm with 0.05 bandwidths, imposing the common support condition. Due to the high quality data and the reasonable sample sizes, matching quality is very high. Table A.3 in the Appendix displays the mean values of the variables used in the analysis for the treatment and control groups after matching in each of the matching procedures performed – one for each testing period. Overall, the figures in Table A.3 confirm that our treatment and comparison groups, though initially somewhat different, look extremely similar after matching, with no significant differences in any of the 36 background variables in the PSM for cognitive outcomes, and very small differences for only two of the variables in the PSM for noncognitive outcomes.

The PSM estimates of the ATT (column 4) are essentially the same as the OLS coefficients with a large set of control variables. In general the significance levels are a little lower, but not for all outcome measures. However, the estimated relationships are still significant at all ages after school entry level, confirming our previous result that controlling for a wide range of factors, children breastfed for four weeks or more do better than children breastfed for less than four weeks by about one tenth of a standard deviation (slightly less at younger ages, and slightly more at older ages). PSM estimates tend to be slightly lower than OLS estimates controlling for the same set of background variables. This suggests that not taking into account the selection effect of breastfeeding may overestimate the actual returns to breastfeeding, but in this case only very slightly. However, the PSM results control for the mother's selection into breastfeeding on the basis of observables, i.e., for the fact that more able mothers (who are more likely also to have more able children) are also the ones who are more likely to breastfeed. Therefore, the difference between the OLS and PSM results gives an indication of the size and direction of the selection effect.

Column 5 in Table 4 presents results for the ATU. Differences between ATT and ATU suggest that the relationships between breastfeeding and child outcomes may differ between those babies who are currently likely to be breastfed, and those babies who are not currently likely to be breastfed. In fact, the ATU estimates are slightly higher than ATT estimates, suggesting that to the extent that breastfeeding does have a causal effect

on cognitive outcomes, that effect would be somewhat higher for children whose mothers are less likely to breastfeed them.

Overall, our results from Table 4 provide evidence that breastfeeding is responsible a statistically significant, increase in SATs scores – around 10 per cent of a standard deviation, which translates roughly into about two positions in a class of 30 children. These results are consistent with previous findings. In particular, the magnitude of the coefficients coincides with the 9.1%-10.7% of a standard deviation reported by Denny and Doyle (2010), who also analyze breastfeeding for at least four weeks in UK children (National Child Development Study). These differences are visible at all the ages we consider (5, 7, 11 and 14) and across a range of subject areas. In fact, coefficients seem to increase in magnitude with time, in line with Cunha and Heckman's (2010) hypothesis of early investments in children having multiplier effects in later childhood. This fact corresponds to the results reported by Del Bono and Rabe (2011), also for United Kingdom (Millennium Cohort Study).

There remains a question however about the reliability of SATS scores. The presence of a random component of measurement error strengthens rather than weakens our conclusions, since this random element would serve to bias estimated coefficients towards zero. We have additionally aggregated the scores from all the tests taken in a particular key stage, since this process is known to reduce bias due to measurement error and the positive results of a comparable magnitude persist.⁸ There is an additional problem of what SATS actually measure. SATS scores are highly influenced by social class, as children from advantageous backgrounds are more likely to be prepared for the tests, or to go to schools which better coach them for these tests. We have controlled for a wide set of socio-economic household variables during the matching process, which may take care of the fact that children from more advantageous backgrounds may be formally or informally coached at home. Also, given that in the ALSPAC data there are lots of children attending the same schools, and there are not many schools, the second problem is likely to be less of a concern. As a robustness check we included school fixed effects in the six outcome variables measured after school entry. We obtained the same results, so we are confident that what we are capturing here is not the effect of better schooling.⁹

Additionally, one may argue whether the observed relationships could be due to the absence of a direct measure of cognition, such as maternal IQ. We estimated the model on a restricted sample of mothers for which maternal IQ information is available (2817 obs.). We did not find any significant results, either including or excluding the maternal IQ variable in this sample. We interpret that this small sample is highly selected among motivated, willing mothers, where breastfeeding-for-4-weeks rates are more than 67%, instead of 55 percent of

⁸Results available upon request.

⁹Results available upon request.

our general sample. This fact together with sample size may prevent us from finding any significant relationship between breastfeeding and cognitive outcomes.¹⁰

Table 5 presents estimates of the relationship between breastfeeding and noncognitive outcomes (Rutter scores and SDQ scores) from OLS and PSM regressions. Three OLS specifications are presented. Each controls for whether the child was breastfed or not at four weeks of age. The first specification includes very few additional controls – only the child’s sex and his or her age at the time of taking the test. The second specification adds in a range of variables indicating parental education, while the third specification controls for the full set of variables described in Section 2. Estimates in the first specification indicate that breastfeeding is associated with better noncognitive skills, although the relationship is much smaller and less precisely estimated than in the case of cognitive skills presented in Table 4. Breastfeeding accounts for less than one fourth of a standard deviation, with significant coefficients ranging in absolute value from 0.196 (hyperactivity in Year 3) to 0.013 (conduct problems at 42 months). As soon as parental education is controlled for (Specification 2) these estimates drop sharply, to approximately two thirds of their original size. Significant coefficients now range in absolute value from 0.098 (conduct problems in Year 3) to 0.005 (peer problems in Year 6). The statistical significance of the coefficients is further reduced, and many stop being statistically significant altogether. As we would expect, controlling for additional factors reduces the estimates still further – although as in the case of cognitive outputs in Table 4, the decrease is not as large as when introducing controls for parental education. Significance levels drop quite substantially, however, and only hyperactivity and peer problems in Year 3 remain significant at the 5% level, with coefficients of -0.068 and -0.077, respectively. In addition, two measures not shown to be influenced by breastfeeding in previous results – conduct problems and prosocial behaviour at 42 months – now show a significant relationship, although, surprisingly, with a counterintuitive sign: it appears that breastfeeding may be negatively influencing the behaviour of relatively young children in these domains.

¹⁰ Results available upon request.

Table 5: Breastfeeding and Noncognitive outcomes

		(1)	(2)	(3)	(4)	(5)
		OLS: Specific. 1	OLS: Specific. 2	OLS: Specific. 3	PSM: ATT	PSM: ATU
Parent, 42 months	Emotional difficulties	-0.092 *** (0.021)	-0.062 ** (0.022)	-0.012 (0.024)	0.033 (0.033)	-0.017 (0.026)
	Conduct problems	0.013 (0.021)	0.027 (0.022)	0.075 ** (0.024)	0.099 *** (0.028)	0.080 * (0.031)
	Hyperactivity	-0.095 *** (0.021)	-0.069 ** (0.023)	-0.043 (0.024)	-0.054 * (0.025)	-0.006 (0.028)
	Prosocial	-0.029 (0.021)	-0.043 (0.022)	-0.054 * (0.024)	-0.092 ** (0.030)	-0.054 * (0.026)
Teacher, Year 3	Emotional Symptoms	-0.089 ** (0.027)	-0.047 (0.029)	-0.029 (0.031)	-0.041 (0.037)	-0.043 (0.039)
	Conduct Problems	-0.172 *** (0.025)	-0.098 *** (0.027)	-0.056 (0.028)	-0.021 (0.031)	-0.089 ** (0.034)
	Hyperactivity	-0.196 *** (0.026)	-0.093 *** (0.027)	-0.068 * (0.029)	-0.017 (0.033)	-0.115 ** (0.036)
	Peer problems	-0.094 *** (0.027)	-0.084 ** (0.029)	-0.077 * (0.031)	-0.064 (0.037)	-0.080 * (0.033)
	Prosocial	0.086 *** (0.026)	0.051 (0.028)	0.015 (0.030)	-0.050 (0.033)	0.038 (0.034)
Teacher, Year 6	Emotional Symptoms	-0.091 *** (0.025)	-0.034 (0.027)	-0.006 (0.029)	-0.011 (0.029)	-0.019 (0.039)
	Conduct Problems	-0.129 *** (0.023)	-0.034 (0.025)	0.019 (0.026)	0.000 (0.030)	0.026 (0.040)
	Hyperactivity	-0.154 *** (0.023)	-0.028 (0.025)	0.011 (0.026)	0.016 (0.034)	-0.015 (0.035)
	Peer problems	-0.034 (0.025)	0.005 (0.027)	0.012 (0.029)	0.025 (0.036)	0.023 (0.036)
	Prosocial	0.020 (0.024)	-0.030 (0.026)	-0.052 (0.028)	-0.061 (0.034)	-0.046 (0.036)

Notes:

Columns 1, 2, and 3 show OLS regression coefficients of the different test scores on breastfeeding for at least 4 weeks. Specification 1 controls for the child's age and sex. Specification 2 also controls for maternal education. Specification 3 additionally includes the baby's birth weight, crown-heel length, and head circumference; gestation and mode of delivery (vaginal or caesarean section); whether the child is twin; the mother's age at birth; the mother's race and marital status; the education levels of both parents, housing tenure, the size of the home, neighbourhood characteristics, whether the mother had been in care as a child, whether she had divorced parents, whether her carer died prematurely; the mother's health, whether mother had smoked during pregnancy, the mother's mental health; the mother's labour market participation; and whether the mother and father had been breastfed themselves as babies and on their attitudes towards breastfeeding, measured prenatally. Standard errors in parentheses.

Columns 4 and 5 show ATT- and ATU- PSM estimates of the effect of breastfeeding for at least 4 weeks including all the above variables as controls. We use the Epanechnikov kernel algorithm with 0.05 bandwidths, imposing the common support condition with the psmatch2 Stata command (Leuven and Sianesi 2003). PSM standard errors in parentheses computed by bootstrapping with 100 repetitions.

Significance denoted by asterisks: * = 5%, ** = 1%, *** = 0.1%

Source: ALSPAC core sample.

Columns 4 and 5 in Table 5 present PSM estimates for non-cognitive outcomes.¹¹ Most relationships become statistically insignificant when controlling for selection on observables, so the first impression is that breastfeeding does not have a clear relationship with non-cognitive outcomes. This agrees with the limited existing evidence on the subject for the United Kingdom. Heikkila et al. (2011), using logistic regression to estimate the probability of abnormal behaviour on 5-year-old children, find no significant relationship between breastfeeding and any of the computed SDQ scores. They only find some evidence of a positive influence of breastfeeding on some noncognitive measures for children exclusively breastfed for more than 4 months. Kramer et al. (2008), using data from a large scale, randomized experiment in the Republic of Belarus, also report no consistent and significant differences in behavioural strengths and difficulties (SDQ) in 6-year-old children whose mothers were encouraged to breastfeed exclusively and for a longer duration. In the United States, however, Belfield and Kelly (2010) find some positive relationships in the case of motor scores and maternal attachment, irrespective of the breastfeeding measure used.

A few significant coefficients are evident. In column 4, which presents ATT estimates, and column 5, which present ATU estimates, the surprising negative coefficients on breastfeeding for parental assessments of conduct difficulties and prosocial behaviour at 42 months of age remain. Both coefficients are somewhat below 10 per cent of a standard deviation. So apparently, when selection on observables is controlled for, we find that breastfeeding is *negatively* associated, in the short run, with children's development in the domains of antisocial behaviour and conduct difficulties. However, the fact that the significance of this result diminishes when the ATU sample is used calls for additional analysis. A potential explanation may be found in the fact that measurement error is likely to be larger in the case of parental assessments when a child is young. Johnston et al. (2011) show that expert psychiatric assessment places greatest weight on teachers' views and rather less on those of parents, indicating a larger measurement error for parents' assessments. Moreover, they also show that observers, both parents and teachers, tend to overstate problems of younger children relative to those of older children and adolescents, by the standards of the fully-informed expert psychiatric diagnosis. In the same vein, Kramer et al. (2008) note that, compared to teachers' ratings, parents' ratings tend to be higher and more extreme for all SDQ scales.

Two other significant results must be commented on. The ATU estimates in column 5 indicate that breastfeeding is associated with a lower risk of conduct problems and hyperactivity in Year 3, but that these results do not persist to Year 6. These results, together with the beneficial results found on hyperactivity and peer problems for Year 3 (but not Year 6) children for the OLS method (column 3), may indicate that breastfeeding has a favourable effect over the medium term. Overall, nonetheless, we find little evidence of significant effects of breastfeeding at least four weeks on noncognitive outcomes.

¹¹ Table A.4 in the Appendix analyses the quality of matching for the procedures used in this table.

3.1 *Heterogeneity of results*

By and large, results from Tables 4 and 5 show that whereas the relationship between breastfeeding and improved outcomes is persistent and statistically significant for all cognitive outcomes, the results for non-cognitive outcomes are much less conclusive. It is possible, of course, that a degree of bias may still be driving either or both sets of results. Recent developments in matching methods emphasize that combining exact matching on key covariates with propensity score matching is a design analogous to blocking in a randomized experiment and can lead to large reductions in bias (Dehejia and Wahba 1999; Stuart and Rubin 2007). Therefore in this section, we stratify the sample with respect to especially relevant confounding factors and re-estimate the model. In particular we investigate whether our estimated results vary across maternal education levels.¹²

Table 6 reports the ATT- and ATU- PSM estimates of the relationship between breastfeeding and cognitive and noncognitive outcomes for each population group. We expect that stratifying by maternal education will have a fairly significant effect on results, given the effect mentioned earlier of including this variable as a control in OLS estimates. Columns 1 and 2 in each table present estimated ATT and ATU results for higher educated mothers (Degree and A-level) while columns 3 and 4 offer the corresponding results for lower educated mothers (O-level, vocational and CSE).¹³ With respect to cognitive outcomes, more substantial positive relationships emerge for lower-educated mothers. This is in contrast to previous results from Gibson-Davis and Brooks-Gunn (2006) for United States who only find positive breastfeeding effects for mothers with at least some post-secondary education. Rothstein (2011) obtains larger effects for children of medium- to low-educated mothers for some outcomes, but not others.

In respect of noncognitive outcomes, our results show, interestingly, that although breastfeeding tends to be associated with worse short-run noncognitive outcomes for children of both high- and low-educated mothers, there are several positive medium- to long-term noncognitive effects for children of low-educated mothers. Thus, we find a decrease in prosocial behaviour at Year 6 for children of high-educated mothers, but we find a decrease in conduct problems, hyperactivity and peer problems for the children of low-educated mothers at Year 3. These results may be considered analogous to the benefits found by Belfield and Kelly (2010) for children below the poverty threshold and by McCrory and Layte (2011) for socially disadvantaged groups.¹⁴

¹² We also experimented with distinguishing pre-term births, as Heikkila et al. (2010), and ethnic origin, as Rothstein (2011), but the standard errors were too large, suggesting that those subdivisions of the sample were too demanding of the data.

¹³ We tried maintaining the five original education levels but sample sizes for most outcome measures precluded obtaining any reliable results.

¹⁴ We also found analogous results stratifying by maternal marital status, with children of non-married mothers obtaining greater benefits from breastfeeding. Results available upon request.

Table 6: Heterogeneity of results by maternal education. Cognitive and noncognitive outcomes. PSM estimates

			Degree, A-level		O-lev., Vocat., CSE					
			(1)	(2)	(3)	(4)				
			ATT	ATU	ATT	ATU				
Cognitive Outcomes	Entry level	Reading	0.068 (0.065)	0.107 (0.049)	*	0.075 (0.033)	*	0.067 (0.030)	*	
		Writing	0.011 (0.055)	0.108 (0.050)	*	0.017 (0.033)		-0.024 (0.037)		
		Maths	0.001 (0.056)	0.060 (0.040)		0.032 (0.033)		0.020 (0.034)		
	Key Stage1	Reading	0.156 (0.076)	0.156 (0.072)	*	0.041 (0.040)		0.066 (0.045)		
		Writing	0.054 (0.046)	0.107 (0.043)	*	0.062 (0.030)	*	0.069 (0.032)	*	
		Maths	0.091 (0.048)	0.138 (0.047)	**	0.090 (0.034)	**	0.089 (0.031)	**	
	Key Stage 2	English	0.000 (0.040)	0.055 (0.041)		0.080 (0.032)	*	0.122 (0.035)	***	
		Maths	0.044 (0.043)	0.101 (0.043)	*	0.096 (0.034)	**	0.114 (0.032)	***	
		Science	0.025 (0.034)	0.083 (0.038)	*	0.142 (0.032)	***	0.154 (0.034)	***	
	Key Stage3	English	0.006 (0.051)	0.081 (0.050)		0.125 (0.031)	***	0.100 (0.029)	***	
		Maths	0.074 (0.058)	0.107 (0.060)		0.093 (0.031)	**	0.112 (0.036)	**	
		Science	0.061 (0.058)	0.073 (0.051)		0.083 (0.035)	*	0.110 (0.038)	**	
	Noncognitive Outcomes	42 months	Emotional difficulties	0.062 (0.053)	-0.004 (0.046)		0.008 (0.030)		-0.020 (0.035)	
			Conduct problems	0.136 (0.054)	0.076 (0.046)	*	0.054 (0.034)		0.075 (0.042)	
			Hyperactivity	-0.068 (0.057)	-0.082 (0.045)		-0.033 (0.035)		0.042 (0.037)	
Prosocial			-0.096 (0.046)	-0.057 (0.042)	*	-0.072 (0.031)	*	-0.052 (0.032)		
Year 3		Emotional Difficulties	-0.077 (0.060)	-0.030 (0.056)		-0.010 (0.039)		-0.044 (0.052)		
		Conduct Problems	0.066 (0.042)	0.017 (0.055)		-0.114 (0.042)	**	-0.130 (0.051)	*	
		Hyperactivity	0.060 (0.058)	0.013 (0.052)		-0.085 (0.044)		-0.158 (0.044)	***	
		Peer problems	-0.038 (0.070)	-0.098 (0.059)		-0.094 (0.041)	*	-0.083 (0.052)		
		Prosocial	-0.087 (0.059)	-0.041 (0.052)		-0.015 (0.044)		0.045 (0.047)		

Table 6: Heterogeneity of results by maternal education. Cognitive and noncognitive outcomes. PSM estimates (Cont.)

		Degree, A-level		O-lev., Vocat., CSE		
		(1)	(2)	(3)	(4)	
		ATT	ATU	ATT	ATU	
Noncognitive Outcomes	Year 6	Emotional	0.012	0.035	-0.017	0.002
		Symptoms	(0.054)	(0.047)	(0.046)	(0.046)
		Conduct Problems	0.001	0.032	0.013	0.046
			(0.044)	(0.065)	(0.036)	(0.053)
		Hyperactivity	0.013	0.024	0.036	-0.016
			(0.056)	(0.062)	(0.039)	(0.045)
		Peer problems	0.079	0.006	-0.005	0.041
	(0.059)	(0.049)	(0.037)	(0.045)		
	Prosocial	-0.108 *	-0.117 *	0.000	-0.022	
		(0.054)	(0.055)	(0.041)	(0.045)	

Notes:

The table shows ATT- and ATU- PSM estimates of the effect of breastfeeding for at least 4 weeks including all as controls the child's age and sex, the baby's birth weight, crown-heel length, and head circumference; gestation and mode of delivery (vaginal or caesarean section); whether the child is twin; the mother's age at birth; the mother's race and marital status; the education levels of both parents, housing tenure, the size of the home, neighbourhood characteristics, whether the mother had been in care as a child, whether she had divorced parents, whether her carer died prematurely; the mother's health, whether mother had smoked during pregnancy, the mother's mental health; the mother's labour market participation; and whether the mother and father had been breastfed themselves as babies and on their attitudes towards breastfeeding, measured prenatally. We use the Epanechnikov kernel algorithm with 0.05 bandwidths, imposing the common support condition, with the psmatch2 Stata command (Leuven and Sianesi 2003). Columns 1 and 2 show ATT and ATU, respectively, computed in the sample of higher educated mothers (Degree and A-level). Columns 3 and 4 show the corresponding result for lower educated mothers (O-level, Vocational, and CSE).

Standard errors in parentheses computed by bootstrapping with 100 repetitions.

Significance denoted by asterisks: * = 5%, ** = 1%, *** = 0.1%

Source: ALSPAC core sample.

4 Deviations from CIA

The previous stratified analysis reveals potentially larger effects of breastfeeding for disadvantaged children and a clear divergence between cognitive vs. non-cognitive outcomes. Our selection-on-observables method requires the conditional independence assumption (CIA) to give estimates a causal interpretation. Selection on observables is more plausible if an ample set of control variables is available. We contend that the very rich data we use mean that a high proportion of the propensity to breastfeed is explained by observable factors and the risk of bias caused by unobservable heterogeneity is low. The CIA is not a testable assumption with non-experimental data. Nevertheless, we provide some evidence on its plausibility using a double strategy: performing a falsification exercise (Heckman and Hotz 1989) and implementing a sensitivity analysis similar to that proposed by Ichino et al. (2008).

4.1 *Falsification Exercise*

In this section we evaluate whether unobservables correlated with breastfeeding incidence and child outcomes could drive the previous results, by performing a falsification exercise (Heckman and Hotz 1989). The exercise consists of identifying a set of pre-treatment outcomes that could be related to these unobservables. If outcomes are measured before the treatment they cannot possibly be affected by the treatment unless there remains some selection bias not properly accounted for by the PSM method. Therefore, CIA is more plausible if the results indicate that the placebo outcome is not influenced by the treatment variable. Imbens and Wooldridge (2009) argue that if the variables used in this text are closely related to the outcome of interest, the test has more power. The placebo outcomes used for this exercise are a child's birth weight, head circumference, and crown-heel length measures. Given that breastfeeding occurs posterior to the birth of the child, the only way in which breastfeeding could be associated with any of these outcomes would be the presence of selection effects, i.e., the presence common factors that affect the placebo outcome and the selection into breastfeeding that are not accounted for in the PSM. One such factor may be the mother's IQ, for example. If more able mothers are also the ones more likely to breastfeed, failure to account for this unobservable will lead to over-estimation of the relationship between breastfeeding and a child's cognitive development. In the Imbens and Wooldridge (2009) framework, birth weight is closely related to cognitive and noncognitive measures and therefore birth weight stands as our main pre-treatment proxy outcome.

We also use maternal characteristics as dependent variables in this placebo exercise. Variables like alcohol and smoking intake before and during pregnancy partly reflect maternal attitudes towards children and could well be related to unobservables influencing children's outcomes. It may be illuminating to consider whether the rest of observable characteristics we use as controls already control for them.

Table 9 presents estimates of the relationship between breastfeeding and the placebo outcomes, for the different methods used. For most of the placebo variables considered, the estimated effects are insignificant, thus evidencing no selection on unobservables. However, we do find a negative association between breastfeeding and the number of cigarettes smoked at the 32nd week of pregnancy, with mothers who go on to breastfeed for four weeks smoking on average half a cigarette less per day than those who do not. This does not invalidate our methodology – it may simply demonstrate that cigarette smoking captures differences between mothers which other variables do not capture, and therefore that this variable is a useful component of our matching models. However, it may indicate a need for additional checks to the model.

Table 7: Falsification exercise.

	(1)	(2)	(3)
	OLS	PSM:ATT	PSM: ATU
Birth weight	-0.324 (7.255)	13.722 (8.849)	6.349 (11.547)
Head circumference	-0.038 (0.023)	-0.037 (0.029)	-0.023 (0.036)
Crown heel length	0.043 (0.035)	0.096 * (0.046)	0.033 (0.044)
Smoked previously	-0.002 (0.008)	-0.005 (0.009)	0.014 (0.011)
Cigarettes at 32 w	-0.524 *** (0.087)	-0.478 *** (0.093)	-0.587 *** (0.136)
Previous alcohol cons.	0.030 (0.017)	0.020 (0.019)	0.031 (0.020)
Alcoholic drinks at 8 w	-0.098 (0.088)	-0.092 (0.092)	-0.139 (0.114)

Notes:

Column 1 shows OLS regression coefficients of the different test scores on breastfeeding for at least 4 weeks. Columns 2 and 3 show ATT- and ATU- PSM estimates of the effect of breastfeeding for at least 4 weeks using the Epanechnikov kernel algorithm with 0.05 bandwidths, imposing the common support condition with the `psmatch2` Stata command (Leuven and Sianesi 2003). For each dependent variable, the rest of covariates in Table 3 are included as controls. These are the child's age and sex, the baby's birth weight, crown-heel length, and head circumference; gestation and mode of delivery (vaginal or caesarean section); whether the child is twin; the mother's age at birth; the mother's race and marital status; the education levels of both parents, housing tenure, the size of the home, neighbourhood characteristics, whether the mother had been in care as a child, whether she had divorced parents, whether her carer died prematurely; the mother's health, whether mother had smoked during pregnancy, the mother's mental health; the mother's labour market participation; and whether the mother and father had been breastfed themselves as babies and on their attitudes towards breastfeeding, measured prenatally. Standard errors in parentheses. PSM standard errors computed by bootstrapping with 100 repetitions.

Significance denoted by asterisks: * = 5%, ** = 1%, *** = 0.1%

Source: ALSPAC core sample.

4.2 Sensitivity to Potential Confounders

Since it is not possible to estimate the magnitude of selection bias with non-experimental data, in this section we address the problem by conducting a sensitivity analysis. The analysis builds on Rosenbaum and Rubin (1983b) and assumes that the CIA is not satisfied given the available observables, but would be satisfied if one could observe an additional binary variable U . Ichino et al. (2008) suggest simulating this potential confounder in the data and using it as an additional covariate in combination with the preferred matching estimator. By comparing the estimates obtained with and without matching on the simulated confounder, the robustness of the baseline results with respect to specific sources of failure of the CIA can be assessed. In particular, the distribution of the potential confounder U is fully characterized by the choice of four parameters,

$p_{ij} \equiv \Pr(U = 1 | D = i, I(Y > y^*) = j)$ with $i, j \in \{0, 1\}$, where I is an indicator function which takes value 1 when the outcome value Y is greater than the mean outcome y^* . The parameters p_{ij} give the probability that $U=1$ in each of the four groups defined by the treatment status i and the outcome value j . For example, if we consider that the unobserved factor is maternal skills, p_{11} can be defined as the fraction of skilled mothers among those who breastfed for at least four weeks and whose children's test scores are greater than the mean. A value of U is attributed to each observation, according to its belonging to one of the four groups defined by the treatment status and the outcome value. The simulated U is then treated as any other observed covariate and is included in the set of matching variables used to estimate the propensity score and to compute the ATT according to the chosen matching estimator (Epanechnikov kernel in our case). Given the probabilities p_{ij} , the simulation of U and the matching estimation is repeated many times (200 in our case) and a simulated estimate of the ATT is retrieved by averaging over the estimated ATTs. This exercise may be repeated for different distributions of U . If conclusions are relatively insensitive over a range of plausible assumptions about the distribution of U , causal inference is more defensible.

Table 10 displays the results for our sensitivity analysis.¹⁵ This analysis only makes sense if the original estimates of the treatment effects are significant, so we only perform it for the cognitive measures. For each column, the first four rows give the four probabilities p_{ij} which define the distribution of the unobservable factor U by breastfeeding and outcome statuses.¹⁶ Column (1) gives the initial matching estimates using only observables as covariates. The rest of the columns present ATT estimates for different distributions of U . Following the available literature (Ichino et al. 2008), the distribution of the binary confounding factor is assumed to mimic different observable covariates. We have chosen especially influential or potentially problematic variables such as whether the child is a twin, whether the mother is cohabiting, the mother's education level, maternal consumption of cigarettes during pregnancy, and whether the mother was breastfed as a child. For most of these assumptions on the distribution of the unobserved factor, ATT estimates remain very close to their originally estimated values. The estimated effects decrease on average by less than 1 per cent of a standard deviation when the distribution of the unobserved confounder resembles that of the variables Twin, Cohabiting, and Mother Breastfed, and by less than 2 per cent of a standard deviation when it mimics that of A-level education and Smoking at 32 weeks pregnant. However, when the distribution of the potential confounder resembles that of the variable University degree, the estimated effects are cut in half (dropping between 4 and 5 per cent of a standard deviation). Therefore, if we believe that there may be a potential unobservable with as much explanatory power as that of a university degree, our results could drop by half. Notice, however, that not

¹⁵ We adapted a Stata program developed by Nanncini (2007) to perform the calculations.

¹⁶ The actual numbers provided by the table correspond to the probabilities calculated for Maths test scores at Key Stage 1. They change slightly for each outcome measure as the distributions of the outcome variables, even though highly correlated, differ among themselves.

even a confounder as strong as the one mimicking University degree overturns our main results. Even in this case, breastfeeding still accounts for an average 5 percent of a standard deviation increase in cognitive test scores.

Table 8: Sensitivity to Unconfoundedness.

		(1)	(2)	(3)	(4)	(5)	(6)	(7)
		No confounder	Confounder like Twin	Confounder like Cohabiting	Confounder like Degree	Confounder like A-level	Confounder like Smoking at 32w	Confounder like Mother breastfed
	p11	0.00	0.01	0.15	0.23	0.30	0.11	0.60
	p10	0.00	0.02	0.18	0.09	0.26	0.15	0.55
	p01	0.00	0.03	0.19	0.05	0.19	0.20	0.38
	p00	0.00	0.03	0.25	0.02	0.11	0.29	0.36
Entry level	Reading	0.096	0.095	0.091	0.051	0.082	0.076	0.083
	Writing	0.040	0.038	0.036	0.013	0.030	0.027	0.032
	Maths	0.052	0.049	0.046	0.019	0.029	0.037	0.039
Key Stage 1	Reading	0.062	0.061	0.056	0.052	0.045	0.044	0.052
	Writing	0.076	0.076	0.069	0.035	0.055	0.057	0.076
	Maths	0.098	0.098	0.091	0.052	0.072	0.077	0.094
Key Stage 2	English	0.066	0.065	0.059	0.010	0.039	0.049	0.056
	Maths	0.103	0.105	0.095	0.056	0.079	0.088	0.092
	Science	0.117	0.117	0.111	0.064	0.089	0.105	0.105
Key Stage 3	English	0.113	0.113	0.107	0.064	0.085	0.097	0.108
	Maths	0.109	0.108	0.103	0.071	0.083	0.097	0.101
	Science	0.082	0.079	0.081	0.048	0.062	0.073	0.082

Notes:

The first four rows contain the parameters p_{ij} used to simulate the binary confounders (U) in each column. For each column, the subsequent rows contain the simulated Epanechnikov-kernel ATT estimators of the effect of breastfeeding for at least four weeks when controlling for U, together with all other variables in Table 3, over 200 repetitions. Column 1 shows our baseline results. In column 2 the confounder simulates the distribution of the variable Twin; in column 3, that of Mother cohabiting; in column 4, that of University degree; in column 5, that of A-level education; in column 6, that of Smoking during pregnancy; and in column 7, that of Mother breastfed as baby. Calculations performed using an adapted version of the Stata program *sensatt* (Nannicini 2007).

Source: ALSPAC core sample.

Taken together, the placebo exercise and the sensitivity analysis convey an impression of considerable robustness of our baseline matching results. On the one hand we find no effect of breastfeeding on birth weight, our preferred pre-treatment proxy outcome. On the other, when introducing simulated confounders, in most cases our baseline results show virtually no changes. All this supports our view that our results are not due to unobserved factors. This is not surprising given the high quality data that we use. Only when the distribution of

the simulated factor resembles that of University degree do we find a significant drop in estimated effects. Therefore the possibility that an unobserved confounder biases our estimated effects cannot be completely ruled out. However, the influence of this unobserved factor would have to be as strong as the influence of having a University degree, on both breastfeeding participation and children outcomes. Therefore if we believe that it is unlikely that another variable as important as maternal education exists, and has been omitted from our analysis, we may conclude that the main results presented in the paper are not driven by unobserved heterogeneity.

5 Alternative Definitions of Breastfeeding

A feature of particular interest in ALSPAC is the very detailed information it contains on breastfeeding. Before their babies are born, both parents are asked about their attitudes to breastfeeding and whether they had been breastfed themselves. At intervals following their child's birth, mothers are asked how they are feeding their babies, including the stages at which infant formula and solid foods were introduced. From this, we are able to compute the duration of exclusive breastfeeding; the duration of breastfeeding; and to identify babies who were fed supplementary foods generally considered to be unsuitable (for example, coffee, coca-cola, and chocolate). In this section we study first whether there is a dose response to breastfeeding, and second whether what the baby is fed matters beyond breastfeeding.

Considering breastfeeding as a dichotomous decision may be too simplistic. If the dose-response relationship is non-linear this methodology may not be adequately capturing the true impact of breastfeeding on cognitive and noncognitive outcomes. Furthermore, it may be argued that the absence of a dose-response relationship questions whether there is a causal relationship at all because if more is not better then it may be that the treatment effect is just due to correlation with some unobserved maternal characteristic.¹⁷ In addition, breastfeeding is not a monolithic activity of given duration and intensity. It is often undertaken in conjunction with some formula-feeding and, after a certain age, solid foods. In the first subsection we consider a continuous measure of breastfeeding and use OLS and the generalized propensity score (GPS) to capture the impact of length of exposure to breastfeeding on cognitive and noncognitive scores. In the second subsection we consider different indicators of exclusive and non-exclusive breastfeeding together with a measure of exposure to unsuitable foods at 6 months old.

5.1 Breastfeeding Duration and Dose-Response Functions

In this subsection, we examine the relationship between children's outcomes and the duration of breastfeeding, including a quadratic term in duration in order to capture nonlinearities in the estimated

¹⁷ We thank one of the referees for pointing this out.

relationships. Table 9 presents estimates from OLS regressions, with the first two columns showing estimates over the whole sample of children, and the third and fourth columns showing estimates over the subsample of children who were ever breastfed. The average duration of breastfeeding is about 4 months (17 weeks) across all children, and more than 5 months (22 weeks), when only breastfed children are considered.

The signs and significance levels of estimated coefficients are consistent with the results in the previous dichotomous analysis, but they demonstrate marked nonlinearities, with clear diminishing returns to duration. Results for cognitive outcomes for the full sample at Key Stages 2 and 3 translate to an effect of a little over 3% of a standard deviation from first month's breastfeeding, flattening off progressively until 8 or 12 months of age, after which there is little additional effect¹⁸.

Interestingly, the few studies which have examined breastfeeding duration tend not to allow for nonlinearities. Rees and Sabia (2009) find that an additional month of breastfeeding is associated with a 4.2 percent increase in 12 year-old children's cognitive scores with data from the US National Longitudinal Study of Adolescent Health, while Denny and Doyle (2010) report that each month of exclusive breastfeeding is associated with an increase of cognitive ability between 1.9% and 1% of a standard deviation for three- and five-year-old children in the UK Millennium Cohort Study. The estimates of Denny and Doyle (2010) would be consistent with a population average composed of a larger marginal effect at lower durations, and a lower marginal effect at longer durations.

The second column shows results for the same model estimated on the restricted sample of breastfed children. The results are similar to those for the whole sample, although the estimated effects are rather smaller – perhaps because the sample is less heterogeneous, but also perhaps because the estimates exclude the margin between children who received no breastmilk and those who received any – a marginal effect which is likely to be relatively large.

¹⁸ Our coefficients imply that cognitive scores will decrease after around 12 months' breastfeeding, but this finding is almost certainly driven by functional form: observations are concentrated at lower breastfeeding durations and hardly any mothers in the sample breastfed for over 12 months.

Table 9: OLS regressions: coefficients on breastfeeding duration and breastfeeding duration squared. Cognitive and noncognitive outcomes.

			(1)		(2)						
			All children		Only breastfed children						
			Breastf. Time	Breastf. Time Sq.	Breastf. Time	Breastf. Time Sq.					
Cognitive outcomes	Entry level	Reading	0.025 (0.008)	** (0.001)	-0.001 (0.001)	*	0.020 (0.009)	*	-0.001 (0.001)		
		Writing	0.013 (0.008)		-0.001 (0.001)		0.005 (0.009)		0.000 (0.001)		
		Maths	0.01 (0.008)		0.000 (0.001)		0.002 (0.008)		0.000 (0.001)		
	Key Stage1	Reading	0.016 (0.009)		-0.001 (0.001)		0.000 (0.010)		0.000 (0.001)		
		Writing	0.022 (0.007)	***	-0.001 (0.000)	**	0.012 (0.008)		-0.001 (0.001)		
		Maths	0.028 (0.007)	***	-0.001 (0.000)	**	0.022 (0.008)	**	-0.001 (0.001)	*	
	Key Stage 2	English	0.031 (0.007)	***	-0.002 (0.000)	***	0.020 (0.007)	**	-0.001 (0.000)	*	
		Maths	0.025 (0.007)	***	-0.001 (0.000)	*	0.020 (0.007)	**	-0.001 (0.000)		
		Science	0.038 (0.007)	***	-0.002 (0.000)	***	0.026 (0.007)	***	-0.001 (0.000)	*	
	Key Stage3	English	0.035 (0.007)	***	-0.002 (0.000)	***	0.027 (0.008)	***	-0.001 (0.001)	**	
		Maths	0.033 (0.007)	***	-0.001 (0.001)	**	0.024 (0.008)	**	-0.001 (0.001)		
		Science	0.03 (0.008)	***	-0.002 (0.001)	**	0.015 (0.009)		-0.001 (0.001)		
	Noncognitive outcomes	42 months	Emotional difficulties	0.001 (0.008)		0.000 (0.001)		0.000 (0.008)		0.000 (0.001)	
			Conduct problems	0.025 (0.007)	***	-0.002 (0.001)	**	0.018 (0.008)	*	-0.001 (0.001)	*
			Hyperactivity	0.001 (0.008)		-0.001 (0.001)		-0.012 (0.008)		0.000 (0.001)	
Prosocial			-0.013 (0.008)		0.001 (0.001)		-0.014 (0.008)		0.001 (0.001)		
Year 3		Emotional Symptoms	-0.01 (0.010)		0.001 (0.001)		-0.007 (0.011)		0.001 (0.001)		
		Conduct Problems	-0.024 (0.009)	*	0.002 (0.001)	*	-0.024 (0.010)	*	0.002 (0.001)	*	
		Hyperactivity	-0.009 (0.009)		0.000 (0.001)		-0.006 (0.010)		0.000 (0.001)		
		Peer problems	-0.015 (0.010)		0.001 (0.001)		-0.017 (0.011)		0.001 (0.001)		
		Prosocial	0.003 (0.010)		0.000 (0.001)		-0.001 (0.011)		0.000 (0.001)		

Table 9: OLS regressions: coefficients on breastfeeding time and breastfeeding time square. Cognitive and noncognitive outcomes.

		(1)		(2)		
		All children		Only breastfed children		
		Breasf. Time	Breasf. Time Sq.	Breasf. Time	Breasf. Time Sq.	
Noncognitive outcomes	Year 6	Emotional Symptoms	-0.014 (0.009)	0.001 (0.001)	-0.013 (0.010)	0.001 (0.001)
		Conduct Problems	0.005 (0.008)	-0.001 (0.001)	0.007 (0.009)	-0.001 (0.001)
		Hyperactivity	0.001 (0.008)	0.000 (0.001)	0.003 (0.009)	0.000 (0.001)
		Peer problems	0.002 (0.009)	0.000 (0.001)	0.001 (0.010)	0.000 (0.001)
		Prosocial	-0.008 (0.009)	0.000 (0.001)	-0.005 (0.010)	0.000 (0.001)

Notes:

The table shows OLS regression coefficients of breastfeeding duration and breastfeeding duration square (measured in months) on the different cognitive and noncognitive outcome measures used. Both specifications include as controls the child's age and sex, the baby's birth weight, crown-heel length, and head circumference; gestation and mode of delivery (vaginal or caesarean section); whether the child is twin; the mother's age at birth; the mother's race and marital status; the education levels of both parents, housing tenure, the size of the home, neighbourhood characteristics, whether the mother had been in care as a child, whether she had divorced parents, whether her carer died prematurely; the mother's health, whether mother had smoked during pregnancy, the mother's mental health; the mother's labour market participation; and whether the mother and father had been breastfed themselves as babies and on their attitudes towards breastfeeding, measured prenatally. Estimates in column 1 are obtained from whole sample of breastfed and non-breastfed children. Estimates in column 2 come from the restricted sample of breastfed children.

Standard errors in parentheses. Significance denoted by asterisks: * = 5%, ** = 1%, *** = 0.1%

Source: ALSPAC core sample.

In order to assess the robustness of these results we implement an alternative method: the generalized propensity score (GPS). This approach, suggested initially by Hirano and Imbens (2004), extends the usual propensity score method to the continuous treatment case to analyze the effect of treatment intensity. Formally, denoting by t a potential treatment level, we are interested in estimating the average dose-response function $\mu(t)=E\{Y_i(t)\}$. The key identifying assumption is that selection into levels of the treatment is random, conditional on a rich set of observable covariates. And just as in the binary treatment case, Hirano and Imbens (2004) demonstrate that, adjusting for the generalized propensity score (GPS) (merely a scalar function of the covariates) removes all biases associated with differences in the covariates.¹⁹

GPS methods are specifically designed to estimate the effect of different lengths of exposure to treatment. Therefore they refer to the subpopulation of treated units. We will thus use the restricted sample of

¹⁹ In essence this method consists of three steps (Imbens and Wooldridge 2009). In the first step, we estimate the GPS $r(t, x)$, that is, the conditional density of the treatment t given the covariates x . In the second step, we estimate the conditional expectation of the outcome as a function of the treatment level T and the score R : $\beta(t,r)=E\{Y|T=t,R=r\}$. In the third step, we estimate the dose-response function, $\mu(t)=E\{\beta(t,r)\}$, by averaging the estimated conditional expectation over the GPS at each level of the treatment we are interested in.

breastfed children, bearing in mind that these results may not be fully generalizable to the sample of never-breastfed children.

Figure 1 shows the dose-response functions, that is, the effect of different durations of breastfeeding on the aggregated cognitive and noncognitive outcomes considered, together with the corresponding 95% confidence bounds.²⁰ For the sake of brevity, we only show graphs for the aggregated measures. We find consistent results for each of the individual measures. Panel A shows the dose-response functions for cognitive outcomes. The effect is clearly nonlinear, with breastfeeding duration increasing test scores at a decreasing rate. The graphs show a peak at breastfeeding durations lower than those suggested in the OLS regressions: again, results towards this end of the distribution need to be treated with caution, owing to relatively small numbers of mothers breastfeeding at these durations. This aside, we are reasonably confident in claiming a positive and nonlinear relationship between breastfeeding and cognitive outcomes, which is much less pronounced at durations higher than 8 months.

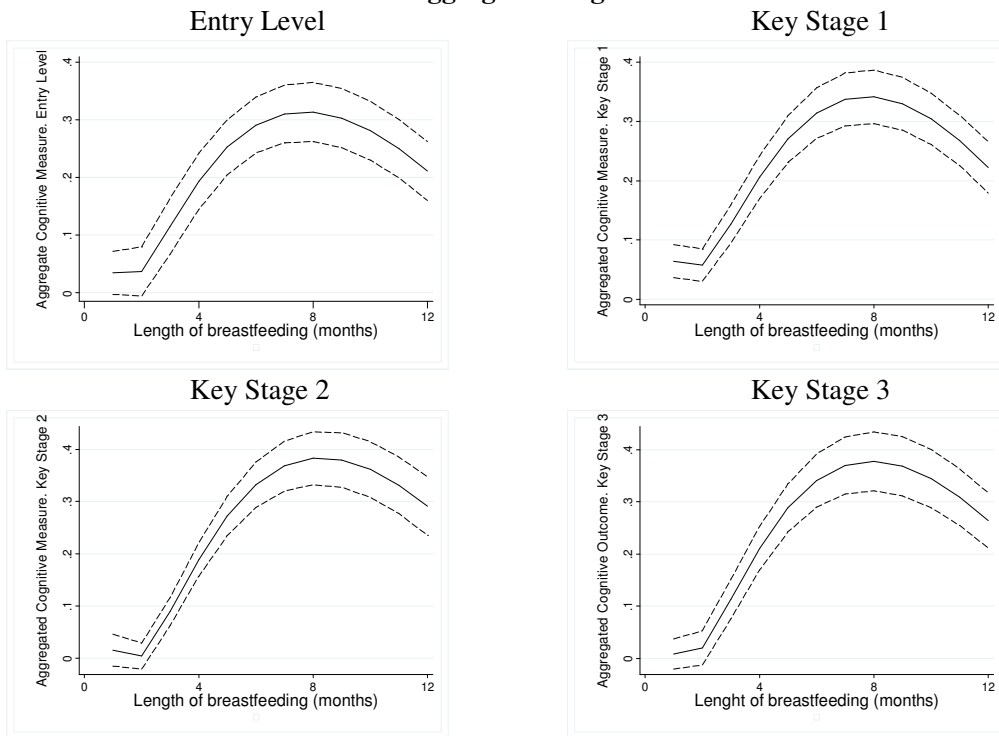
Panel B presents the dose-response functions for the aggregated noncognitive scores. Following the ALSPAC Study Team (2008a) these aggregate measures include only the emotional difficulties, conduct problems, hyperactivity, and peer problem domains, associated with worse child behaviour, and thus exclude the prosocial domain.

These results are interesting in the light of our previous estimates, which suggested very little in the way of relationships between breastfeeding and noncognitive outcome measures. Here, we see a possible explanation for this: the curves are extremely flat at the lower end (indeed in the case of the parental measures at 42 months, the curve slopes up at low durations), and it is not until two or three months' breastfeeding duration that outcomes for breastfed babies appear better than outcomes for babies fed for the minimum duration. This may be the reason that we see so few differences in outcomes when we comparing babies breastfed for less than 4 months with babies breastfed for 4 months or longer: it appears that the positive relationship between breastfeeding and the noncognitive outcomes we consider do not kick in until longer breastfeeding durations.

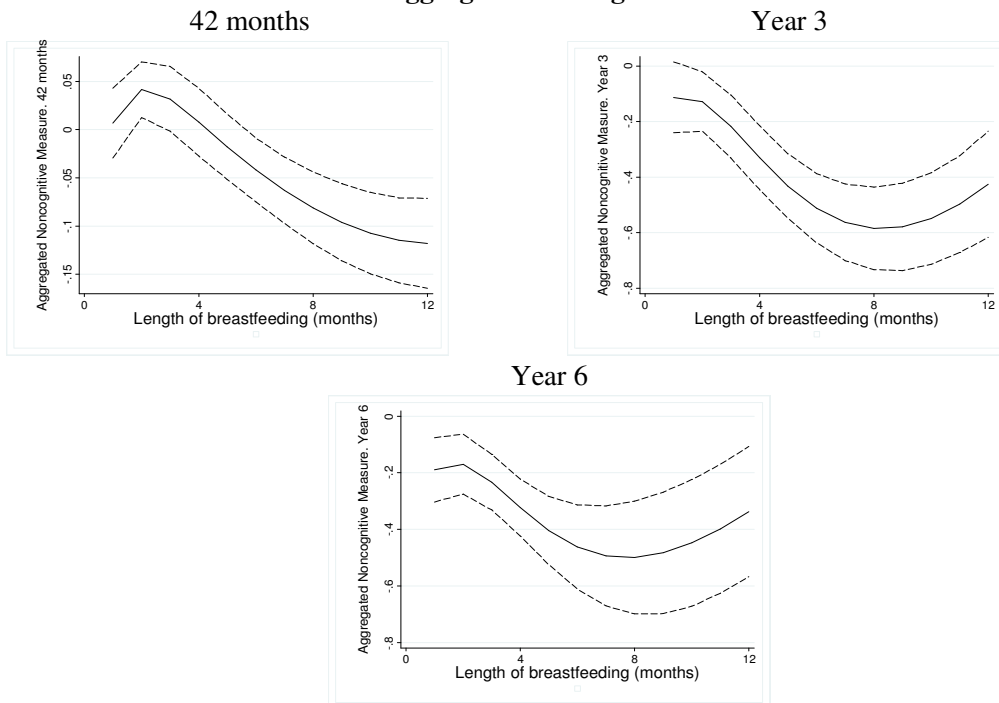
²⁰ The GPS is estimated using a normal distribution of the logarithm of the length of breastfeeding, given the covariates. The balancing property is tested using 11 strata and four treatment intervals. The potential outcome at each treatment level is estimated using a quadratic approximation of the treatment and the GPS. The dose-response function is estimated at $T=1,2,\dots,12$ months. Confidence bounds at 95% level are estimated using bootstrapping. We use the Stata programs `gpscore` and `doseresponse` (Bia and Mattei 2008) to perform the calculations.

Figure 1: Dose-response Functions for Breastfeeding Duration

Panel A. Aggregated Cognitive Outcomes



Panel B. Aggregated Noncognitive Outcomes



Notes:

The figures show the dose-response function for the corresponding outcome score and the 95% confidence bounds. All controls of Table 3 included in the estimation of the GPS. The GPS is estimated using a normal distribution of the logarithm of the length of breastfeeding, given the covariates. The balancing property is tested using 11 strata and four

treatment intervals. The potential outcome at each treatment level is estimated using a quadratic approximation of the treatment and the GPS. The dose-response function is estimated at $T=1,2,\dots,15$ months. Confidence bounds at 95% level are estimated using bootstrapping. We use the Stata programs `gpscore` and `doseresponse` (Bia and Mattei 2008) to perform the calculations.

5.2 *Feeding indicators*

In Table 10 we provide ATT results obtained through propensity score matched samples for the different breastfeeding indicators. Our first indicator (column 1) is an initiation measure, whether the child was ever breastfed. The incidence of ever breastfeeding is 75% in the sample. The prevalence of any breastfeeding at 4 weeks (our main indicator, on which all previous analyses were based, column 2) is very close to the prevalence of exclusive breastfeeding (column 3); these stand at 55% and 42% respectively. Following Belfield and Kelly (2010) we also use whether the child was formula-fed at birth (column 4). The rate is 30%, which indicates moderate overlap between breast- and formula-feeding (about 5% of all children were both breast- and formula-fed from birth). An additional indicator is a measure of whether the child was fed supplementary foods generally considered to be unsuitable when he was 6 months old (column 5): this includes children who had ever been given coffee, tea, sweets chocolate, coca-cola, and other fizzy drinks at that age. The rate of use of unsuitable foods in the sample is 27%.

For cognitive outcomes, all the estimated effects are consistent, of expected sign, and many are statistically significant, as previously found. Results for “ever breastfed” are broadly coincident with previous studies, though it should be mentioned that most previous studies report just an aggregate cognitive measure (Belfield and Kelly, 2010) or the results of pre-school tests (Quigley et al. 2011, Del Bono and Rabe 2011). Compared to those studies which report estimates of the effect of breastfeeding initiation on tests comparable to those we use, our results, ranging from 0.07 to 0.14 of a standard deviation, are closer to those of Rothstein (2011) for children 5 to 6 years old (0.06 to 0.10 of a standard deviation) than to the findings of Jiang et al. (2011) for 7-year olds (between 0.11 and 0.21 of a standard deviation), both for the United States.

Turning to the results on exclusive breastfeeding, we do not find that exclusive breastfeeding is significantly more beneficial than any breastfeeding: when we compare estimated effects at four weeks, the coefficient on exclusive breastfeeding is not always larger than the coefficient on any breastfeeding, and sometimes it is actually smaller. This is in contrast with previous findings by Del Bono and Rabe (2011) who report significantly positive effects for exclusive breastfeeding at 4 weeks, but no significant effect for any breastfeeding at 4 weeks. Note, however, that they signal problems of weak identification for this latter measure. Our results are more similar to those of Quigley et al. (2011), who report that the association between exclusive breastfeeding and cognitive scores is broadly similar to that for any breastfeeding.

The negative relationship between cognitive outcomes and having been formula-fed at birth coincide with estimates reported by Belfield and Kelly (2010) on their aggregated cognitive measure. As the measure

refers to children exclusively bottle-fed from birth, its effects are identically symmetrical to exclusive breastfeeding at birth. Since the effects of being formula-fed at birth are not always larger, in absolute terms, than the positive effects of ever being breastfed or put to the breast, this is further evidence that exclusive breastfeeding may not differ significantly from any breastfeeding in this respect. Therefore, we may concur with Quigley et al. (2011) that both exclusive breastfeeding and any breastfeeding are equally positive for children's cognitive abilities. The finding that the intake of unsuitable foods is rather weakly related to the cognitive outcomes considered, may also reinforce the hypothesis that it is any breastfeeding rather than exclusive breastfeeding which matters in this case.

It is also evident from these results that the relationship between breastfeeding and cognitive outcomes persists over the long term. In fact, the relatively increasing magnitude of the estimated effects is in line with Cunha and Heckman's (2007, 2008, 2010) hypothesis of early investment in children having multiplier effects in later childhood.

Results for noncognitive outcomes are also shown in Table 10. Again, not many of the estimated coefficients are significantly different from zero. This is consistent with the results of Kramer et al. (2008) from a large-scale, randomized experiment for the Republic of Belarus. Being formula-fed at birth shows significant harmful effects for Year 3 children, but these disappear in the longer term. The intake of supplementary unsuitable foods is also related to certain increases in behavioural problems, but only in the very short term (42 months).

Overall these results suggest that regardless of the measure used, breastfeeding shows a positive relationship with children's cognitive outcomes, and that the magnitude of this relationship increases with age. On the contrary, infant feeding seems to have no such a clear association with children's later behaviour. Our results for cognitive outcomes agree with those of Rothstein (2011) and Jiang et al. (2011), with a coefficient on breastfeeding initiation of less than one tenth of a standard deviation, and with Denny and Doyle (2010), with a coefficient on "any breastfeeding for at least four weeks" of about one tenth of a standard deviation. Unlike Del Bono and Rabe (2011), we do not find that exclusive breastfeeding differs substantially from any breastfeeding.

Table 10: Other breastfeeding measures.Cognitive and noncognitive outcomes. PSM estimates.

		(1)	(2)	(3)	(4)	(5)		
		Ever breastfed	Any breastf. 4 weeks	Exclusive breastf. 4 weeks	Formula-fed at birth	Unsuitable foods 6 months		
Cognitive Outcomes	Entry level	Reading	0.069 (0.036)	0.103 *** (0.027)	0.115 *** (0.026)	-0.066 * (0.031)	-0.037 (0.025)	
		Writing	0.031 (0.037)	0.064 * (0.028)	0.050 (0.027)	-0.059 (0.031)	-0.016 (0.025)	
		Maths	0.102 ** (0.035)	0.043 (0.027)	0.028 (0.026)	-0.070 * (0.030)	-0.016 (0.024)	
	Key Stage 1	Reading	0.082 * (0.041)	0.076 * (0.033)	0.043 (0.032)	-0.092 * (0.036)	-0.034 (0.029)	
		Writing	0.071 * (0.031)	0.076 ** (0.024)	0.052 * (0.023)	-0.074 ** (0.027)	-0.082 *** (0.022)	
		Maths	0.117 *** (0.032)	0.119 *** (0.025)	0.071 ** (0.023)	-0.120 *** (0.027)	0.003 (0.022)	
	Key Stage 2	English	0.084 ** (0.031)	0.090 *** (0.023)	0.078 *** (0.022)	-0.100 *** (0.026)	-0.082 *** (0.021)	
		Maths	0.078 * (0.032)	0.126 *** (0.024)	0.089 *** (0.022)	-0.096 *** (0.026)	-0.028 (0.022)	
		Science	0.128 *** (0.030)	0.117 *** (0.023)	0.088 *** (0.021)	-0.111 *** (0.025)	-0.005 (0.021)	
	Key Stage 3	English	0.112 *** (0.033)	0.146 *** (0.025)	0.101 *** (0.024)	-0.105 *** (0.028)	-0.114 *** (0.023)	
		Maths	0.066 (0.035)	0.110 *** (0.027)	0.055 * (0.025)	-0.080 ** (0.030)	-0.037 (0.024)	
		Science	0.141 *** (0.036)	0.127 *** (0.028)	0.084 ** (0.026)	-0.121 *** (0.030)	0.004 (0.025)	
	Noncognitive Outcomes	42 months	Emotional difficulties	-0.004 (0.037)	0.007 (0.027)	-0.012 (0.025)	0.039 (0.031)	-0.002 (0.025)
			Conduct problems	0.037 (0.035)	0.075 ** (0.025)	0.026 (0.023)	-0.017 (0.029)	0.067 ** (0.023)
			Hyperactivity	0.091 * (0.037)	-0.043 (0.027)	-0.055 * (0.025)	-0.058 (0.031)	-0.009 (0.025)
Prosocial			-0.046 (0.037)	-0.089 *** (0.026)	-0.054 * (0.025)	0.024 (0.031)	0.044 (0.025)	
Year 3		Emotional Symptoms	-0.044 (0.052)	0.008 (0.037)	0.001 (0.035)	0.061 (0.043)	-0.006 (0.034)	
		Conduct Problems	-0.005 (0.047)	-0.032 (0.032)	-0.087 ** (0.030)	0.078 * (0.037)	0.037 (0.030)	
		Hyperactivity	-0.068 (0.051)	-0.047 (0.034)	-0.062 (0.033)	0.089 * (0.041)	-0.012 (0.033)	
		Peer problems	0.053 (0.052)	-0.052 (0.037)	-0.046 (0.034)	0.095 * (0.042)	-0.032 (0.033)	
		Prosocial	0.029 (0.051)	-0.005 (0.035)	0.064 (0.034)	-0.130 ** (0.043)	0.056 (0.034)	

Table 10: Other breastfeeding measures.Cognitive and noncognitive outcomes. PSM estimates. (cont)

		(1)	(2)	(3)	(4)	(5)	
		Ever breastfed	Any breastf. 4 weeks	Exclusive breastf. 4 weeks	Formula-fed at birth	Unsuitable foods 6 months	
Noncognitive Outcomes	Year 6	Emotional Symptoms	-0.021 (0.046)	0.007 (0.034)	-0.015 (0.032)	0.033 (0.038)	-0.01 (0.031)
		Conduct Problems	-0.017 (0.041)	0.032 (0.030)	0.005 (0.029)	0.034 (0.035)	0.011 (0.028)
		Hyperactivity	0.017 (0.044)	0.034 (0.031)	-0.023 (0.030)	0.067 (0.037)	0.03 (0.030)
		Peer problems	0.024 (0.046)	0.035 (0.034)	-0.020 (0.032)	0.041 (0.039)	0.01 (0.031)
		Prosocial	-0.005 (0.047)	-0.046 (0.033)	0.025 (0.032)	-0.005 (0.040)	-0.02 (0.032)

Notes:

The table shows ATT estimates of the effect of (1) ever breastfed, (2) breastfeeding for at least 4 weeks, (3) exclusive breastfeeding for at least 4 weeks, (4) formula-fed at birth, and (5) fed unsuitable foods at 6 months. The specifications include as controls the child's age and sex, the baby's birth weight, crown-heel length, and head circumference; gestation and mode of delivery (vaginal or caesarean section); whether the child is twin; the mother's age at birth; the mother's race and marital status; the education levels of both parents, housing tenure, the size of the home, neighbourhood characteristics, whether the mother had been in care as a child, whether she had divorced parents, whether her carer died prematurely; the mother's health, whether mother had smoked during pregnancy, the mother's mental health; the mother's labour market participation; and whether the mother and father had been breastfed themselves as babies and on their attitudes towards breastfeeding, measured prenatally. We use the Epanechnikov kernel algorithm with 0.05 bandwidths, imposing the common support condition, with the psmatch2 Stata command (Becker and Ichino 2002). Columns 1 and 2 show ATT and ATU, respectively, computed in the sample of higher educated mothers (Degree and A-level). Columns 3 and 4 show the corresponding result for lower educated mothers (O-level, Vocational, and CSE).

Standard errors in parentheses computed by bootstrapping with 100 repetitions.

Significance denoted by asterisks: * = 5%, ** = 1%, *** = 0.1%

Source: ALSPAC core sample.

6 Conclusion

This paper implements a rigorous analysis of the relationship between breastfeeding and children's later cognitive and non-cognitive outcomes, attempting to disentangle the effects of breastfeeding from the effects of mother's characteristics and other observable factors. Using ALSPAC, a very rich longitudinal data set, we implement PSM techniques and a range of sensitivity analyses. We find statistically significant relationships between breastfeeding and cognitive skills at all ages between school entry and Key Stage 3 (age 14); these are very similar in magnitude regardless of the selection-on-observables method implemented (OLS, PSM-ATT and PSM-ATU). Children breastfed for four weeks or more do better than children breastfed for less than four weeks by about one tenth of a standard deviation across all the cognitive measures we examine. In fact, the coefficient becomes larger over time, with larger magnitudes found for older children. This is in

agreement with Cunha and Heckman's (2008, 2010) hypothesis of early investments in children having multiplier effects in later childhood.

On the contrary, we find much less evidence, under any of the selection-on-observables methods used, of a relationship between breastfeeding and noncognitive outcomes - although when we stratify the analysis by maternal education, we do find that breastfeeding is positively related with some of the behavioural outcomes we consider, for the group of children born to less-educated mothers.

When considering different definitions of breastfeeding and breastfeeding duration, we find that the relationship between breastfeeding and both cognitive and noncognitive development is highly non-linear, with diminishing returns to duration. With respect to cognitive skills we find no significant differences between exclusive breastfeeding and any breastfeeding.

Although we view our results as strong evidence that mother's socio-economic background cannot entirely explain the observed effect of breastfeeding on children's cognitive outputs, we acknowledge that unobserved heterogeneity may remain an issue - for example, one determinant of children's cognitive outcomes which we are not able to consider in our analysis is the mother's IQ. We note, however, that the richness of the data allows us to control for observed heterogeneity in a very precise way. Furthermore, we find that a very large amount of the heterogeneity between breastfeeding and non-breastfeeding families is captured by parental education, with all the other controls together capturing only a small proportion after education is controlled for. We contend that any remaining unobserved heterogeneity is likely to be low, as indicated by the falsification exercise and the sensitivity analysis performed to test the plausibility of CIA.

The divergence in the observed relationships between breastfeeding and cognitive development on the one hand, and non-cognitive development on the other, suggests different channels via which breastfeeding may affect different aspects of child development. It may be that the contents of breast milk are a determinant of cognitive ability, whereas contact with the mother may be more important for developing the non-cognitive skills measured here. To the extent that breastfed and bottle-fed babies have a similar degree of contact with the mother, we may not expect breastfeeding to be closely related to non-cognitive outcomes. Nevertheless further research is needed to disentangle these issues. This would require large studies collecting detailed infant feeding data, mode of receiving breast milk, and, importantly, mother's time spent with the child, feeding and not feeding the baby.

Our results provide insights into the short term and long term relationships between breastfeeding and a range of child outcomes. We have not established beyond doubt that these relationships are causal, but we have taken systematic steps to disentangle the effects of breastfeeding from the confounding effects of other factors. To the extent that the estimated relationships are causal, the relevance of this research to policy is clear. The World Health Organisation recommends exclusive breastfeeding during the first six months, and the continuation of breastfeeding alongside solid foods for two years (World Health Organisation 2003). The UK

Department of Health has identified breastfeeding promotion as a key strategy in reducing inequalities in health, and has funded several initiatives and projects which aim to increase breastfeeding rates, particularly among women from disadvantaged groups (Department of Health 2003). Given the huge social gradient in breastfeeding rates, with the most privileged mothers currently being many times more likely to breastfeed than the least privileged mothers, our results show that breastfeeding may well be a significant route for the intergenerational transmission of human capital, and for policy-making aimed at reducing inequality.

References

- ALSPAC Study Team, 2008a. Data Collected from the Questionnaires 'Child's Behaviour and Abilities', 'About Your Class', 'About Your School'. ALSPAC Documentation.
- ALSPAC Study Team, 2008b. Data Collected from the Questionnaires 'The Developing Child', 'Questionnaire for the Class Teacher', 'Questionnaire for the Head Teacher'. ALSPAC Documentation.
- Anderson J.W., Johnstone B. M., Remley D.T., 1999. Breast-feeding and cognitive development: a meta-analysis, *American Journal of Clinical Nutrition*, Vol.70, Issue 4 pp.525-535.
- Baker, M. and K. Milligan 2007. "Maternal employment, breastfeeding, and health: Evidence from maternity leave mandates", NBER Working Paper 13188. NBER, Cambridge: Mass.
- Belfield, C. R., Kelly, I.R., 2010., The Benefits of Breastfeeding across the Early Years of Childhood. NBER Working Paper 16496.
- Bia, M., Mattei, A., 2008. "A Stata package for the estimation of the dose-response function through adjustment for the generalized propensity score," *Stata Journal*, 8(3), 354-373.
- Britton, J.R., Britton, H.L., Gronwaldt, V., 2006. Breastfeeding, sensitivity, and attachment. *Pediatrics* 118, 1436-43.
- Carneiro, P. and J. Heckman 2003. Human capital policy. IZA Discussion Paper n. 821. IZA: Bonn
- Chapkin, R. S., Chen Zhao, C., Ivanov, I., Davidson, L. A. Jennifer S. et al. (2010) Noninvasive stool-based detection of infant gastrointestinal development using gene expression profiles from exfoliated epithelial cells. *Am J Physiol Gastrointest Liver Physiol* 298: G582-G589.
- Crown S, Crisp AH. 1979. *Manual of the Crown-Crisp Experimental Index*. London: Hodder & Stoughton.
- Cunha, F., J. Heckman and S. Schennach, 2010. Estimating the Technology of Cognitive and Noncognitive Skill Formation. *Econometrica* 78(3), 883-931.
- Dehejia, R., Wahba, S., 2002. Propensity Score-Matching Methods For Nonexperimental Causal Studies. *Review of Economics and Statistics* 84(1), 151-161.
- Del Bono, E., Ermisch, J. F. And M. Francesconi 2008. Intra-family resource allocations: A dynamic model of children's birth weight. IZA Discussion Papers 3704.
- Del Bono, E., Rabe, B., 2011. Estimating the Causal Effects of Breastfeeding on Early Child Outcomes. Institute for Social and Economic Research, mimeo.
- Denny K., Doyle O., 2010. The causal effect of breastfeeding on children's cognitive development: A quasi-experimental design, University College Dublin WP10/05
- Department of Health 2003. Infant Feeding Initiative. A report evaluating the breastfeeding practice projects 1999-2002. Department of Health, TSO: London

- Der, G., Batty, G. D. and I. J. Deary 2006. "Effect of breast feeding on intelligence in children: prospective study, sibling pairs analysis, and meta-analysis", *British Medical Journal*, vol. 333(7575). 945-xx.
- Duncan, G. and R. Dunifon, 1998. Soft skills and long-run labor market success. *Research in Labor Economics* 17, 123-150
- Dyson, L., Renfrew, M., McFadden, A., et al. 2006. Promotion of breastfeeding initiation and duration. Evidence into practice briefing. National Health Service, National Institute for Health and Clinical Excellence: London
- Elander J, Rutter M, 1996. Use and development of the Rutter parents' and teachers' scales. *International Journal of Methods in Psychiatric Research* 6, 63-78.
- Evenhouse E., Reilly S., 2005. Improved estimates of the benefits of breastfeeding using sibling comparisons to reduce selection bias, *Health Services Research* 40(6), 1781-1802
- Feinstein, L. 2003. Inequality in the early cognitive development of British children in the 1970 cohort. *Economica* 70, 73-98
- Fergusson D. M., Woodward L.J., (1999) Breast feeding and later psychosocial adjustment , *Paediatric and Perinatal Epidemiology* 13(2), 144-157.
- Gibson-Davis, C. M., Brooks-Gunn, J., 2006. Breastfeeding and the Verbal Ability of Three-year-olds in a Multi-city Sample. *Pediatrics* 118, 1444-51.
- Golding, J., Pembrey, M., Jones, R. and the ALSPAC Study Team, 2001. ALSPAC-The Avon Longitudinal Study of Parents and Children. *Paediatric and Perinatal Epidemiology* 15, 74-87.
- Golding, J. Bickerstaffe, I. Heron, J., Stowe, B. and Bishop, J. 2003. Data Collected From The Questionnaire 'My Son/Daughter's Health and Behaviour'. ALSPAC documentation.
- Goodman, R. 1997. The strengths and difficulties questionnaire: A research note. *Journal of Child Psychology and Psychiatry* 38, 581-586.
- Gregg, P. Washbrook, E., Propper, C., Burgess, S. 2005. The effects of a mother's return to work decision on child development in the UK. *Economic Journal* 115(501), F48-F80
- Health Promotion Agency for Northern Ireland 2006. Promoting breastfeeding for mothers returning to work. Health Promotion Agency for Northern Ireland: Belfast
- Heckman, J.H., Hotz, V.J., 1989. Choosing among alternative nonexperimental methods for estimating the impact of social programs: The case of Manpower training. *Journal of the American Statistical Association* 84, 862-874.
- Heckman, J., Ichimura, H. Todd, P., 1997. Matching as an econometric evaluation estimator: evidence from evaluating a job training program. *Review of Economic Studies* 64 (4), 605-654.
- Heckman, J., Stixrud, J. and U. Sergio, 2006. Effects of cognitive and noncognitive abilities on labor market outcomes and social behaviour. *Journal of Labor Economics* 24(3), 411-482.
- Heikkilä, K. Sacker, A., Kelly, Y. Renfrew, M. J., Quigley, M. A., 2011. Breast feeding and child behaviour in the Millennium Cohort Study. *Archives of Disease in Childhood* 96, 635-642.
- Hirano, K., and G. W. Imbens. 2004. The propensity score with continuous treatments. In *Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives*, ed. A. Gelman and X.-L. Meng, 73-84. West Sussex, England: Wiley InterScience.
- Ichino, A., Mealli, F. and Nannicini, T. (2008). 'From temporary help jobs to permanent employment: what can we learn from matching estimators and their sensitivity?' *Journal of Applied Econometrics*, No. 23, pp. 305-327.

- Imbens; G. W. and Wooldridge, J. M. (2009). "Recent Developments in the Econometrics of Program Evaluation," *Journal of Economic Literature*, American Economic Association, vol. 47(1), pages 5-86, March.
- Illsley, R. 2002. A city's schools: From inequality of input to inequality of outcome. *Oxford Review of Education* 28, 427-445
- Innis SM., 2004. Polyunsaturated fatty acids in human milk: an essential role in infant development. *Advanced Experiments in Medical Biology* 554:27-43.
- Ip, S., Chung, M., Raman, G., et al., 2007. Breastfeeding and maternal and infant health outcomes in developed countries. Agency for Healthcare Research and Quality Publication n. 07-E007.
- Jiang, M., Foster, E.M., Gibson-Davis, C.M., 2011. Breastfeeding and the Child Cognitive Outcomes: A Propensity Score Matching Approach. *Maternal and Child Health Journal* 15(8), 1296-307.
- Johnston, D., Propper, C., Pudney, S., Shields, M., 2011. Child mental health and educational attainment: multiple observers and the measurement error problem. Centre for Microdata Methods and Practice working paper CWP27/11.
- Julvezet. al., 2006. Attention behaviour and hyperactivity at age 4 and duration of breast-feeding, *Acta Paediatrica* Vol.96, Issue 6, pp. 842-847
- Kelly, Y. and R. Watt, 2005. Breast feeding initiation and exclusive duration at 6 months by social class: Results from the Millennium Cohort Study. *Public Health Nutrition* 8, 417-421
- Kramer, M. S., Chalmers, B., Hodnett, E.D. et al., 2001. Promotion of Breastfeeding Intervention Trial (PROBIT). A randomized trial in the republic of Belarus. *Journal of the American Medical Association* 285(4): 413-420
- Kramer, M.S., Fombonne, E., Igumnov, S., et al., 2008. Effects of prolonged and exclusive breastfeeding on child behavior and maternal adjustment: Evidence from a large randomized trial. *Pediatrics* 121(3), 435-440.
- Leuven, E. and B. Sianesi., 2003. "PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing". <http://ideas.repec.org/c/boc/bocode/s432001.html>
- Marshall, A., 1980. *Principles of Economics*. Macmillan.
- McCrary, C., Layte, R., 2011. The effect of breastfeeding on children's educational test scores at nine years of age: Results of an Irish cohort study. *Social Science and Medicine* 72, 1515-1521.
- Mortensen et. al., 2002. The Association Between Duration of Breastfeeding and Adult Intelligence, *The Journal of the American Medical Association* 287, 2365-2371
- Mueller, G. and E. Plug, 2006. Estimating the effect of personality on male-female earnings. *Industrial and Labor Relations Review* 60(1), 3-22
- Nannicini, T. (2007). 'Simulation-based sensitivity analysis for matching estimators', *Stata Journal*, Vol. 7, pp. 334-350.
- Duke, M.P., & Nowicki, S., Jr. (1973). Personality correlates of the Nowicki-Strickland locus of control scale for adults. *Psychological Reports*, 33, 267-270.
- Oddy, W. H., 2006. Fatty acid nutrition, immune and mental health development from infancy through childhood. In J. D. Huang (ed.) *Frontiers in Nutrition Research*. Nova Science Publishers

- Oddy W. H., Kendall, G.E., Blair, E., et al., 2010. The Long-Term Effects of Breastfeeding on Child and Adolescent Mental Health: A Pregnancy Cohort Study Followed for 14 Years. *Journal of Pediatrics* 156(4), 568-574.
- Petryk, A., Harris, S., Jongbloed, L., 2007. Breastfeeding and Neurodevelopment: A Literature Review. *Infants & Young Children* 20(2), 120-134.
- Quigley, M. A., Hockley, C., Carson, C., Kelly, Y., Renfrew, M. J., Sacker, A., 2011. Breastfeeding is Associated with Improved Child Cognitive Development: A Population-Based Cohort Study *Journal of Pediatrics* (forthcoming).
- Rees, D., I., Sabia, J.J., 2009. The Effect of Breast Feeding on Educational Attainment: Evidence from Sibling Data. *Journal of Human Capital* 3(1), 43-72.
- Rosenbaum, P., Rubin, D. B., 1983a. The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika* 70(1), 41-55.
- Rosenbaum P, Rubin D. 1983b. Assessing sensitivity to an unobserved binary covariate in an observational study with binary outcome. *Journal of the Royal Statistical Society Series B* 45 : 212–218.
- Rothstein, D., 2011. Breastfeeding and children's early cognitive outcomes. U.S. Bureau of Labor Statistics, mimeo.
- Sacker et al., 2006. Breastfeeding and Developmental Delay: Findings From the Millennium Cohort Study, *Pediatrics* 118(3), e682-689
- World Health Organization, 2003. Global strategy for infant and young child feeding. World Health Organization: Geneva
- Zhao, Z., 2008. "Sensitivity of propensity score methods to the specifications," *Economics Letters*, 98(3), 309-319.

Appendix A

Table A.1 Effect of Sample Selection on Selected Variables

	(1) Core Sample	(2) Bf sample	(3) Entry	(4) KS1	(5) KS2	(6) KS3	(7) 42 m.	(8) Y3	(9) Y6
Sample size	12268	11067	7738	9178	9964	8769	9320	5249	6018
Breastfeeding at least 4w.	54%	55%	51%	52%	54%	52%	57%	55%	54%
Breastfeeding duration	16.32	16.94	15.31	15.82	16.59	15.65	18.25	16.90	16.51
Birth weight	3406.83	3418.78	3412.35	3414.90	3421.32	3411.14	3428.13	3423.07	3424.60
Maternal education									
Degree	13%	13%	9%	10%	12%	10%	15%	13%	12%
A-level	22%	23%	21%	21%	23%	22%	24%	24%	23%
O-level	44%	41%	44%	43%	42%	43%	40%	41%	43%
Vocational	9%	10%	11%	10%	10%	10%	9%	9%	10%
No. Cigarettes, 32w preg.	1.86	1.92	2.07	2.07	1.90	2.01	1.68	1.77	1.96

Notes:

The table shows the number of observations, the proportions for dichotomous variables and the means for continuous variables used in the analyses. Column (1) corresponds to the total initial sample. Column (2) refers to the sample of observations with breastfeeding information. Columns (3) through (6) offer the calculations for the different samples of the cognitive analysis: at Entry level, Key Stage 1, Key Stage 2, and Key Stage 3. Columns (7) through (9) analyze the different samples used in the noncognitive analysis: at 42 months, Year 3, and Year 6.

Source: ALSPAC Core Sample.

Table A.2 List of control variables used in the analysis

Control Variables:
Breastfeeding variables:
<u>Main control:</u> A dummy variable (feed4wany) that takes value 1 if the child was given any breast milk at week 4 after birth.
<u>Attitudes and intentions:</u> An indicator variable (M_bf) that takes value one if the mother/father was breastfed as a baby; and a variable indicating breastfeeding attitudes of the mother and father (M_bfatt).
Age at the time of the test:
Four variables (tstEage, tst1age tst2age tst3age) that indicate the age of the child (in months) at which he or she sited the test.
Socio-demographic variables (at or during pregnancy):
Two dummies for house tenure (tenure_own and tenure_rent) that take value 1 if the mother owned the house or rented the house during pregnancy, the number of living rooms in the house during pregnancy (rooms), neighborhood indicators with higher values indicating a better neighborhood (neigh_q), a dummy indicating the mother's race (white); three dummies (mastat) indicating the marital status of the mother at the time of pregnancy (married, cohabiting, or separated/divorce), 5 dummies (M_ed and F_ed) indicating the mother's and father's education level (CSE, Vocational, O level, A level, degree), an indicator variable that takes value 1 if the mother was in local authority care (M_care),an indicator variable that takes value 1 if she had divorced parents by age 17 (M_div17), an indicator variable that takes value 1 if the mother's main carer died by age 17 (M_carer17); and an indicator variable (M_work18w) that takes value 1 if the mother was working at 18 weeks of the pregnancy.
Health, pregnancy and delivery information:
A dummy that takes value 1 if the child is a female (fem),a dummy that takes value 1 if the child is twin (twin); an indicator taking value one if the mother is in bad health status (M_badh), mother's age at birth (M_age_b), number of cigarettes a day (measured at 23 weeks of pregnancy, M_ncigs_AN32w), an indicator variable (M_csec*) that takes value 1 if the mother had a cesarean section; the length of the gestation period (gest);and two measures of the mother's mental health: the adult version of the Nowicki-Strickand locus of control scale (Duke and Nowicki, 1973) and the Total Crown Crisp score (Crown and Crisp 1979).

Table A.3. Quality of matching procedures: cognitive outcomes

	Entry Level				Key Stage 1				Key Stage 2				Key Stage 3			
	Breastfed	Not breastfed	% bias	t-test	Breastfed	Not breastfed	% bias	t-test	Breastfed	Not breastfed	% bias	t-test	Breastfed	Not breastfed	% bias	t-test
<i>Child's characteristics</i>																
Age	54.511	54.487	0.6	0.28	87.735	87.725	0.3	0.10	133.880	133.870	0.3	0.14	169.500	169.540	-1.2	-0.58
Female child	0.488	0.502	2.9	-1.23	0.438	0.453	-3.0	-1.10	0.500	0.507	-1.5	-0.74	0.505	0.505	-0.1	-0.03
Twin	0.015	0.014	0.7	0.40	0.017	0.015	1.2	0.60	0.015	0.014	0.8	0.51	0.015	0.015	0.1	0.04
<i>Pregnancy</i>																
Mother working at 18w.	0.607	0.624	-3.6	-1.60	0.607	0.633	-5.2	-2.55*	0.606	0.629	-4.6	-2.41*	0.605	0.630	-5.1	-2.44*
No. Cigarrets at 32w	1.283	1.310	-0.5	-0.30	1.522	1.660	-2.5	-1.16	1.154	1.248	-1.9	-1.24	1.199	1.305	-2.1	-1.24
Alcohol consump. Before	2.611	2.603	1	0.46	2.635	2.625	1.2	0.59	2.648	2.654	-0.7	-0.38	2.621	2.623	-0.2	-0.1
Mother bad health	1.750	1.754	-0.7	-0.33	1.756	1.764	-1.3	-0.47	1.737	1.731	1.0	0.51	1.745	1.733	1.9	0.91
Mother locus/control	3.984	3.985	0.0	0.00	4.203	4.239	-1.8	-0.68	3.863	3.877	-0.7	-0.36	3.925	3.954	-1.4	-0.69
Mother CC at 32w.	14.490	14.390	1.2	0.57	14.924	14.959	-0.4	-0.16	14.282	14.083	2.5	1.31	14.330	14.159	2.2	1.02
<i>Delivery</i>																
Mother's age at birth	29.237	29.117	2.6	1.19	29.034	28.870	3.5	1.34	29.433	29.345	1.9	1.01	29.273	29.158	2.5	1.20
Gestation	39.496	39.513	-0.9	-0.44	39.482	39.508	-1.4	-0.56	39.520	39.532	-0.7	-0.39	39.492	39.515	-1.2	-0.64
C-section	0.088	0.096	-2.7	-1.22	0.091	0.099	-2.6	-1.00	0.090	0.098	-2.9	-1.51	0.090	0.094	-1.5	-0.70
<i>Birth</i>																
Birth Weight	3436.800	3428.700	1.5	0.70	3432.600	3423.400	1.6	0.66	3447.900	3440.000	1.5	0.79	3441.400	3440.700	0.1	0.07
Head Circumference	34.825	34.803	1.6	0.72	34.817	34.797	1.5	0.57	34.862	34.843	1.4	0.73	34.833	34.823	0.7	0.34
Crownheel length	50.579	50.556	1.0	0.48	50.541	50.449	4.1	1.61	50.625	50.603	1.0	0.53	50.601	50.588	0.6	0.29
<i>Demographic variables</i>																
White mother	0.958	0.955	2.0	0.85	0.951	0.939	5.4	1.82	0.959	0.953	3.1	1.50	0.961	0.957	2.4	1.08
Mother cohabiting	0.163	0.176	-3.1	-1.42	0.165	0.171	-1.6	-0.81	0.158	0.167	-2.3	-1.26	0.161	0.179	-4.5	-2.23*
Mother single	0.019	0.018	0.3	0.18	0.025	0.020	2.5	1.16	0.019	0.020	-0.5	-0.30	0.019	0.018	0.0	0.03
<i>Socioeconomic variables</i>																
Own house	0.831	0.824	1.5	0.74	0.792	0.781	2.4	0.95	0.833	0.832	0.3	0.15	0.833	0.828	1.2	0.60
Private rented	0.053	0.053	-0.2	-0.10	0.057	0.059	-0.6	-0.22	0.057	0.055	1.1	0.54	0.056	0.056	-0.1	-0.06
Number of rooms	1.662	1.683	-2.5	-0.97	1.633	1.618	1.8	0.63	1.698	1.700	-0.2	-0.11	1.664	1.665	-0.2	-0.08
Neighbourhood qual.	8.409	8.434	-1.1	-0.52	8.300	8.376	-3.4	-1.31	8.430	8.474	-2.0	-1.06	8.434	8.484	-2.3	-1.11
<i>Psicosocial variables</i>																
Mother in care	0.018	0.023	-3.4	-1.67	0.020	0.020	0.0	0.00	0.016	0.019	-1.7	-0.97	0.017	0.022	-3.5	-1.68
Mother: divorced parents	0.135	0.139	-1.1	-0.53	0.145	0.145	0.2	0.08	0.129	0.130	-0.4	-0.23	0.132	0.135	-0.7	-0.35
Mother: carer died	0.098	0.089	3.1	1.41	0.100	0.099	0.3	0.11	0.099	0.095	1.3	0.66	0.102	0.097	1.6	0.74
<i>Mother's Education</i>																
Mother degree	0.150	0.141	3.2	1.13	0.101	0.097	1.9	0.55	0.178	0.168	3.3	1.37	0.154	0.145	3.0	1.12
Mother A-level	0.281	0.277	0.9	0.38	0.256	0.252	1.1	0.35	0.293	0.296	-0.8	-0.35	0.291	0.290	0.1	0.03
Mother O-level	0.399	0.410	-2.2	-0.99	0.422	0.427	-0.9	-0.33	0.383	0.391	-1.6	-0.83	0.402	0.410	-1.7	-0.78
Mother vocational	0.081	0.081	-0.3	-0.14	0.097	0.099	-0.8	-0.32	0.071	0.070	0.2	0.12	0.072	0.070	0.5	0.26
<i>Father's Education</i>																
Father degree	0.193	0.178	4.6	1.68	0.211	0.194	5.3	2.10*	0.238	0.226	3.4	1.48	0.210	0.192	5.6	2.19*
Father A-level	0.287	0.296	-2.2	-0.94	0.267	0.274	-1.7	-0.59	0.283	0.288	-1.4	-0.67	0.282	0.293	-2.4	-1.09
Father O-level	0.316	0.320	-1.0	-0.43	0.349	0.367	-3.6	-1.35	0.299	0.308	-1.7	-0.91	0.316	0.322	-1.3	-0.61
Father vocational	0.079	0.083	-1.4	-0.63	0.094	0.091	0.7	0.28	0.072	0.073	-0.6	-0.32	0.077	0.079	-0.6	-0.28
<i>Breastfeeding Attitudes</i>																
Mother breastfed	0.571	0.570	0.3	0.13	0.563	0.565	-0.4	-0.16	0.570	0.566	0.9	0.46	0.563	0.563	0.0	0.01
Father breastfed	0.357	0.358	-0.1	-0.06	0.338	0.345	-1.5	-0.53	0.369	0.366	0.7	0.35	0.359	0.356	0.8	0.34
Father's Attitud to BF	16.044	16.056	-0.5	-0.24	15.993	16.009	-0.7	-0.26	16.027	16.063	-1.6	-0.81	15.973	16.028	-2.4	-1.12

Notes: Calculations performed with the pstest module in STATA (Leuven and Sianesi 2003).