ISER Working Paper Series

Breastfeeding and child cognitive outcomes: Evidence from a hospital-based breastfeeding support policy

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

Birgitta Rabe
Institute for Social and Economic Research
University of Essex

No. 2012-29
December 2012

INSTITUTE FOR SOCIAL & ECONOMIC RESEARCH

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

www.iser.essex.ac.uk

Non-technical summary

A positive association between breastfeeding and many health outcomes for children has been extensively documented in the medical and epidemiological literature over several decades. As a result, public health agencies around the world promote initiatives to increase the initiation and duration of breastfeeding. The World Health Organization recommends exclusive breastfeeding for six months, and the continuation of breastfeeding alongside solid foods for up to two years.

Recent empirical evidence also suggests that breastfeeding positively affects cognitive and non-cognitive childhood development. As rates of breastfeeding are usually higher for more educated mothers, breastfeeding could be seen as an important factor in explaining the existence of large socio-economic differences in early measures of child development. It also follows that policies aiming at promoting breastfeeding among less educated women could be very effective at increasing rates of social mobility.

The decision to breastfeed rather than formula-feed an infant as well as the duration of doing so is not random but reflects maternal characteristics and preferences. This implies that it is difficult to ascertain how much of the positive effect of breastfeeding is causal, and researchers have only recently begun to examine this issue in more detail. The aim of this paper is to isolate the true impact of breastfeeding on child cognitive development from an association which might arise because of differences between breastfeeding and non-breastfeeding mothers.

In this paper we exploit variation in breastfeeding rates brought about by hospital-level differences in breastfeeding support policies. In recognition of the fact that breastfeeding needs to be established and supported from an early stage after birth to be successful, the World Health Organisation and UNICEF launched the UNICEF Baby Friendly Initiative (BFI) program in 1991. The BFI program defines Ten Steps to Successful Breastfeeding deemed best practice for breastfeeding support at the hospital level, for example: training all health care staff in the skills necessary to implement the policy; helping mothers initiate breastfeeding soon after birth; give newborn infants no food or drink other than breast milk, unless medically indicated; encourage breastfeeding on demand.

Based on data for the UK Millennium Cohort Study, we find that that women giving birth in hospitals that participated in the BF Initiative were up to 15 percentage points more likely to initiate breastfeeding and between 8 and 9 percentage points more likely to breastfeed exclusively at 4 and 8 weeks than comparable mothers giving birth in non-participating hospitals.

We then compare the outcomes of children who were breastfed as a result of the BFI program with those of otherwise similar non-breastfed children using appropriate estimation techniques. We find significant effects of breastfeeding on cognitive outcomes throughout childhood, and in particular between ages 3 and 7. Our results indicate a positive effect of exclusive breastfeeding at 4 and 8 weeks on interviewer-administered tests measuring the childs vocabulary, word reading ability and progress in maths. These results also hold for other cognitive outcomes such as school test scores at age 7. In contrast to the previous literature we find no statistically significant effect of breastfeeding on a number of health outcomes observed in our data, but we see an improvement in child emotional development and maternal mental health.

Breastfeeding and child cognitive outcomes: Evidence from a hospital-based breastfeeding support policy*

Emilia Del Bono[†], and Birgitta Rabe[‡]

December 14, 2012

Abstract

This paper estimates the causal effects of breastfeeding on early child development using exogenous variation in breastfeeding support policies across UK maternity hospitals. Based on data from the Millennium Cohort Study, we find that mothers giving birth in hospitals where such policies are implemented are between 8 and 9 percentage points more likely to breastfeed exclusively at 4 and 8 weeks than mothers who give birth in other hospitals. The effect of breastfeeding are found to be large and positive on many different measures of child cognitive development throughout early childhood. In contrast to the previous literature, we find no statistically significant impact of breastfeeding on a number of health outcomes, but we see an improvement in child emotional development and maternal mental health.

Keywords: breastfeeding, child outcomes, hospital policies, instrumental variables
JEL classification: J13, C26, I18

---

*We would like to thank UNICEF UK Baby Friendly Initiative for giving us access to their records, Suzanne Bartington, Lucy Griffiths and Carol Dezateux for comparing with us their data on the Baby Friendly status of hospitals, Jane Thomas and Shantini Paranjothy for making available to us data from the 2001 National Sentinel Caesarean Section Audit. We also would like to thank Jon Johnson of the CLS for his assistance with the UK Millennium Cohort Study data and the hospital identifiers. We received very useful comments and feedback from our colleagues at ISER, particularly Maria Iacovou and Steve Pudney, as well as Sonia Bhalotra, Helmut Rainer, Almudena Sevilla-Sanz, Andrea Weber and participants in the Forum on Greater Gender Equality at the University of Cambridge 2010, the European Society for Population Economics conference in Hangzhou 2011, the European Association of Labour Economists conference in Cyprus 2011, the Health and Human Capital workshop at ZEW Mannheim 2011, the Center for Market and Public Organization (CMPO) at the University of Bristol 2011, and the Economics Department at the University of Mannheim 2011. All mistakes are the authors' sole responsibility. This work was funded by the ESRC under grant RES-062-23-1693.

[†]ISER, University of Essex, and IZA, Bonn. Corresponding author: edelbono@essex.ac.uk

[‡]ISER, University of Essex. E-mail: brabe@essex.ac.uk

1 Introduction

A positive association between breastfeeding and many health outcomes for children has been extensively documented in the medical and epidemiological literature over several decades. As a result, public health agencies around the world promote initiatives to increase the initiation and duration of breastfeeding. The American Academy of Pediatrics (American Academy of Pediatrics, 2005) recommends exclusive breastfeeding for the first 6 months of a child's life and then continued breastfeeding throughout at least the first year. This is in line with the World Health Organization which recommends exclusive breastfeeding for six months, and the continuation of breastfeeding alongside solid foods for up to two years (World Health Assembly, 2001). More recently, statistically significant relationships have also been reported between breastfeeding and early child development indicators, spanning both cognitive and non-cognitive domains (e.g. Anderson et al., 1999; Heikkilä et al., 2011).

Research has underlined the importance of timely parental investments into child development, as differences in children's cognitive and non-cognitive development emerge at early ages and parental inputs can have a large impact (Almond and Currie, 2011; Carneiro and Heckman, 2004; Cunha and Heckman, 2008). Breastfeeding is a very early intervention, and one where there is great scope to increase uptake: in the U.S. in 2002 about 70% of mothers initiated breastfeeding in hospital, but only 33% of babies were still breastfed at 6 months (American Academy of Pediatrics, 2005). In the UK, only 28% of babies are being exclusively breastfed at four weeks of life (Bolling et al., 2007).

Not much is known about the mechanisms through which breastfeeding might affect child development. Several breast milk components have been suggested to explain the advantages held by breastfed children, most prominently long-chain polyunsaturated fatty acids. Regarding cognitive development, these are known to accumulate in the brain and retina through ingestion of breast milk and are thought to be essential in cellular differentiation and synaptogenesis of the maturing brain and retina (Petryk et al., 2007).

Development of the nervous system also depends on the amount, quality, and timing of sensory stimulation of the developing infant. Several components of the breastfeeding relationship have been suggested as enhancers of infant stimulation, such as the skin-to-skin contact involved in breastfeeding. Infant sucking also releases prolactin and oxytocin in the mother, which are thought to contribute to mothering behavior.

The decision to breastfeed rather than formula-feed an infant as well as the duration of doing so is not random but reflects maternal characteristics and preferences. This implies that it is difficult to ascertain how much of the positive effect of breastfeeding is causal, and researchers have only recently begun to examine this in detail. Some papers use sibling differences to control for unobserved family characteristics that may affect both breastfeeding and child cognitive outcomes (e.g. Belfield and Kelly, 2010; Der et al., 2006; Evenhouse and Reilly, 2005; Rees and Sabia, 2009; Rothstein, 2012). This method has its limitations in that mothers are likely to make the same breastfeeding decisions for both siblings so that identifying variation is likely to be limited, and where mothers breastfeed siblings differentially this may be related to baby's characteristics at birth that also affect outcomes. Most of the papers using within-sibling variation find no effect of breastfeeding. Several recent papers seek to identify any causal effects of breastfeeding using propensity score matching (Belfield and Kelly, 2010; Borra et al., 2012b; McCrory and Layte, 2011; Rothstein, 2012), relying on the assumption that all factors that affect selection into breastfeeding are observed. These papers all find positive but quite small effects of breastfeeding on cognitive outcomes at different ages. The only experimental study to investigate the effect of breastfeeding on cognitive outcomes is an analysis conducted in the Republic of Belarus where maternity hospitals were randomly selected to participate in a breastfeeding promotion intervention (Kramer et al., 2008, 2001). Here the authors find positive and significant effects of eclusive breastfeeding at 3 months on various IQ measures, with estimates ranging from 0.2 to 0.5 of a standard deviation.

Previous papers using instrumental variables techniques have relied on caesarean sections, mother's smoking and alcohol consumption at the individual level; or state and county-level variation in factors such as laws about breastfeeding in public and the prevalence of establishments related to health care and social assistance (Belfield and Kelly, 2010; Denny and Doyle, 2010; Rothstein, 2012). However, it is questionable whether the individual-level variables are excludable from the main equation of interest (Belfield and Kelly, 2010)[1] and the area-level instruments often turn out to be weak (Rothstein, 2012).

This paper paper overcomes these difficulties by using exogenous hospital-level differences in breastfeeding support offered to mothers after birth to estimate the causal effects of breastfeeding on child cognitive development. In recognition of the fact that breastfeeding needs to be established and supported from an early stage after birth to be successful, the World Health Organisation and UNICEF launched the UNICEF Baby Friendly Initiative (BFI) program in 1991.[2] We use differences in the maternity hospital's participation in the BFI program for identification. Because hospital's participation in the program was non-random, we use the distance between the mother's address and the nearest hospital implementing the BFI program as our preferred version of the instrument, assuming that residential selection occurs independently of hospital's BFI participation. We provide evidence about the credibility of the underlying assumptions in a dedicated section of the paper.

The UNICEF BFI program has been implemented in 134 countries around the world, and there are a number of studies evaluating its effectiveness. Most studies show a sizeable and statistically significant increase in breastfeeding initiation and duration among women in treated hospitals (Broadfoot et al., 2005; Cattaneo and Buzzetti, 2001; Dulon

[1]For example, there is empirical evidence that caesarean delivery is correlated with a range of child health outcomes such as asthma (Thavagnanam et al., 2008), type I diabetes (Cardwell et al., 2008) and diarrhoea. Recent research indicates a relationship with psychopathological problems (Li et al., 2011). Moreover, it is likely that mothers who choose to have a caesarean section (elective caesarean) have characteristics that are correlated with child outcomes, such as higher propensity for post-natal depression.

[2]The BFI program defines Ten Steps to Successful Breastfeeding deemed best practice for breastfeeding support at the hospital level, for example: training all health care staff in the skills necessary to implement the policy; helping mothers initiate breastfeeding soon after birth; give newborn infants no food or drink other than breast milk, unless medically indicated; encourage breastfeeding on demand.

and Kersting, 2003; Kramer et al., 2001; Merten et al., 2005). There is only one study on BFI effectiveness for the UK, based on the same data as this paper (Bartington et al., 2006). This study does find an effect of BFI accreditation on breastfeeding initiation, but not on breastfeeding at one month, so that there seems to be no effect on duration of breastfeeding after controlling for personal characteristics and country of residence. After controlling for region, area deprivation and hospital characteristics we do find statistically significant effects of BFI participation on breastfeeding initiation and incidence at four and eight weeks.

This paper makes several important contributions to the existing literature on skill formation and early child development. The most important is that we comprehensively address the issue of causality. We do this by exploiting the UNICEF Baby Friendly Initiative as a source of identification of the effect of breastfeeding. This, we will argue, is more convincing than using individual-level or area-level variables. For our identification strategy to hold, we have to consider the possibility that the adoption of these policies was not random. Therefore we take into account any systematic differences across hospitals that did or did not participate in the BF Initiative by combining survey data with data collected for a medical audit of hospitals. These hospital level data which have not previously been exploited for research allow us to comprehensively control for a wide range of hospital characteristics. We conduct a series of checks on the instrument and the model specification to satisfy ourselves of the robustness of our results. We also add to the previous literature by looking at a wide range of cognitive outcomes. These include interviewer-administered ability tests in different domains, assessed between ages 3 and 7, which we complement with school readiness tests as well as assessments and tests taken in school at ages 5 and 7. We also consider the effect of breastfeeding on non-cognitive and health outcomes.

We find significant effects of breastfeeding in the cognitive domain throughout childhood, as our results indicate a positive effect of exclusive breastfeeding at 4 and 8 weeks of the infant's age on several British Ability Scale measures. These results also hold for other cognitive outcomes as well as across alternative specifications of the instrument and selection of the estimation sample. In contrast to the previous literature we find no statistically significant effect of breastfeeding on a number of health outcomes observed in our data, but see an improvement in child emotional development and maternal mental health.

Another contribution of this paper is that we offer an evaluation of the UNICEF Baby Friendly Initiative. In contrast to a previous UK study(Bartington et al., 2006) we control comprehensively for area and hospital level characteristics and find that this breastfeeding support program was successful not only at stimulating breastfeeding initiation among mothers, but also at sustaining breastfeeding for longer periods. We find that women giving birth in hospitals that participated in the BF initiative were 8 to 9 percentage points more likely tobreastfeed exclusivelyat 4 and 8 weeks than comparable mothers giving birth in non-participating hospitals.

The remainder of the paper is structured as follows. The next section gives an overview of the data we use. Section 3 discusses the empirical strategy and gives details on our estimation method, provides evidence on the validity of our identification strategy and reports the first stage results. Section 4 contains the results on the effects of breastfeeding on cognitive child outcomes. It also reports robustness checks and discusses heterogeneity, as well as results obtained for other outcomes. The final section concludes.

2 Data

2.1 The Millennium Cohort Study

Our analysis is based on the UK Millennium Cohort Study (MCS), a longitudinal birth cohort study of infants born between September 2000 and August 2001 in England and Wales, and between November 2000 and January 2002 in Scotland and Northern

Ireland. The study offers detailed information about the infant, maternal breastfeeding and repeated measurements of child cognitive outcomes as well as information on a range of socio-economic characteristics of the family. This makes this study particularly well-suited to our purposes.

The MCS sampling frame is based on the UK electoral wards. The sample is clustered geographically and disproportionately stratified to over-represent: (1) the three smaller countries of the UK (Wales, Scotland and Northern Ireland); (2) areas in England with higher minority ethnic populations in 1991 (where at least 30 per cent of the population were Black or Asian); and (3) disadvantaged areas (drawn from the poorest 25 per cent of wards based on the Child Poverty Index). A list of all nine month old children living in the sampled wards was derived from Child Benefit records provided by the Department of Social Security. In the UK Child Benefit claims cover virtually all of the child population except those ineligible due to recent or temporary immigrant status.

The first wave of data collection took place when the infants were about 9 months old and includes data on 18,818 children in 18,552 families. Subsequent information was collected when the children were about 3, 5 and 7 years old. During each sweep, the interviewers administered physical and cognitive assessments, while the mother (usually the main respondent) was asked to report about the socio-economic circumstances of the family as well as the child's health and emotional development.

Our sample includes all singleton children of mothers interviewed for the first time at 9 months and where the mother was the main respondent (18,143 observations). From this sample we exclude children born at home (less than 2%), and children born in Scotland or Northern Ireland as we do not have complete hospital information for these countries (23%). We then select only cases with non-missing information on hospital of birth, breastfeeding and any of the independent variables used in the analysis, but retain cases with missing information on some, but not all, of the outcomes. Our sample at this point consists of 10,635 children. We further remove children who were low birth weight or were born before 37 weeks of gestation (6.3%). Our final selection excludes families living more than 20 miles from the closest hospital (less than 4%) as our main instrument is based on a measure of distance between area of residence and hospital postcode and is subject to a degree of measurement error. The final sample thus consists of 9.524 mother-child pairs.

Breastfeeding

Detailed information on infant feeding practices was collected at the first interview. Mothers were asked whether they ever tried to breastfeed, the age in days or weeks at which the baby last had breast milk (if no longer breastfed), as well as the age in days or weeks at which supplements (formula milk, cows milk, other milk, solid food) were introduced. Information on the total duration of breastfeeding was not asked again until the third interview, at age 5 of the child, but by 9 months most breastfeeding mothers had already ceased to do so.

Figure 1 gives an overview of feeding practices in our sample of MCS infants. It shows the proportion of babies fed any breast milk, formula, cows or other milk as well as solid foods from birth to 9 months of age. There is a marked decline in the proportion of children who are receiving *any* breastmilk during the first month of life, with a corresponding increase in the proportion of formula-fed babies.[3] After the first month the changes in feeding patterns are much smoother over time, with the incidence of any breastfeeding gradually declining to below 20% and the incidence of formula feeding increasing to almost 80%. The figure also shows that solid foods were introduced by many mothers at 4 months, which is in line with the recommendations in place in the UK at the time (Department of Health, 1994).[4]

[3]Note that there is a decline of almost 20% points in the incidence of any breastfeeding, while the percentage of formula fed infants increases by 10%. This is because the largest decline in breastfeeding occurs among mothers who do not breastfeed exclusively and who have introduced formula early on.

[4]The World Health Organization (WHO) undertook a systematic review on the Optimal Duration of Exclusive Breastfeeding, which was completed in April 2001. The WHO expert committee recommended that exclusive breastfeeding for six months was not associated with growth faltering at the population level, and was beneficial to infant health in both developing and industrialised countries. The conclusions and the recommendations of the WHO review on the duration of exclusive breastfeeding were incorporated into the WHO Global Strategy for Infant and Young Child Feeding (World Health Assembly, 2001). The WHO strategy was endorsed by the UK Scientific Advisory Committee on Nutrition (SACN) in September 2001.

In measuring breastfeeding three dimensions are usually distinguished in the literature: breastfeeding initiation, duration and prevalence or exclusivity. In addition to breastfeeding initiation, which we define as breastfeeding for at least one week, we consider (i) any breastfeeding at 4 weeks, (ii) exclusive breastfeeding at 4 weeks, (iii) any breastfeeding at 8 weeks and (iv) exclusive breastfeeding at 8 weeks. We chose to focus on 4 weeks because Figure 1 shows that 4 weeks marks a turning point in infant feeding with the sharpest decline in breastfeeding and the steepest increase in formula feeding taking place. Another advantage of choosing breastfeeding at 4 weeks is that this circumvents the issue of solid foods, as a low proportion of infants are given solid foods at that age. Finally, our measure is comparable to similar measures used in the UK literature on the subject (Borra et al., 2012b). We perform our analyses also using breastfeeding at 8 weeks in order to check the robustness of our results. We decided against considering longer durations as we found little evidence that the breastfeeding support policies operating at the hospital level we exploit for the identification of the effects of interest had a long term impact on feeding patterns.

Figure 2 shows how rates of breastfeeding initiation, any breastfeeding, and exclusive breastfeeding at 4 and 8 weeks vary across the English regions and Wales and over time within our sample. Overall, breastfeeding rates tend to be higher in England than in Wales, and higher in the South and the East of England than in the North or the Midlands. In all the observed parts of the UK rates of breastfeeding decline rapidly over time, but there is a lot of variation in the rate of decline. Some regions – such as the Eastern region – experience a very steep decrease in breastfeeding rates by 8 weeks, while in other regions – such as the South East region – there is more persistence over time. We also observe geographic differences in the relationship between exclusive breastfeeding and any breastfeeding. Most notable is the situation in London, where we see the highest rates of any breastfeeding in the whole sample, but where levels of exclusive breastfeeding are comparable to those in the South East and the South West.

Cognitive outcomes

The MCS records a number of standard tests of cognitive development. These are mainly taken from the British Ability Scales (BAS). The BAS are a set of standard age-appropriate tests of cognitive abilities and educational achievements suitable for use with young children (Elliott et al., 1996, 1997). The MCS offers information on: the BAS Naming Vocabulary test at ages 3 and 5, the BAS Picture Similarity test at age 3, the BAS Pattern Comprehension test at ages 5 and 7, and the BAS Word Reading test at age 7. The Naming Vocabulary test is a test where children are shown pictures of objects and are asked to identify them. In the Picture Similarities test the child is shown a row of four pictures and is given a card with a fifth picture, the child then places the card under the picture with which the card shares an element or concept. In the Pattern Construction test the child is asked to construct a design by combining flat squares or solid cubes with black and yellow patterns on each side. In the Word Reading test the child reads aloud a series of words presented on a card.

In addition to these tests, the survey children were assessed at age 3 according to some of the components of the Bracken School Readiness Assessment (Bracken, 2002). The score was derived from the total number of correct answers to six sub-tests, including: colors, letters, number/counting, sizes, comparisons and shapes. This indicator is thought to be directly related to early childhood education and to predict readiness for more formal education. We also use information obtained at age 7 from a variant of the National Foundation for Educational Research Progress in Maths (PiM) test in which a range of tasks covering numbers, shapes, space, measures and data are assessed.

In each case the tests were administered by Computer Assisted Personal Interviewing (CAPI) by interviewers who were specifically trained, but did not have a psychology background. Where appropiate, our analysis uses age-adjusted ability scores, which reflect the raw score and the difficulty of the items administered. All the scores are converted into z-scores, with mean zero and standard deviation one. The density plots shown in Figure 3 indicate that in most cases the test scores follow a normal distribution although there are differences across tests and ages of the child. The main exception is represented here by the Progress in Maths score, which appears to be truncated to the right.

Characteristics of mothers and children

Information on the mother's characteristics, the pregnancy, and delivery is derived from the first wave of the survey, when the infants were approximately 9 months old. The maternal characteristics we consider include: mother's age, educational qualification, annual family income (split into the categories 0-10,400; 10,400-20,800; 20,800-31,200; over 31,200 GBP), smoking status during pregnancy, partnership status at birth, and type of delivery (vaginal, assisted, planned caesarean, elective caesarean). Information on the child includes his or her birthweight, gestation, ethnicity, gender and parity. In all our regressions we also control for age of the child (in days) at the interview. Moreover, and as we will discuss below, we add a set of area characteristics, which include the region of the hospital of birth as well as the level of deprivation of the mother's neighbourhood.[5]

2.2 The UNICEF Baby Friendly Initiative

For the purposes of this paper we need to establish whether or not maternity hospitals had implemented the Ten Steps to Successful Breastfeeding of the UNICEF Baby Friendly Initiative at the time of birth of MCS study children. We use the staged process under which hospitals received certification and then accreditation with the BFI program as measures of hospital implementation of the program.

The first stage towards BFI accreditation is a certificate of commitment which is given to maternity units that have (i) adopted a breastfeeding policy, (ii) developed an action plan, (iii) had an action planning visit from UNICEF Baby Friendly and (iv) formal mechanisms for recording breastfeeding statistics. The full accreditation as Baby Friendly hospital is awarded when the Ten Steps are fully implemented. Full

[5]Neighborhood deprivation is assessed using Indices of Multiple Deprivation at the level of Lower Level Super Output Areas, a statistical geography comprising roughly 1500 households. Indices of Deprivation describe deprivation in different domains such as employment, education, health, and living environment.

accreditation is subject to annual audits of compliance, re-assessment 24 months after the initial award and collection and submission of breastfeeding statistics.

Data on the BFI accreditation and certification status of maternity units in the UK was collected from historic records of the UNICEF UK Baby Friendly Initiative office. The records show the dates at which maternity units received certification or accreditation respectively, and for hospitals that lost either of this status, the dates at which it was withdrawn. We assume that a hospital that received certification on a particular date will have already been on the way to implementing breastfeeding friendly policies before this date. Likewise, hospitals moving from certification to full accreditation will have fully implemented the policies before the accreditation date, which will also depend on the availability of UNICEF UK to visit the hospital. Therefore we follow Bartington et al. (2006) in assigning BFI certification and accreditation status, respectively, if that status was awarded by the midpoint of the birth period relevant for inclusion of infants into the MCS. This is 1 March 2001 for England and Wales, the two UK countries included in our sample. If the status was not awarded until this midpoint, the births were coded as not in a hospital with the respective BFI status.

Table 2 shows the BFI status of MCS infants' hospital of birth separately for England and Wales. Roughly 4% of maternity units were fully accredited at the relevant time, and 3% of babies were delivered in these hospitals. 24% of maternity units had certification status (34% of births), and in total 28% of maternity units (37% of births) were participating in the BFI program, either in the certification or the accreditation stage. The BFI status varies between the countries, with maternity units in Wales having a higher participation in the UNICEF UK Baby Friendly Initiative than in England.

2.3 Hospital characteristics

Detailed hospital-level variables are of particular importance for this analysis as we need to control for the possibility that hospitals with BFI certification or accreditation status had different characteristics that might affect child outcomes or might indicate the presence of selection issues. Unfortunately in the early 2000's the "maintenance of a comprehensive and accurate statistical evidence base for maternity care is $(\cdots)$ seriously impaired, not only by inadequate or nonexistent data systems and by inconsistent use of terminology, but also by a lack of IT and analytical support for maternity units." (House of Commons Health Committee, 2003a, p. 3).

We were able to trace and gain access to a data source previously not exploited for research that provides extensive information on maternity units for England and Wales for the time-period of interest. This is data collected for the National Sentinel Caesarean Section Audit (NSCSA) which was commissioned by the Department of Health to determine the current caesarean section rate and explore factors associated with variation in the Section rate and quality of care (Thomas et al., 2001). It contains data on all births within a reference period, detailed information about every caesarean delivery, as well as supplementary surveys covering midwifery, obstetric and anaesthetic issues. Moreover, there is a classification of mothers into 10 so-called Robson Groups according to clinical characteristics: parity, previous caesarean section, multiple pregnancy, presentation, gestation and labor onset. These groups represent distinct clinical risk groups that can account for differences in caesarean section rates. The data was collected in May - July 2000 in maternity units in England and Wales so that the timing corresponds very closely with the time-period of birth dates chosen for inclusion in the MCS.

Based on the NSCSA data we are able to characterize hospitals both in terms of the types of mothers and babies that typically use the hospitals and in terms of the services provided by the hospitals. Among the characteristics of mothers and babies we can distinguish the hospital-level mean maternal age; the percentage of mothers with white ethnicity, first parity, multiple births and cephalic presentation; the mode of delivery (caesarean section, instrumental and spontaneous delivery rates); percentage of infants with low birth weight and born preterm, as well as percentage of mothers in the 10 Robson Groups. Among the characteristics that describe how hospitals deliver their services we are able to distinguish the size of the maternity units according to number of births per quarter. We also use a set of 9 binary variables capturing the type of organization of midwifery care in the maternity unit. These are non-mutually exclusive variables, since a hospital could adopt a system which combines rotation and case-load schemes, for example.

Hospital characteristics from the NSCSA data set were merged to the MCS data using names and postcodes of the hospital of birth. We were able to match hospital-level data to 87% of maternity units used by mothers in the Millennium Cohort Study. Matching was successful for 98.5% of mothers, however, because unmatched hospitals were very small units with low numbers of births.

3 Empirical strategy

3.1 Estimation

In order to estimate the causal effect of breastfeeding on child cognitive outcomes we consider the following linear model:

$$Y_{ih} = \alpha + X_i'\beta + W_h'\gamma + \lambda_h + \theta_i + u_{ih}, \qquad (1)$$

where Y represents the child cognitive outcome, i is the mother-child pair and h is the hospital of birth. X is a vector of variables that affect cognitive development, such as breastfeeding and maternal schooling for example, and W is a vector of observed hospital characteristics. The next two terms capture unobserved effects at the hospital level and at the individual level, while u_{ih} is the error term which is assumed to be i.i.d.

Standard models of human capital formation do not usually take into account hospital level characteristics. However, these are particularly relevant in our context because the identification of the effect of breastfeeeding on child cognitive outcomes relies on variation in breastfeeding support policies at the hospital level. It is therefore important to take into account systematic differences across hospitals that did or did not participate in the UNICEF BF Initiative. These differences might reflect hospital-level characteristics correlated with the variables in X and affecting Y and should be explicitly included in our main equation.

The first problem we need to address is that our data consists of a single cross section of children which are followed over time. This implies that we observe each hospital of birth only at one point in time. As our instrument varies at the hospital level, we cannot net out hospital level unobservables by means of a fixed-effects estimator. This implies that λ_h will be absorbed in the error term. If we write the latter as $\varepsilon_{ih} = \lambda_h + u_{ih}$, our model becomes:

$$Y_{ih} = \alpha + X_i'\beta + W_h'\gamma + \theta_i + \varepsilon_{ih}. \tag{2}$$

The second problem we have and which is common to most studies in this area, is that although we might observe a rich set of individual variables there is always the possibility that some mother or child specific relevant characteristics are omitted from the model. If these unobservables are also correlated with the included regressors, then our estimates of the parameters in (2) will be biased. In the absence of observations on siblings, a mother fixed-effects strategy is not feasible and θ_i will inevitably end up in the error term, $\phi_{ih} = \theta_i + \varepsilon_{ih}$, so that $Cov(X_i, \phi_{ih}|W_h) \neq 0$.

In these circumstances a standard solution is to find a set of instruments correlated with the endogenous regressors and orthogonal to ϕ. Suppose for the sake of exposition that the vector X consists only of three sets of variables, B maternal breastfeeding, a vector of exogenous variables (such as the sex of the child), X_1, and a vector of mother or child characteristics which could be potentially endogenous (such as mother's schooling), X_2. In order to estimate the parameters of (3) we would need to find at least as many instruments as the number of potentially endogenous variables.

$$Y_{ih} = \alpha + \beta_0 B_i + X_{i1}'\beta_1 + X_{i2}'\beta_2 + W_h'\gamma + \phi_{ih}. \tag{3}$$

Usually researchers can rely on one plausible instrument for the endogenous variable of interest, occasionally there will be more than one instrument for the same single endogenous variable, but in most cases there are not enough instruments for all the endogenous variables in (3). This leaves us with the question of what to do with the endogenous regressors for which we cannot find a suitable instrument. Most of the literature in this area includes these other regressors in the model and focuses exclusively on the endogeneity of breastfeeding, effectively treating the other regressors as exogenous.

Here we follow a different approach and exclude all the other endogenous regressors from the main equation. The model we estimate can be specified as follows:

$$Y_{hi} = \alpha + \beta_0 B_i + X_{i1}' \beta_1 + W_h' \gamma + \nu_{ih}. \tag{4}$$

where $\nu_{ih} = \phi_{ih} + X_{i2}' \beta_1$, so also includes the vector of endogenous variables, X_2. Note, however, that although we do not control for individual-level endogenous variables, we do control for a number of maternal characteristics such as maternal age and type of delivery at the hospital level, thereby taking into account some heterogeneity across mothers.

Performing OLS on (4) will clearly lead to inconsistent estimates of the effect of breastfeeding on child cognitive outcomes, as many of the elements in X_2, such as mother's schooling, will be important determinants of child cognitive outcomes and will also be strongly correlated with breastfeeding, so that $Cov(B_i, \nu_{ih} | X_{i1}, W_h) \neq 0$. However, if we had a variable Z that is correlated with breastfeeding and is uncorrelated with the error term, ν_{ih}, then the effect of breastfeeding estimated via IV will be consistent despite the fact that we have excluded from the main equation potentially important determinants of child cognitive outcomes.

In our case this variable Z is given by the participation of the hospital of birth in the UNICEF BF Initiative. The whole identification strtegy relies on $Cov(Z_h, \nu_{hi} | X_{i1}, W_h) = 0$. In particular, we need to assume that $Cov(Z_h, \lambda_h | X_{i1}, W_h) = 0$, or that after controlling for observable hospital characteristics there is no residual correlation between the BFI status of the hospital of birth and unobserved hospital effects. We will also assume that $Cov(Z_h, X_{i2} + \theta_i | X_{i1}, W_h) = 0$, which means that that there are no omitted individual level variables which are relevant for explaining child cognitive outcomes and are correlated to the BFI status of the hospital of birth.

3.2 Evidence on the validity of the identification strategy

As discussed above, two conditions are required for our identification strategy to be valid. First, we need to be confident that the BFI status of hospitals is unrelated to unobserved hospital-level characteristics that could directly affect child outcomes. We show below that there are almost no significant differences across hospitals according to their BFI status when we look at a large set of hospital-level variables. Second, we need to satisfy that the BFI status of the birth hospital is uncorrelated with a set of observed - but possibly endogenous - and unobserved mother and child characteristics that may directly influence child outcomes independently of the effect of breastfeeding. We provide support for this assumption by arguing that UK mothers had very little or no choice about the hospital of birth during the period covered by this study and by showing that there is very little correlation between BFI status of the hospital of birth and a large set of mother and child characteristics which are usually thought to be relevant determinants of child cognitive outcomes, such as maternal education and child birth weight.

3.2.1 Selection of hospitals into the BFI program

As was shown in Table 1, only some maternity units in England and Wales were participating in the UNICEF BFI program in the period of interest. Unfortunately there is not much information available on the likely motivation of hospitals to sign up to the program or to refrain from doing so. According to UNICEF UK, hospitals with a determined infant feeding advisor or in areas of existing good breastfeeding practice may have chosen to become Baby Friendly, but sometimes awareness of low breastfeeding rates triggered the initiative to join the BFI program.[6]

In order to understand whether hospitals that participated in the BFI program were systematically different from non-participating hospitals we need to examine the available evidence. Table 2a shows whether hospitals differ systematically according to their BFI status. As we can see, hospitals that hold a certificate of committment or are fully accredited are on average larger than hospitals that did not participate, although this difference is very small and not statistically significant. There is also some indication that average maternal age is lower in certificated or fully accredited maternity units than in non-participating hospitals, but again this difference is not very large. We analyse other maternal characteristics, such as her ethnicity and previous parity status, but no significant differences emerge.

The next set of variables explores the possibility that hospitals that participate in the BFI are those that provide more specialised care. So, we compare the percentage of multiple births, the incidence of the most common type of presentation at labour (cephalic), the incidence of the most common type of labour onset (spontaneous), the incidence of caesarean sections, as well as the percentage of low birthweight babies and preterm babies born in the hospital. In none of these cases we find any evidence of systematic differences across hospitals according to BFI status.

Finally, we look at a set of variables that describe the type of midwifery care available at the hospital. We have reason to think that these variables could be important in our analysis. For example, the staffing model used to organize midwifery care could impact the extent to which one-to-one care is available in the labor ward, and this has been shown to reduce the rates of some obstetric interventions in labor (e.g. Page et al., 1999). However, Table 2a shows that although some of the differences across fully accredited, certified and non-participating hospitals appear large in magnitude, these are

[6]Personal conversation between the investigators and the director of the UNICEF BFI program in the UK.

not statistically relevant.[7] Overall, since we find no systematic differences across a wide range of observed hospital characteristics according to BFI status, we think it is unlikely that there are omitted hospital level effects correlated with our instrument.

3.2.2 Selection of mothers into hospitals with different BFI status

Our second key identifying assumption is that the BFI status of the birth hospital is uncorrelated with observed – but possibly endogenous – and unobserved mother or child characteristics that may directly influence child outcomes. Stated differently, we need to assume that assignment of the mother-baby pair to the treatment group (delivery in a hospital with BFI certification or accreditation) or the control group (delivery in a hospital with no BFI status) is as good as random. The assumption would be violated if women selected into hospitals according to the hospital's breastfeeding-friendly status, and for example if women with higher levels of education or more knowledge about the benefits of breastfeeding chose the give birth at a Baby Friendly hospital. Although this seems a plausible scenario for women in the UK today, we argue that the situation in the relevant time-period put serious restrictions on the choice of the hospital of birth.

The possibility for pregnant women to select into a hospital of choice will depend on (1) information on the main characteristics of hospitals as basis for any choice (2) reasonable access to more than one maternity unit. In 2002/2003 the House of Commons Health Committee published two reports on the provision of maternity services and on choice in maternity services which highlight problems in both these areas (House of Commons Health Committee, 2003a,b). Regarding information on hospital characteristics that could form the basis for any choice of women the situation was particularly

[7] In our set of preliminary analyses we also controlled for a much wider range of hospital-level observable characteristics, including whether the hospital provided ante-natal classes or clinics, and whether it complied with a set of so-called *auditable standards* specifically related to maternity care. The latter could be interpreted as indicators of the hospital propensity to follow the best and more up-to-date practices. We could never detect any significant differences in any of these variables according to the BFI status of the hospital. As most of these variables showed no relationship with either breastfeeding or later child outcomes, we decided to adopt a parsimonious specification, and exclude them from our main specification.

difficult in the time-period of interest (2000-2002). The House of Commons Health Committee (2003a) dedicated a large section of their report to investigating why there were no reliable hospital-level statistics on the most fundamental hospital characteristics available. It urged for more information to be made available to mothers, particularly on the caesarean section rates of hospitals and on particular consultants (p. 37). The variability of caesarean section rates was at the forefront of discussions on maternal choice at the time, as it was perceived that clinical practice varied hugely and put mothers at unequal risk of major surgery according to a so-called postcode lottery, i.e. dependent on where they live (ibid.). This seems to indicate that BFI status of hospitals was not generally known by mothers nor was it a major concern at the time. The lack of maternity statistics also prompted the organization Birth Choice UK in 2001 to set up a web site collating hospital-level information and to support mothers to "start thinking about making a choice" (email from Birth Choice UK, 5 Jan 2011).

Regarding access to maternity units, the committee wrote that for the vast majority of pregnant women, the first point of contact for accessing maternity services is their GP (General Practitioner, i.e. family doctor). Although women should have had a free choice between hospital, birth center or home birth, the women were usually referred by their GPs to particular hospitals, and they found it hard to access maternity care without GP referral. Women were not offered a choice of different acute units (House of Commons Health Committee 2003b, p. 9). Moreover, the National Health Service had started, in the late 1990's, a process of centralization of maternity units as this was deemed more efficient. This restricted the number of maternity wards that could reasonably be accessed from the home within a journey time of up to 30 minutes while in active labor (Dodwell and Gibson 2009).

The proportion of women that accessed the hospital closest to their home to deliver their baby could further inform the question of whether women were able to exercise choice in hospitals, if we quite plausibly assume that women did not choose their homes to be close to BFI hospitals and if we assume that GPs systematically referred women to their closest hospital. A high proportion would indicate that there was little choice. Unfortunately we cannot derive this proportion accurately with the MCS data, as information on the location of residence is available for the first time at the interview which took place at age 9 months of the cohort child, and a sizeable number of households may have moved since having their baby as there is high residential mobility around the time of child birth. Using the distances between the residence at 9 months and the birth hospital we find that 73% of women gave birth in the hospital nearest to them.[8] This is a large proportion given the noise in the measurement and uncertainty about referral practices of GPs, and we take it as confirmation of limited hospital choice.

To investigate this issue further, we present in Table 2b descriptive statistics for a number of mother and child characteristics according to the BFI status of the hospital of birth. The first set of variables consists of (i) mother's age at the birth of the child, (ii) her level of educational qualification, (iii) the level of income of the family (if below GPB 20,800 per year), (iv) the percentage of mothers who smoked during the pregnancy, (v) the percentage of intact families, (vi) the percentage of normal deliveries, (vii) the child's weight and (viii) gestation at birth. Although these variables are important determinants of early child outcomes, if we were to include them in the model we would face additional endogeneity problems. As explained above, we will omit these variables from our analysis. This implies that these variables will be absorbed in the error term and it is therefore very important to check wether they are correlated with our instrument.

As we can see, there is no evidence of systematic differences in these variables according to whether the hospital of birth was certified, fully accredited or non-participating. The only exception is to be found for the variable gestation, where we see that children born in fully accredited hospitals are generally born earlier than children in non-

[8]The location of each maternity unit is given by its postcode which has an associated grid reference defining its location on the British National Grid. For MCS households the most detailed available geographical identifier is the Output Area. This is a small scale statistical geography containing a minimum of 40 and an average of about 125 households. Using the Output Area centroid as an approximation of the location of the household on the relevant National Grid we calculated distances using the Pythagorean formula, i.e. using a planar approximation to the surface of the earth which is justified for small distances.

participating hospitals. This might be due to a few outlier observations in the very small group of children born in fully accredited hospitals, however. Indeed, if we combine accredited and certified hospitals in a single category, there is no significant difference in gestation between participating and non-participating hospitals.

The second set of variables includes characteristics of the child which are plausibly exogenous, such as (i) ethnicity, (ii) sex, (iii) parity, and (iv) age at the interview. Some of these variables will be strongly correlated with breastfeeding, like ethnicity for example, while others will be largely independent of it, like age at the interview. In any case, Table 2b shows that there are no significant differences in the distribution of these variables between participating and non-participating hospitals.

Finally, we consider variables related to the geographic distribution of the sample. In particular we analyse whether mothers who give birth in BFI hospitals are less likely to come from disadvantaged heighborhoods (defined by the lowest five deciles of the index of multiple deprivation). As we can see, there is no evidence that this is the case. We also look at whether the average distance from the closest hospital is different for mothers who gave birth in participating or non-participating hospitals. If BFI hospitals were concentrated in metropolitan or urban areas, for example, we should observe on average shorter distances for women giving birth in participating hospitals. There is some evidence that this is the case, particularly when comparing hospitals holding a certificate of commitment with non-BFI hospitals. So, in the analysis that follows we will always include a set of controls for the geographical distribution of the sample, such as regional dummies, dummies for area of deprivation, and distance to the closest hospital. We will also perform some robustness checks to consider the potential effect of including/excluding certain areas of the country, or individuals living in close proximity to a hospital.

3.3 First stage results

Table 3 presents our first-stage estimates of the effect of the BFI status of the hospital of birth on the likelihood of breastfeeding, where the latter is defined in terms of both its duration and exclusivity. In particular, we distinguish breastfeeding initiation - breastfeeding for at least one week - from any breastfeeding and exclusive breastfeeding at both 4 and 8 weeks. The definition of exclusive breastfeeding is derived using information on breastfeeding duration and information on the timing of the introduction of formula milk, cows milk, other milk, and solid foods. The standard definition of exclusive breastfeeding usually requires information on the administration of water, juices and other liquids different from breastmilk, but this is not available in the MCS. In our analysis we should therefore consider exclusivity simply as a measure of the prevalence of breastmilk over other types of liquids.

We begin by estimating logit regressions on the outcome variable, which is always defined as binary. In the first specification we do not include any control variables and simply analyse the effect of the BFI status of the hospital of birth. As we can see, we find that there are almost no statistically significant differences in breastfeeding rates across accredited, certified and non participating hospitals. We find some evidence that mothers who gave birth in a certified hospital were slightly more likely to breastfeed exclusively at 4 and 8 weeks, but this effect is only significant at the 10% level. This result is not surprising, since rates of breastfeeding vary substantially across different parts of the UK and the reach of the BFI program was certainly not uniform across the whole country.

The picture changes after we introduce regional controls (column two). Here we see statistically significant effects of both certification and accreditation on breastfeeding initiation, some statistically significant effects of certification on exclusive breastfeeding at 4 and 8 weeks, but no effects on any breastfeeding at either of the two points in time considered. Controlling for the level of deprivation of the area where the mother lived at the 9 months interview (column three), reveals slightly larger and more precisely estimated effects of both certification and accreditation. In column four we introduce the vector of hospital-level variables described in Section 3.1. Here we see that while the effect of certification becomes slightly larger in magnitude the effect of accreditation remains the same. Adding controls for individual level characteristics that can be considered exogenous, such as sex of the child, age of the child at the interview, ethnicity, parity, and distance to the closest hospital (column five) does not change much in the estimated coefficients. This confirms what we have already seen in Table 2b, i.e. that in large part these individual characteristics are orthogonal to BFI status of the hospital of birth.

The results presented in Table 3 indicate that BFI certification had strong and significant effects on all our measures of breastfeeding. More precisely, we can say that the odds of breastfeeding at 4 weeks, for example, where about 1.5 higher for mothers who gave birth in certified hospitals than for mothers who gave birth in non-participating hospitals. The effects of accreditation are much less easy to ascertain, however. While we find that the odds of initiating breastfeeding are almost 2.5 higher in an accredited hospitals, accreditation does not significantly affect breastfeeding rates at 4 weeks, and has much weaker effects on the other measures of breastfeeding. Given that the number of fully accredited hospitals is very low in our sample (only 7 in total), and that we are relying on variation within regions, we are not surprised to see that the effect of accreditation is not very robust.

Our evidence contrasts with the results in Bartington et al. (2006), who find no effects of certification and only some effects of accreditation on breastfeeding initiation but not on breastfeeding at one month. Our analysis is however based on a different sample, which excludes Scotland and Northern Ireland, and includes regional, local area and hospital level controls. The exclusion of Scotland, which at the time our survey took place featured 4 accredited maternity units is likely to account for the fact that we find weaker effects of accreditation, and we have seen that regional, local area and and hospital level controls are important to reveal the effects of certification. We conducted an extensive range of checks of our data. Using the same specification and the same sample as in Bartington et al. (2006) we were able to replicate their results very closely, so that we are confident that our findings are not driven by differences in the attribution of BFI status or by the definitions of the main variables.[9] When entering region of birth and indicators of local area deprivation into the specification and sample used by Bartington et al. (2006) we find a statistically significant effect of the BFI certification and accreditation on breastfeeding initiation, and of BFI certification on exclusive breastfeeding at four weeks. We were unable to include hospital level variables in their sample as these are only available for England and Wales.

In the last two columns we present the marginal effects calculated from the logit model in column 5, and compare them to what we would obtain if we were to use a linear regression. As we can see, the coefficients of the dummies for certification and accreditation are very similar in both magnitude and level of statistical significance in the two columns. We take this as evidence that a linear approximation is adequate in our case, and will proceed using linear regressions, effectively ignoring the binary nature of the breastfeeding variable.[10]

In what follows, we also restrict our attention to two definitions of breastfeeding: exclusive breastfeeding at 4 weeks and exclusive breastfeeding at 8 weeks. This is because it would be very difficult to imagine that breastfeeding a child for one week only (breastfeeding initiation) could lead to substantial effects on cognitive achievements. At the same time, our measure of "any breastfeeding" includes too many different breastfeeding practices – including formula supplementation, for example – which would make the interpretation of the results very difficult.

[9] Although we have shared with the authors of the study in Bartington et al. (2006) information on the BFI status on maternity units, so that these coincide almost perfectly, there remain small differences in how individuals are allocated to maternity units. Bartington et al. (2006) carried out their own coding of maternity units, whereas we use the official information available through a special licence version of the MCS.

[10] We also estimate a models with a non-linear first stage. In particular we used the suggestion in Vella (1993, 1998) and estimate a probit for the first stage relationship, obtain the generalised residuals and then use these in the second stage equation in addition to the endogenous regressor following a control function approach. The estimates we obtained were in large part very close to what we present in this analysis.

Having thus defined our specification, both in terms of functional form and control variables, and chosen the measure of breastfeeding to focus on, we carry out some tests aimed at analysing the strength of our instrument. We compare three different ways of measuring the effects of the BFI program. First we use separate dummies for certified and accredited hospitals; then we consider a single dummy capturing the effect of the BFI program as a whole; and finally we consider distance to the closest BFI participating hospital (and its square). For each of these different versions of our instrument we present in Table 4 the F statistics and partial R-squared obtained after performing a weighted OLS regression. Our standard errors are clustered at the hospital level, therefore the weak instrument test statistics we use is the Kleibergen Paap Wald F statistics (Kleibergen and Paap, 2006).

Given that we have a very low number of accredited hospitals in our sample, we are not really in a position to distinguish different levels of participation in the initiative, and as we can see the F statistic is always much higher when we consider a single BFI dummy rather than separate controls for certified and accredited hospitals (column one and two, respectively). When we take into consideration distance to the closest BFI hospital as our instrument, the explanatory power of the variable drops somewhat in the equation measuring exclusive breastfeeding at 4 weeks, but it remains high in the equation measuring exclusive breastfeeding at 8 weeks. In all cases the value of the F-test is relatively high, or at least higher than the level which is usually considered critical in the literature which is the the 10% maximal IV size statistic from Stock and Yogo (2005). However, the partial R^2 are relatively low (about 7.5% of the total R^2) and this acts as a warning that our results might still be affected by a weak instrument problem.

If we were to choose the most precise way of measururing the impact of the BFI program on the basis of these tests, we would probably consider a single dummy capturing both certified and accredited hospitals, as the specification which makes use of this variable features the highest F statistics. However, despite the fact that women in the UK had almost no choice about the hospital of birth at the time the survey took place, and that we include a very large set of controls for hospital characteristics, one might still be worried that an instrument based on the actual hospital of birth could be endogenous. To address this concern we choose the distance to the closest BFI hospital as our preferred version of the instrument. This assumes that after we take into account residential choices (measured by distance to the closest hospital), the distance to the closest BFI hospital is orthogonal to unmeasured individual characteristics which will be captured by the error term.

4 Results

4.1 Cognitive outcomes

The main aim of this study is to investigate whether maternal breastfeeding has a causal effect on child cognitive outcomes. The earliest measures of cognitive ability available in the MCS are collected at age 3 of the study child. As we have seen before, these measures mainly consist of interviewer-administered tests which assess the verbal and comprehension skills of very young children. Similar tests are available at age 5 and 7, and this allows us to see whether the effects of breastfeeding persist over time. Each test is normalized and expressed as a z-score, so that the estimated coefficients can be interpreted as fractions of a standard deviations.

Tables 5a and 5b present our main results, showing OLS and 2SLS estimates of the effects of exclusive breastfeeding at 4 and 8 weeks, respectively, on the available measures of cognitive ability. These results indicate that OLS estimates of the effects of breastfeeding are always statistically significant and positively related to cognitive outcomes. For example, looking at the first column in Table 5a we see that infants who were exclusively breastfed for 4 weeks or longer exhibit values of the Bracken test which is 0.26 standard deviations higher than that of children who were breastfed for less than 4 weeks or not breastfed at all. Similarly, in column two of Table 5a we see that breastfeeding is on average associated with 0.17 points of a standard deviation higher BAS naming vocabulary scores at age 3. Overall, the OLS effects of breastfeeding on our measures of child cognitive development vary between 0.14 and 0.29 points of a standard deviation, depending on the outcome and definition of breastfeeding considered.

We expect the OLS effects to be biased upwards because of the omission of relevant observed and unobserved characteristics of the mother and the child. Our 2SLS estimates are however almost always much larger in magnitude than the corresponding OLS estimates. So, for example, we see that the effect of breastfeeding exclusively for 4 weeks or more on measures of cognitive ability varies between 1 and 1.5 points of a standard deviation at age 3, between 0 and 0.7 points of a standard deviation at age 5, and between 0.7 and 1.1 points of a standard deviation at age 7. It is not surprising to observe such a variability across 2SLS estimates, given that we observe very large standard errors. What is interesting, however, is that in many instances the 2SLS estimates are statistically different from zero and also statistically different from their OLS counterparts.

We look at the possibility that our outcomes are correlated with maternal characteristics. The Bracken and BAS scores analysed in Tables 5a and 5b are measured by the interviewer, but we cannot exclude the possibility that mothers might directly influence the performance of the child in the test. Therefore we consider in Table 6 the effects of breastfeeding on teacher assessments at age 5 and externally marked National Curriculum tests at age 7 that are arguably unaffected by maternal characteristics. These are the Early Years Foundation Stage Profile (EYPSP) measured at age 5 and the child's reading, writing and mathematical ability tested at age 7. [11] By analysing outcomes reported by teachers we reduce the likelihood that mothers have a direct influence on the outcomes, and therefore the possibility that omitted maternal characteristics may affect

[11] The EYFSP score sums up and describes each childs development and learning achievements at the end of their first year of school as assessed by the teacher. We express it as a z-score. The child's reading, writing and mathematical ability is tested in English schools in National Curriculum assessments that are externally marked. We use a 0/1 dummy for achieving level 3 or 4, as level 2 is the level expected at age 7. Note that both outcomes are only available for pupils in England.

the results. As we can see in Table 6, the findings are broadly in line with what we saw earlier, in that the 2SLS estimates are much larger and in some cases significantly higher than the OLS estimates. This is true in particular for the EYFSP scores and for top scores in writing, which are closely related to the the Bracken, BAS naming vocabulary and word reading skills analysed in Tables 5a and 5b.

There are at least three reasons that might explain why the 2SLS estimates are so large in magnitude. The first is that the breastfeeding variable is affected by measurement error. Usually this implies that the OLS estimates are downward biased, and the corresponding 2SLS estimates correct for this. Although there is likely a problem of measurement error in the breastfeeding variable, as this is mother-reported and affected by recollection bias, the noise to signal ratio would have to be very substantial indeed to explain the large difference between OLS and 2SLS estimates we see in our results.[12]

A second possibility is that our 2SLS estimates identify a local average treatment effect (LATE) on a subset of mothers that are induced by the presence of breastfeeding support policies at the hospital level to change their breastfeeding behaviour and for whose children breastfeeding is particularly effective. In the presence of heterogeneous effects it is possible that the LATE is higher than the average treatment effect (ATE) identified by the OLS coefficients. This might be an important element of the explanation, and we will offer more information about the heterogeneity of the effects later on.

The third explanation is that our instrument is too weak, so that even small levels of correlation between the BFI status of the maternity unit and the error term in the structural equation will result in a large bias in the 2SLS estimate. The direction of the bias would be the same as for the OLS estimates, i.e. this problem will cause the 2SLS effects to be too large. Although we have many reasons to think that maternal information about breastfeeding policies was very limited at the time of the survey, mothers

[12] There is also the non-trivial issue that measurement error in the breastfeeding variable would most likely be non-classical, so that standard results would not apply in this case.

had little choice about the hospital in which they gave birth, and we control for a large vector of hospital level variables, we cannot completely rule out some residual correlation between omitted individual or hospital level characteristics and our instrument. As the relationship between the BFI status of the maternity unit and breastfeeding is strong, but still explains only a small part of the variation in breastfeeding rates (see section 4.3), a weak instrument problem is a possible cause for concern.

The weak instrument tests presented in Table 4 are fairly encouraging. The values of the Kleibergen Paap F-statistic are not particularly high but in the chosen version of the instrument the F-statistic are above the Stock and Yogo (2005) critical value which is the 10% maximal IV size statistic. We further address the weak instrument concern by conducting two types of checks. First we look at the explanatory power of the instrumental variable in the reduced form equation. When the correlation between the instrument and the error term is small, the reduced form should give approximately unbiased estimates of the effect of the instrument on the outcome (Angrist and Pischke, 2008). If the relationship is statistically significant and the coefficient goes in the expected direction, then we can be more confident in our 2SLS estimates. We run the reduced-form estimations (results are not shown, they are available from the authors on request) and observe that the coefficients on the distance to the closest BFI hospital and its square value have always the same signs as in the first-stage equations, and are always jointly statistically significant at the 5% level except in the case of the BAS picture similarity test at age 5. This confirms that the instrument has the expected effect on the outcomes also in the reduced form equation.

Second, we check for effects of breastfeeding on outcomes which cannot possibly be affected by it, such as birth weight and gestation. A vast epidemiological literature testifies the existence of strong correlations between birth outcomes and maternal characteristics such as education or age. It is therefore not surprising to see a positive and statistically significant relationship between breastfeeding and weight at birth or gestation in the OLS estimates. As shown in the last two columns of Table 6, we see for example that breastfeeding at 4 weeks is associated with having children about 31 grams heavier, and with gestations about 0.7 days longer. The effects of breastfeeding at 8 weeks are very similar. What is interesting to see here is that these correlations change sign and become statistically insignificant in the 2SLS estimates. If our instrumental variable estimates were severely affected by a weak instrument problem, then the 2SLS coefficients would be large, positive and statistically significant also in this case.

Ultimately, we cannot really exclude the possibility that the true "causal" effects of breastfeeding on child cognitive outcomes are indeed quite large. Due to the fact that our estimates are however very imprecise, we take a cautious approach and place more weight on the smaller coefficients. The effect of breastfeeding at 4 weeks on the BAS naming vocabulary at age 5 is estimated to be about 0.7 points of a standard deviation, for example, with a confidence interval which ranges between 0.07 and 1.33. A similar effect is found when considering the coefficient on breastfeeding at 4 weeks on the BAS word reading test at age 7, where the point estimate is 0.89 and the confidence interval ranges between 0.22 and 1.55. When we look at the BAS pattern comprehension scores we find effects of about 0.5 points of a standard deviation at age 5, and 0.72 points of a standard deviation at age 7. In the latter case the confidence interval is between 0.11 and 1.33 points of a standard deviation. Looking at the effects of breastfeeding at 8 weeks provides a very similar picture.

Evidence of the effects of breastfeeding on measures of cognitive ability range between 0.05 and 0.20 points of a standard deviation (Belfield and Kelly, 2010; Borra et al., 2012b; Rothstein, 2012) with some studies pointing out effects up to 0.5 points (Kramer et al., 2008). Although our measures of cognitive ability differ from those used in these other studies, our confidence interval overlaps with these estimates at the lower end, and we find that the effects of breastfeeding on cognitive outcomes are at least in this range, and possibly higher. In what follows we will examine more closely whether the effects we find are robust to a variety of checks, and will present evidence that the effects are heterogeneous across a range of characteristics of the mother and the local area.

For reasons of space we cannot present these checks for all our measures of cognitive ability and/or definitions of breastfeeding, so we will restrict our analysis to the effects of breastfeeding on the BAS naming vocabulary score at age 5, which we take as our benchmark estimate.

4.2 Robustness checks

In Table 7 we check the robustness of our results with respect to a number of changes in the specification and the sample. We show the results of these checks when using the BAS naming vocabulary score measured at age 5 as the reference outcome and, as in previous tables, when considering exclusive breastfeeding at both 4 and 8 months.

The first column shows the effects of breastfeeding on the BAS test score obtained when using as instrumental variable a binary variable indicating whether or not the hospital of birth was either BFI certified or accredited. As we saw in Table 4, this variable has a significant effect on the probability to breastfeed exclusively at 4 or 8 weeks, but could potentially be a choice variable if mothers could select into hospitals. In column 2, we exclude infants born in either very small or very large hospitals. We define small and large by looking at the distribution of hospitals according to size (the number of births in a given period) and eliminating the smallest and largest 5% units.[13] This is because we would like to be confident that our results do not rely on the characteristics of a small number of non-representative maternity units. In column 3 and 4 we exclude individuals who might have located close to a hospital (within 1.6 kilometer or 1 mile) or who live in large metropolitan areas, such as Greater London, where access to more than one hospital is more likely. Finally, in column 5, we exclude individuals living in the least deprived areas, as one might argue that many other forms of breastfeeding support are available in affluent neighbourhoods and our instrument might be confounding hospital with area characteristics.

[13] The specific cutoffs are set at 260 births and 1230 births, so our sample excludes hospitals with less than 260 births and more than 1230 births.

As we can see, the results are in line with what we find in our main specification and in the general sample. If anything, the effect of breastfeeding is sometimes larger and more precisely estimated. Overall, these checks seem to suggest once again that mothers could not easily choose the hospital of birth, or at least did not do so on the basis of information about the availability of breastfeeding support policies, and that selection of individuals into different areas with different levels of access to hospitals or other facilities is not driving our findings.

4.3 Heterogeneity of the effects

It is possible that breastfeeding support policies operating at the hospital level might have a different impact on different groups of mothers. At the same time it is possible that the effects of breastfeeding on cognitive development might be stronger for some children and weaker for others, depending on how the effects of breastfeeding interact with the effects of other types of parental investments and with other characteristics of the child. For these reasons it is interesting to look at the heterogeneity of the effects, and in order to do so we subdivide the sample according to maternal and child characteristics.

In the first two columns of Table 8 we split the sample according to the level of education of the mother. We distinguish between mothers with less than A-levels – the school leaving exam required to access University courses – and mothers with A-levels or higher qualifications. The first thing we notice is that the OLS estimates of the effects of breastfeeding are not statistically significant and very close to zero for the sample of more educated mothers. As these mothers are generally more likely to breastfeed (61.4% of mothers in this group breastfeed exclusively at 4 weeks, against 31.3% of mothers in the other group), but also to foster their children's cognitive development through a variety of other channels, it is plausible that the effects of breastfeeding are not so large for this group. Alternatively, it is possible that more educated mothers who choose not to breastfeed compensate with other types of investments. The 2SLS estimate are in line with the OLS estimates, as they show no impact of breastfeeding on the children of more educated mothers and a strong and significant effect on the children of less educated mothers. Differences in the F-statistic across the two samples suggest that the BFI program had a stronger impact on less educated mothers, although differences in the sample size confound the interpretation.

In the next two columns we split the sample according to the level of income of the family. Here OLS estimates show strong effects of breastfeeding on both high and low income mothers, although there is some suggestion that the effects are higher for the low income group. The 2SLS coefficients show significant and positive effects only for the latter. In this case the sample size is not very different across groups, and the difference in the magnitude of the F-statistic indicates clearly that low income mothers were more responsive to the BFI program. So the lack of effects on high income children is likely to be the result of a weaker response of the high income mothers to the BFI policy.

The last two columns present results for children with birth weight lower or higher than the average (about 3460 grams). OLS effects are clearly significantly different from zero and very similar for these two groups. On the other hand, the 2SLS tell a different story and show that the effects are statistically significant and large only for the heavier and hence likely healthier children. There is no clear indication that this result is due to differences in the first-stage results, so we attribute it to the fact that breastfeeding is more effective the healthier the child is in the first place.

Overall these results seem to suggest that the BFI policy had strong effects on women with lower levels of education and income, who were perhaps less well informed about the benefits of breastfeeding or faced higher costs in acquiring the necessary technology. Women with high level of education or higher income where by contrast less likely to be induced to change their behaviour as a result of hospital level support. This can explain why we find no statistically significant effects of breastfeeding on children from high educated or high income groups. The results also show that children with above average birth weight seem to benefit most. As children of low educated mothers are usually more likely to be lighter at birth, there is a complex interaction between maternal characteristics on the one hand, and child characteristics on the other, which makes it difficult to identify clearly which groups should be targeted by these policies.

4.4 Other outcomes

Epidemiological research makes a strong case that breastfeeding is associated with significant health benefits for children, including lower rates of asthma, allergy and respiratory illness, and fewer infections of the gastrointestinal tract, the middle ear and the urinary tract (American Academy of Pediatrics, 1997, 2005). Most of these findings are however based on observational studies, that is to say on comparisons between breastfed and not breastfed children. Observable maternal and child characteristics are controlled for in the hope that they can capture all the relevant confounding factors.

We know of two relevant exceptions. Kramer et al. (2001) use a randomised controlled trial conducted in Belarus, which assigned breastfeeding support programs to some maternity units and not others. Women who gave birth in the treated hospitals exhibited higher incidence and duration of breastfeeding, and their children showed significant reductions in the risks of gastrointestinal tract infections and atopic eczema, but not in the incidence of a variety of respiratory problems. Baker and Milligan (2008) exploit an extension in paid maternity leave entitlement in Canada, which resulted in large increases in the duration of any breastfeeding (1 month) and exclusive breastfeeding (0.5 months). They find no evidence that the increase in breastfeeding induced by the expanded parental leave policies lead to improvements in child health outcomes including weight, problems of the respiratory system, ear infections and allergies.[14]

The MCS offers a lot of information on child health conditions, including incidence of

[14]Interestingly, a recent epidemiological study conducted by Brion et al. (2011) looks at the association between breastfeeding and a range of child outcomes in countries where there is a negative (or weak) correlation between mother's soecio-economic status and breastfeeding. The study shows that the positive relationship between breastfeeding and health outcomes does not survive the inlcusion of a number of observed characteristics of the mother or the family.

gastrointestinal infections, ear infections, respiratory illnesses, asthma, skin conditions such as eczema, and number of hospitalization episodes. It is important to stress that all these conditions are reported by the mother, so there is no independent observation of whether or not the condition really affects the survey child. Previous studies using the same data have found significant effects of prolonged and exclusive breastfeeding on rates of hospitalization for gastroenteritis and severe respiratory problems such as chest infection and pneumonia (Quigley et al., 2007). As we can see from the upper panel of Table 9, while there are significant differences in the incidence of any health problems at 9 months and 3 years between children who are exclusively breastfed at 4 weeks and those who are not according to OLS regressions, our 2SLS estimates are unable the reject the null hypothesis of no effects.[15] Also, we find no evidence that rates of hospitalization at 3 and 5 years are associated to breastfeeding either in the OLS and the 2SLS estimates.

We also look at whether breastfeeding affects the weight of the child. In particular, we look at the incidence of overweight children in our sample. Height and weight are measured by a nurse, so they are independent of maternal reports. From this information we derive a measure of the Body Mass Index (BMI) and define children with a BMI equal or above 25 as being overweight. Breastfeeding has been shown to reduce overweight and obesity problems in later life in many epidemiological studies (Harder et al., 2005; Scott et al., 2012). Although different criteria have been used in individual studies to define overweight and obesity, the results are quite similar. The only non-observational studies which have looked at this relationship, however, find no evidence of a relationship (Baker and Milligan, 2008; Kramer et al., 2007). Our evidence is consistent with the latter findings, in that we see no significant association between breastfeeding and a child's weight in our 2SLS regressions.[16]

In the next panel we explore the association between breastfeeding and child emotional development. In contrast to what we know about health oucomes, there is much

[15]For space restrictions, we show results for exclusive breastfeeding at 4 weeks, results for exclusive breastfeeding at 8 weeks are very similar. We also show here only evidence on a general indicator for the occurrence of *any* health problem, we also looked at specific conditions and could not find any effects.

[16]We also analyzed the child's weight and BMI as continuous variables, but found the same result.

less evidence of an association between breastfeeding and the psychosocial or emotional development of the child. Although a significant relationship is usually found (Heikkilä et al., 2011), the results are not very robust, hold only for certain groups of mothers or at certain ages of the child (Borra et al., 2012b). We examine this relationship here using information provided in the Strengths and Difficulties Questionnaire[17] and the Child Social Behaviour Questionnaire[18] administered to the MCS mothers.

As we can see from Table 9, the OLS coefficients show a strong and robust association between breastfeeding exclusively at 4 weeks and child emotional development at ages 3, 5 and 7 of the child. The association is always significant at the 1% level and ranges between 0.19 and 0.27 points of a standard deviation, with effects slightly stronger at younger ages and weaker later on. There is a strong similarity between the results obtained when using the Total Difficulty score and the Emotional sub-component of the Self-Regulation score, indicating that both measures are capturing similar aspects of behaviour.[19] After performing 2SLS we find that the effects of breastfeeding predicted using exogenous variation induced by the BFI policy go in the same direction as the OLS estimates, are highly statistically significant and exceed the OLS values. This is consistent with our results on measures of cognitive ability and would seem to suggest that the benefits of breastfeeding extend to non-cognitive aspects of development.

One reason why breastfeeding might have significant effects on the child's cognitive and non-cognitive development is that it may influence maternal behaviours which are

[17]The Strengths and Difficulties Questionnaire (SDQ) is a behavioural screening questionnaire designed to measure psychological adjustment in children aged 3 to 16 (Goodman, 1997). At ages 3, 5 and 7, the MCS parents were asked to answer a battery of 25 questions that identify five different components: (i) hyperactivity/inattention, (ii) conduct problems, (iii) emotional symptoms, (iv) peer problems, and (v) prosocial behaviour. We summed up responses to the first four subscales (i.e. excluding prosocial behaviour) to obtain the Total Difficulty score which we then transformed into a z-score using the mean and the standard deviation observed in our sample.

[18]The Emotional sub-component of the child Self-Regulation score, as measured by selected items from the Child Social Behaviour Questionnaire used on the EPPE study (Sammons et al., 2004), is another measure of emotional development available in the MCS. Higher values on this score indicate a higher ability of the child to concentrate and to control his/her emotional responses. In order to make the comparison with the Total Difficulty score easier, here we have reverse-coded the values of this score and transformed it into a z-score

[19]Indeed, the two scores show a coefficient of correlation which is 0.60 at age 3, 0.67 at age 5, and 0.69 at age 7.

also relevant to these child outcomes. In particular we are interested here in the possibility that breastfeeding may affect maternal employment or maternal well-being. There is a large literature which shows that maternal employment, especially in the first year after birth, is negatively related to child outcomes (James-Burdumy, 2005). At the same time, epidemiological studies suggest that mothers who do not initiate breastfeeding, or who breastfeed for a short time, are more likely than other mothers to become depressed (Dennis and McQueen, 2009; Kendall-Tackett, 2010). Maternal depression is in turn often found to have negative effects on child cognitive abilities and socio-emotional adjustment (Murray et al., 2010). We test these hypotheses using exogenous variation in breastfeeding rates induced by the presence of BFI support in the hospital of birth.

In the lower panel of Table 9 we consider whether breastfeeding affects two indicators of maternal employment: the first is a simple 0/1 dummy for whether or not the mother is working at the time of the first interview, the second is a 0/1 dummy which indicates whether the mother took more than 4 months of maternal leave, which was the maximum amount of paid leave available to women in 2000/01.[20] The OLS estimates reveal that breastfeeding mothers were more likely to be in work at 9 months after the birth of the baby (column 1), and among those who returned to work they were less likely to take more than the maximum amount of paid leave,which was 4 months at the time (column 2). This is consistent with the strong association between breastfeeding and maternal education, as we would expect mothers with higher levels of education to be more likely to return to work and to return to work sooner than lower educated mothers, who have a lower opportunity cost of staying at home. The 2SLS estimates tell a different story, however. The sign on the coefficient indicates that mothers who were induced by the BFI policy to breastfeed were less likely to go back to work at 9 months, and more likely to take longer periods of maternity leave. This goes in the direction we expected, but the coefficients are very close to zero and statistically insignificant.

Finally we turn to the analysis of maternal mental health. Here we measure maternal

[20]This second indicator is calculated only on women who go back to work by 9 months.

mental health or well-being at 9 months by a modified version of the Malaise Inventory Scale (Rutter et al., 1970), and at ages 3, 5 and 7 by the Kessler Scale, a screening device frequently used to diagnose mental illness (Kessler et al., 2003). A higher score on these items indicates the presence of psychological distress or depression. The distribution of the scores is very skewed to the right, so we transform these variables into 0/1 dummies using tresholds identified in clinical studies. As the OLS coefficients show, there is evidence of a negative relationship between these outcomes and breastfeeding. In particular, exclusive breastfeeding at 4 weeks is associated with a decrease in the propensity of showing mental health problems of between 17 and 30%. The 2SLS estimates are generally consistent with these results, although the coefficient are much larger than their OLS counterparts.

It is thus possible that maternal mental health is one of the channels which explains the positive effects of breastfeeding on child cognitive and non-cognitive outcomes. However it is also possible that by *directly* improving child outcomes breastfeeding has *indirect* positive effects on maternal well-being. These indirect effects would be still captured by our 2SLS coefficients. A new study shows that after controlling for antenatal mental health the negative association between breastfeeding and post-natal depression usually found in the epidemiological literature ceases to exist (Borra et al., 2012a), but information on this is not available in the MCS. This evidence leads us to think that rather than being one of the channels through which breastfeeding causes better child cognitive and non-cognitive outcomes, maternal well-being is itself an outcome of the positive link between breastfeeding and child development.

5 Conclusions

This paper has shown that hospital-based breastfeeding support policies can be very effective at increasing breastfeeding initiation and sustaining breastfeeding among mothers who would otherwise not have breastfed, or who would have done so for shorter durations. In particular we find that hospitals' participation in the UNICEF Baby Friendly Initiative increased breastfeeding initiation rates by up to 15 percentage points and increased rates of exclusive breastfeeding at 4 and 8 weeks by about 8-9 percentage points. Given the low levels of breastfeeding observed in the UK these increases are substantial.

We use the BFI program as source of exogenous variation in breastfeeding rates in order to establish the causal effect of breastfeeding on a range of child cognitive outcomes. Convincing causal analyses in this research area are still scarce, as it is difficult to find a valid identification strategy. We dedicate a large part of this paper to examine the validity of our identification strategy and possible threats to it, and we combine three data sources to make sure we control for as many possible confounding factors as possible. To our knowledge, this is the first paper which uses a specific policy intervention to estimate the causal effects of breastfeeding on children cognitive outcomes. This is very important as our results can be immediately translated into policy recommendations.

We find substantial benefits of exclusive breastfeeding at 4 and 8 weeks on a range of cognitive child outcomes observed between ages 3 and 7, including British Ability Scales measures in various domains and the Bracken school readiness test, as well as teacher assessments in schools. The confidence intervals around our point estimates are large, but nonetheless we find evidence of large and positive effects of about 0.7 of a standard deviation.

Heterogeneity analysis reveals that mothers from low income families and with low levels of education are more responsive to the BFI program than highly educated mothers in more affluent families. This shows that hospital-based policies could be effective at attenuating early socio-economic inequalities. However, we also find that children born with above average birth weight - and therefore a better health status - benefit more from being breastfed than children with below average birth weight, so that according to these child characteristics breastfeeding does not seem to have the property of equalizing opportunities between children.

We also look at the effect of breastfeeding on other outcomes. Notably, in contrast to the vast epidemiological literature on the topic but in line with more recent quasi-experimental evidence (Baker and Milligan, 2008; Kramer et al., 2007) we find no statisically significant relationship of breastfeeding with the health outcomes available to us in the Millennium Cohort Survey. On the other hand, when looking at non-cognitive outcomes we do find that breastfeeding reduces child emotional problems as measured in the Strength and Difficulty Questionnaire and is also strongly and significantly associated to maternal mental health.

The Department of Health in England has spent about £4m and £3m on the implementation of the UNICEF BFI program in 2008/09 and 2009/10, respectively (Freedom of Information Request September 2011). This corresponds to £4-5 per baby. This is clearly not a costly intervention, but one that might be quite effective if the returns in terms of enhanced child development and maternal outcomes prove to be as large as our results suggest.

References

Almond, D., Currie, J., 2011. Human Capital Development before Age Five. In: Ashenfelter, O., Card, D. (Eds.), Handbook of Labor Economics. Vol. 4B. Elsevier, Amsterdam.

American Academy of Pediatrics, 1997. Breastfeeding and the Use of Human Milk. Pediatrics 10, 1035–1039.

American Academy of Pediatrics, 2005. Breastfeeding and the Use of Human Milk. Pediatrics 115, 496–506.

Anderson, J. W., Johnstone, B. M., Remley, D. T., 1999. Breast-feeding and Cognitive Development: A Meta-Analysis. American Journal of Clinical Nutrition 70, 525–535.

Angrist, J., Pischke, S., 2008. Mostly Harmless Econometrics: An Empiricist's Companion. University Press, Princeton.

Baker, M., Milligan, K., 2008. Maternal Employment, Breastfeeding, and Health: Evidence from Maternity Leave Mandates. Journal of Health Economics 27, 871–887.

Bartington, S., Griffiths, L. J., Tate, A. R., Dezateux, C., 2006. Are Breastfeeding Rates Higher among Mothers Delivering in Baby Friendly Accredited Maternity Units in the UK? International Journal of Epidemiology 35, 1178–1186.

Belfield, C. R., Kelly, I. R., 2010. The Benefits of Breastfeeding across the Early Years of Childhood. NBER Working Paper 16496.

Bolling, K., Grant, C., Hamlyn, B., Thornton, A., 2007. Infant Feeding Survey 2005. The Information Centre for Health and Social Care, London.

Borra, C., Iacovou, M., Sevilla-Sanz, A., 2012a. Breastfeeding and Maternal Wellbeing, unpublished manuscript.

Borra, C., Iacovou, M., Sevilla-Sanz, A., 2012b. The Effect of Breatsfeeding on Children's Cognitive and Noncognitive Development. Labour Economics 19, 496–515.

Bracken, B. A., 2002. Bracken School Readiness Assessment: Administration Manual. Psychological Corporation, San Antonio Texas.

Brion, M., Matijasevich, A., Horta, B., Anselmi, L., Araujo, C., Menezes, A., Victora, C., Smith, G., 2011. What are the Causal Effects of Breastfeeding on IQ, Obesity, and Blood Pressure? Evidence from Comparing High-Income with Middle-Income Cohort. International Journal of Epidemiology 40, 670–680.

Broadfoot, M., Britten, J., Tappin, D. M., MacKenzie, J. M., 2005. The Baby Friendly Hospital Initiative and Breast Feeding Rates in Scotland. Archives of Disease in Childhood. Fetal and Neonatal Edition 90, F114–116.

Cardwell, C. R., Stene, L. C., Joner, G., Cinek, O., Svensson, J., Goldacre, M., Parslow, R., Pozzilli, P., Brigis, G., Stoyanov, D., Urbonaite, B., Sipetić, S., Schober, E., Ionescu-Tirgoviste, C., Devoti, G., de Beaufort CE, C., Buschard, K., Patterson, C., 2008. Caesarean Section is Associated with an Increased Risk of Childhood-Onset Type 1 Diabetes Mellitus: A Meta-Analysis of Observational Studies. Diabetologia 51, 726–735.

Carneiro, P., Heckman, J., 2004. Human Capital Policy. In: Friedman, B. (Ed.), Inequality in America: What Role for Human Capital Policies. MIT Press, Cambridge.

Cattaneo, A., Buzzetti, R., 2001. Effect on Rates of Breast Feeding of Training for the Baby Friendly Hospital Initiative. British Medical Journal 3213, 1358–1361.

Cunha, F., Heckman, J., 2008. Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation. Journal of Human Resources 43, 738–782.

Dennis, C. L., McQueen, K., 2009. The Relationship between Infant-Feeding Outcomes and Postpartum Depression: A Qualitative Systematic Review. Pediatrics 123, 338–346.

Denny, K., Doyle, O., 2010. The Causal Effect of Breastfeeding on Children's Cognitive Development: A Quasi-Experimental Design. University College Dublin, Working Paper 10/05.

Department of Health, 1994. Weaning and the Weaning Diet. Report on Health and Social Subjects 45. HMSO, London.

Der, G., Batty, G., Deary, I., 2006. Effect of Breast-Feeding on Intelligence in Children: Prospective Study, Sibling Pairs Analysis, and Meta-Analysis. British Medical Journal 333, 945–950.

Dulon, M., Kersting, M., 2003. Breastfeeding Promotion in Non-UNICEF-Certified Hospitals and Long-Term Breastfeeding Success in Germany. Acta Paediatrica 92, 653–658.

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

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

Evenhouse, E., Reilly, S., 2005. Improved Estimates of the Benefits of Breastfeeding Using Sibling Comparisons to Reduce Selection Bias. Health Services Research 40, 1781–1802.

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

Harder, T., Bergmann, R., Kallischnigg, G., Plagemann, A., 2005. Duration of Breast-feeding and Risk of Ooverweight: A Meta-Analysis. Americal Journal of Epidemiology 162, 397–403.

Heikkilä, K., Sacker, A., Kelly, Y., Renfrew, M., Quigley, M., 2011. Breast Feeding and Child Behaviour in the Millennium Cohort Study. Archives of Disease in Childhood 96, 635–642.

House of Commons Health Committee, 2003a. Choices in Maternity Services. Ninth Report of Session 2002-2003. House of Commons, London.

House of Commons Health Committee, 2003b. Provision of Maternity Services. Fourth Report of Session 2002-2003. House of Commons, London.

James-Burdumy, S., 2005. The Effect of Maternal Labor Force Participation on Child Development. Journal of Labor Economics 23, 177–211.

Kendall-Tackett, K. A., 2010. Depression in New Mothers, 2nd Edition. Routledge, Taylor & Francis Group, New York.

Kessler, R., Barker, P., Colpe, L., Epstein, J., Gfroerer, J., Hiripi, E., Howes, M., Normand, S., Manderscheid, R., Walters, E., Zaslavsky, A., 2003. Screening for Serious Mental Illness in the General Population. Archives of General Psychiatry 60, 184–189.

Kleibergen, R., Paap, R., 2006. Generalized Reduced Rank Tests using the Singular Value Decomposition. Journal of Econometrics 133, 97–126.

Kramer, M., Aboud, F., Mironova, E., Vanilovich, I., Platt, R., Matush, L., Igumnov, S., Fombonne, E., Bogdanovich, N., Ducruet, T., Collet, J., Chalmers, B., Hodnett, E., Davidovsky, S., Skugarevsky, O., Trofimovich, O., Kozlova, L., Shapiro, S., 2008. Breastfeeding and Child Cognitive development - New Evidence from a Large Randomized Trial. Archives of General Psychiatry 65, 578–584.

Kramer, M., Hodnett, B. C. E. D., , Sevkovskaya, Z., Dzikovich, I., Shapiro, S., JP, J. C., Mezen, I. V. I. I., Ducruet, T., Shishko, G., Zubovich, V., Mknuik, D., Gluchanina, E., Dombrovskiy, V., Ustinovitch, A., Kot, T., Bogdanovich, N., Ovchinikova, L., Helsing, E., 2001. Promotion of Breastfeeding Intervention Trial (PROBIT). A Randomized Trial in the Republic of Belarus. Journal of the American Medical Association 285, 413–420.

Kramer, M., L, L. M., Vanilovich, I., Platt, R., Bogdanovich, N., Sevkovskaya, Z., Dzikovich, I., Shishko, G., Collet, J., Martin, R., Smith, G. D., Gillman, M., B, B. C., Hodnett, E., Shapiro, S., 2007. Effects of Prolonged and Exclusive Breastfeeding on Child Height, Weight, Adiposity, and Blood Pressure at Aage 6.5y: Evidence from a Large Randomized Trial. American Journal of Clinical Nutrition 86, 1717–1721.

Li, H., Ye, R., Achenbach, M., Ren, A., Pei, L., Zheng, X., Liu, J., 2011. Caesarean Delivery on Maternal Request and Childhood Psychopathology: A Retrospective Cohort Study in China. British Journal of Obstetrics and Gynaecology 118, 424–8.

McCrory, 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.

Merten, S., Dratva, J., Ackermann-Liebrich, U., 2005. Do Baby-Friendly Hospitals Influence Breastfeeding Duration on a National Level? Pediatrics 116, e702–e708.

Murray, L., Halligan, S., Cooper, P., 2010. Effects of Postnatal Depression on Mother-Infant Interactions and Child Development. Wiley-Blackwell.

Page, L., McCourt, C., Beake, S., Hewison, J., Vail, A., 1999. Clinical Interventions and Outcomes of One-to-One Midwifery Practice. Journal of Public Health Medicine 21, 243–248.

Petryk, A., Harris, S., Jongbloed, L., 2007. Breastfeeding and Neurodevelopment: A Literature Review. Infants and Young Children 20, 120–134.

Quigley, M., Kelly, Y., Sacker, A., 2007. Breastfeeding and Hospitalization for Diarrheal and Respiratory Infection in the UK Millennium Cohort Study. Pediatrics 119, e837–e842.

Rees, D., Sabia, J., 2009. The Effect of Breast Feeding on Educational Attainment: Evidence from Sibling Data. Journal of Human Capital 3, 43–72.

Rothstein, D., 2012. Breastfeeding and Childrens Early Cognitive Outcomes. Review of Economics and Statistics forthcoming.

Rutter, M., Tizard, J., KWhitmore, 1970. Education, Health and Behaviour. Longmans, London.

Sammons, P., Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I., Taggart, B., 2004. The Effective Provision of Pre-school Education (EPPE) Project. Technical Paper 11: The Continuing Effect of Pre-School Education at Age 7 Years. Institute of Education, London.

Scott, J. A., Ng, S. Y., Cobiac, L., 2012. The Relationship between Breastfeeding and Weight status in a National Sample of Australian Children and Adolescents. BMC Public Health 12, e682–e689.

Stock, J., Yogo, M., 2005. Testing for Weak Instruments in Linear IV Regression. In: D. W. K. Andrews, and J. H. Stock (Ed.), Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg. Cambridge University Press, Cambridge.

Thavagnanam, S., Fleming, J., Bromley, A., Shields, M. D., Cardwell, C. R., 2008. A Meta-Analysis of the Association Between Caesarean Section and Childhood Asthma. Clinical and Experimental Allergy 38, 629–633.

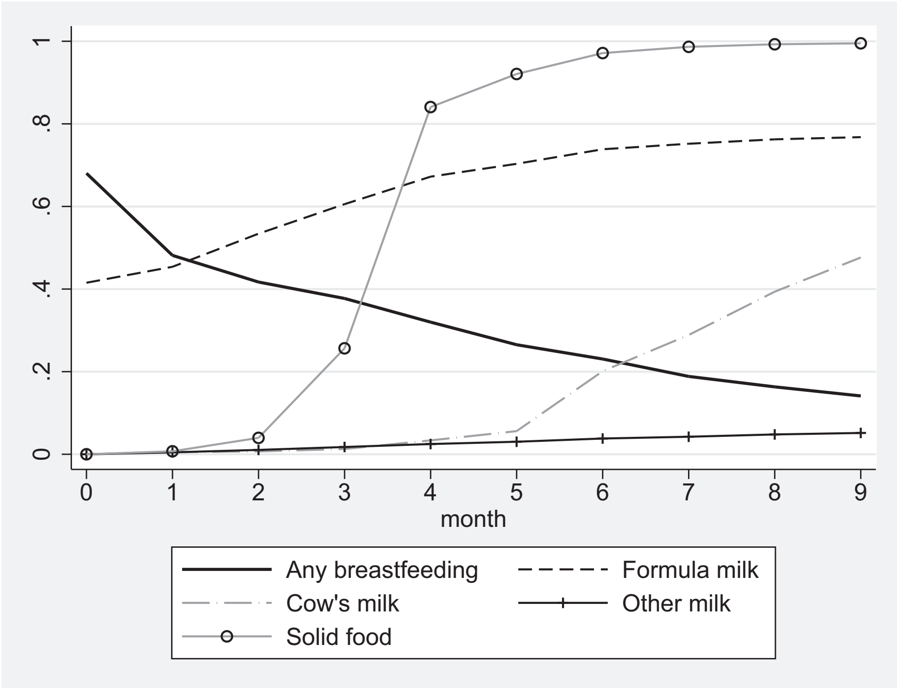
Thomas, J., Paranjothy, S., Royal College of Obstetricians and Gynaecologists, Clinical Effectiveness Support Unit, 2001. National Sentinel Caesarean Section Audit Report. RCOG Press, London.

Vella, F., 1993. A Simple Estimator for Simultaneous Models with Censored Endogenous Regressor. International Economic Review 34, 641–457.

Vella, F., 1998. Estimating Models with Sample Selection Bias: A Survey. Journal of Human Resources 33, 127–169.

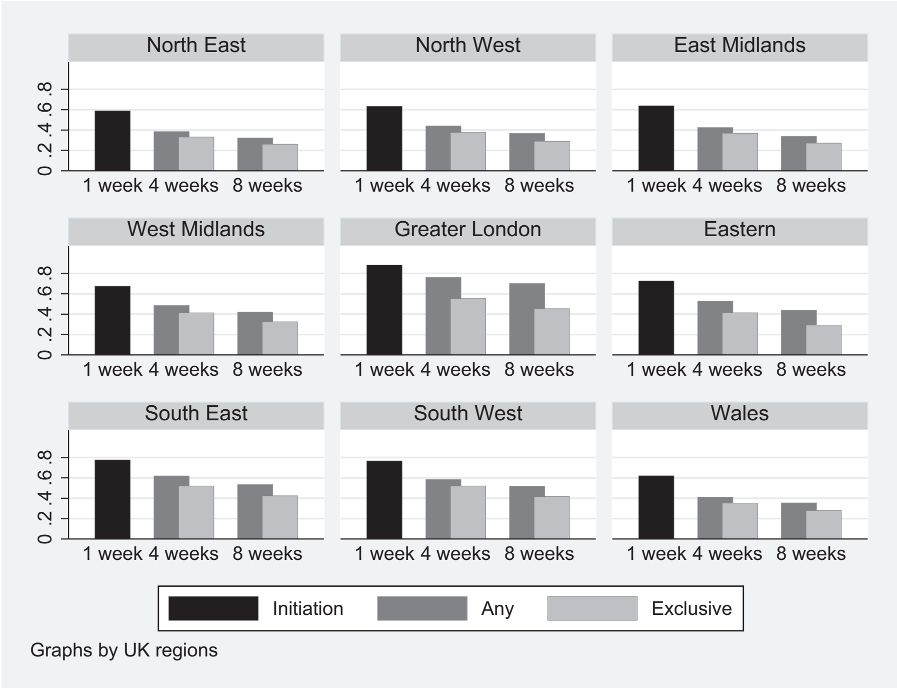
World Health Assembly, 2001. Infant and Young Child Nutrition: Resolution 54.2. The Optimal Duration of Exclusive Breastfeeding. World Health Organization, Geneva.

Figure 1: Feeding practices of MCS infants



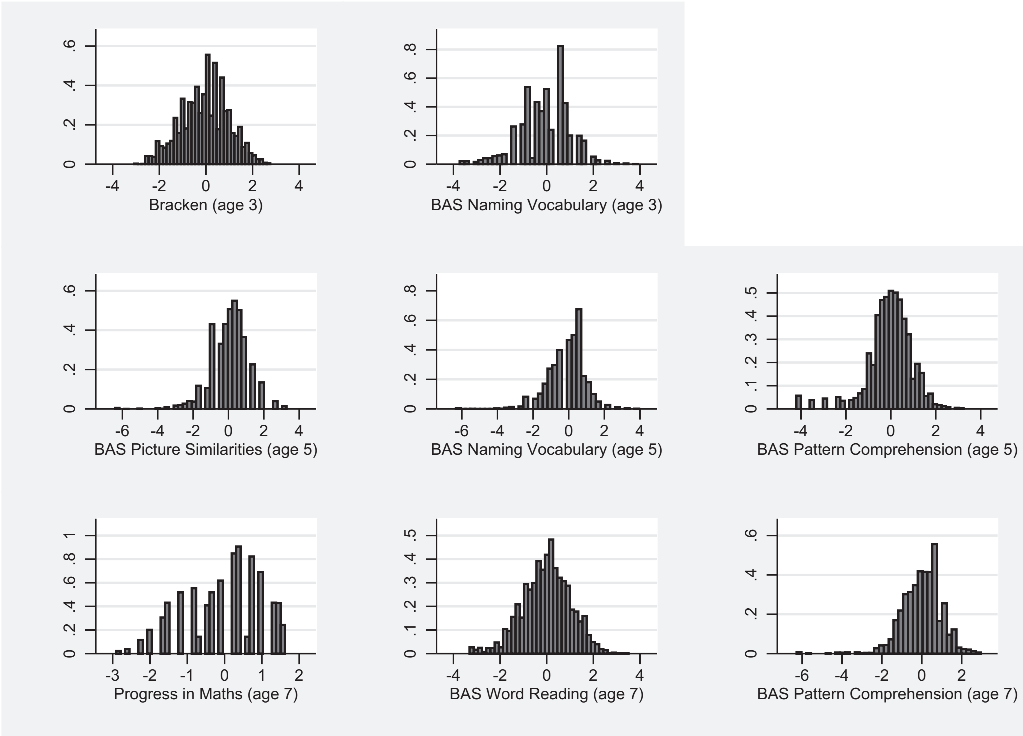
Notes: Millennium Cohort Study estimation sample. All data are weighted using sample weights.

Figure 2: Incidence of breastfeeding by duration and prevalence in regions of England and in Wales



Notes: Millennium Cohort Study estimation sample. All data are weighted using sample weights.

Figure 3: Density plots of cognitive tests at various ages of the child



Notes: Millennium Cohort Study estimation sample. All data are weighted using sample weights.

Table 1: Baby Friendly status of MCS infants by country and BFI status of the maternity unit

		Baby Friendly Status of the maternity unit			
		Full accreditation	**Certif. of commitment**	**Neither award**	**Total**
England	Maternity units	6	39	120	165
	Births	199	2,432	5,140	7,771
Wales	Mat. units	1	5	9	15
	Births	96	801	926	1,823
Total	Maternity units	7	44	129	180
	Births	295	3,233	6,066	9,594

Notes: Millennium Cohort Study estimation sample according to UNICEF UK Baby Friendly status of the hospital of birth.

Table 2a: Are BFI hospitals different?

	Neither award	**Cert. of Commit.**	**Full Accred.**	**t-test diff. in means**	
	mean (sd) (a)	*mean (sd)* (b)	*mean (sd)* (c)	*diff. (se)* (b)-(a)=0	*diff. (se)* (c)-(a)=0
No. of births per quarter	682.08	763.14	726.43	81.06	44.35
	(287.85)	(296.63)	(264.14)	(50.50)	(112.24)
Age mothers: mean	29.07	28.66	29.00	-0.41	-0.07
	(1.44)	(1.52)	(1.73)	(0.26)	(0.57)
Age mothers: 25th perc.	24.84	24.16	24.29	-0.68**	-0.55
	(1.79)	(1.85)	(2.06)	(0.32)	(0.70)
Age mothers: 75th perc.	32.79	32.50	32.43	-0.31	-0.37
	(1.10)	(1.32)	(1.51)	(0.21)	(0.46)
% white	85.17	84.40	86.63	-0.77	1.47
	(19.24)	(17.12)	(12.45)	(3.24)	(7.20)
% nulliparous	41.05	41.95	40.43	0.90	-0.62
	(3.61)	(3.82)	(3.15)	(0.64)	(1.43)
% multiple births	1.42	1.48	1.59	0.06	0.17
	(0.54)	(0.55)	(0.28)	(0.09)	(0.21)
% cephalic presentation	94.47	94.57	94.21	0.09	-0.26
	(1.04)	(1.01)	(0.65)	(0.178)	(0.40)
% spontaneous labour	66.73	66.92	65.45	0.19	-1.29
	(5.59)	(4.46)	(3.84)	(0.92)	(2.05)
% caesarean	21.78	21.42	20.21	-0.36	-1.57
	(4.21)	(3.67)	(2.58)	(0.71)	(1.57)
% births <2500g	5.92	6.42	5.72	0.49	-0.21
	(1.85)	(1.58)	(2.29)	(0.32)	(0.70)
% preterm babies	7.06	7.72	6.65	0.66	-0.42
	(1.88)	(1.99)	(2.59)	(0.34)	(0.75)
Midwifery care scheme					
Midwifery-led care	0.42	0.45	0.57	0.03	0.15
				(0.09)	(0.19)
Core midwives	0.59	0.70	0.71	0.12	0.13
				(0.08)	(0.19)
Rotation	0.71	0.63	0.71	-0.08	0.00
				(0.08)	(0.18)
Hospital-based teams	0.18	0.16	0.29	-0.02	0.11
				(0.07)	(0.15)
Commun.-based teams	0.45	0.45	0.43	0.00	0.02
				(0.09)	(0.19)
Caseload teams	0.09	0.16	0.00	0.07	-0.09
				(0.05)	(0.12)
Other midwifery care	0.28	0.32	0.43	0.04	0.15
				(0.08)	(0.18)
N	129	44	7		

Notes: Millennium Cohort Study estimation sample, UNICEF UK Baby Friendly, National Sentinel Caesarean Section Audit. Mean and standard deviation (only for continuous variables) reported in columns (a) to (c). Column (d) and (e) report the difference between the values (b) and (a), and the values (c) and (a), as well as the standard error of these differences. Symbols: + significant at 10% level; * significant at 5% level; ** significant at 1% level.

Table 2b: Are mothers who give birth and babies born in BFI hospitals different?

	Neither award	**Cert. of Commit.**	**Full Accred.**	**t-test diff. in means**	
	mean (sd)	*mean (sd)*	*mean (sd)*	*diff. (se)*	*diff. (se)*
	(a)	(b)	(c)	(b)-(a)=0	(c)-(a)=0
Mother's age at birth	29.19	29.39	28.55	-0.20	-0.64
	(5.62)	(5.88)	(6.54)	(0.334)	(0.96)
Mother less than A-levels	0.64	0.62	0.62	-0.02	-0.02
				(0.03)	(0.08)
Low income household	0.53	0.52	0.54	-0.01	-0.01
				(0.04)	(0.09)
Mother smoked dur. pregn.	0.36	0.36	0.35	-0.00	-0.001
				(0.02)	(0.04)
Partner at birth	0.84	0.85	0.82	0.00	-0.03
				(0.02)	(0.04)
Normal delivery	0.70	0.70	0.74	-0.00	0.04
				(0.01)	(0.03)
Birthweight	3463.64	3463.23	3441.15	-0.41	-22.50
	(445.09)	(472.65)	(499.08)	(14.06)	(30.60)
Gestation (weeks)	40.01	40.00	39.82	-0.01	-0.19**
	(1.28)	(1.37)	(1.42)	(0.04)	(0.09)
White ethnicity	0.85	0.85	0.90	-0.00	0.05
				(0.04)	(0.05)
Male child	0.52	0.50	0.49	-0.01	-0.02
				(0.01)	(0.03)
Firstborn	0.42	0.43	0.42	0.01	0.00
				(0.01)	(0.03)
Child's age (days)	295.42	295.12	296.91	-0.26	1.49
at 1st interview	(14.33)	(14.15)	(16.07)	(0.69)	(1.55)
More depr. areas	0.51	0.60	0.58	0.09	0.07
				(0.06)	(0.14)
Dist. to closest hosp.	5.96	4.94	7.23	-1.02**	1.26
	(4.37)	(3.29)	(4.61)	(0.48)	(1.33)
N	6,066	3,233	295		

Notes: Millennium Cohort Study estimation sample, UNICEF UK Baby Friendly. Mean and standard deviation (only for continuous variables) reported in columns (a) to (c). Column (d) and (e) report the difference between the values (b) and (a), and the values (c) and (a), as well as the standard error of these differences. All calculations take into account sampling weights. Symbols: + significant at 10% level; * significant at 5% level; ** significant at 1% level.

Table 3: Effect of BFI initiative on various measures of breastfeeding

	Logit (odds ratios)					Logit (m.e.)	OLS (coeff.)
Breastfeeding initiation							
Cert. of Commit.	1.187	1.280*	1.365**	1.547**	1.580**	0.082**	0.086**
	(0.156)	(0.152)	(0.118)	(0.104)	(0.103)	(0.012)	(0.011)
Full Accreditation	1.397	1.890**	2.129**	2.258**	2.456**	0.162**	0.150**
	(0.422)	(0.382)	(0.286)	(0.281)	(0.297)	(0.022)	(0.029)
Mean of dep. var.	0.704	0.704	0.704	0.704	0.704	0.704	0.704
Any breastfeeding at 4 weeks							
Cert. of Commit.	1.143	1.201+	1.269**	1.435**	1.473**	0.084**	0.082**
	(0.150)	(0.128)	(0.107)	(0.102)	(0.101)	(0.015)	(0.015)
Full Accreditation	0.872	1.119	1.208	1.147	1.176	0.035	0.027
	(0.387)	(0.371)	(0.326)	(0.216)	(0.181)	(0.033)	(0.031)
Mean of dep. var.	0.522	0.522	0.522	0.522	0.522	0.522	0.522
Exclusive breastfeeding at 4 weeks							
Cert. of Commit.	1.203+	1.273**	1.322**	1.481**	1.519**	0.094**	0.091**
	(0.129)	(0.116)	(0.097)	(0.092)	(0.093)	(0.014)	(0.014)
Full Accreditation	1.039	1.240	1.333	1.330+	1.337+	0.065+	0.065*
	(0.410)	(0.368)	(0.312)	(0.213)	(0.204)	(0.034)	(0.032)
Mean of dep. var.	0.430	0.430	0.430	0.430	0.430	0.430	0.430
Any breastfeeding at 8 weeks							
Cert. of Commit.	1.158	1.206+	1.273**	1.460**	1.501**	0.087**	0.083**
	(0.154)	(0.129)	(0.112)	(0.104)	(0.106)	(0.015)	(0.016)
Full Accreditation	1.063	1.383	1.500	1.357	1.400+	0.072+	0.065+
	(0.449)	(0.436)	(0.374)	(0.268)	(0.242)	(0.034)	(0.035)
Mean of dep. var.	0.449	0.449	0.449	0.449	0.449	0.449	0.449
Exclusive breastfeeding at 8 weeks							
Cert. of Commit.	1.210+	1.264**	1.306**	1.457**	1.496**	0.084**	0.080**
	(0.129)	(0.110)	(0.099)	(0.096)	(0.100)	(0.014)	(0.014)
Full Accreditation	1.243	1.443	1.541	1.324	1.328	0.059	0.065+
	(0.415)	(0.353)	(0.257)	(0.241)	(0.259)	(0.040)	(0.037)
Mean of dep. var.	0.336	0.336	0.336	0.336	0.336	0.336	0.336
Regional dummies	no	yes	yes	yes	yes	yes	yes
Local area char.	no	no	yes	yes	yes	yes	yes
Hospital char.	no	no	no	yes	yes	yes	yes
Individual char.	no	no	no	no	yes	yes	yes
N	9,524	9,524	9,524	9,524	9,524	9,524	9,524

Notes: Data from Millennium Cohort Study, UNICEF UK Baby Friendly, National Sentinel Caesarean Section Audit. Specifications (1)-(6) estimated by logit regression, specification 7 estimated by linear regression. All estimates take into account sampling weights. Robust standard errors, clustered at the hospital level, in parentheses. Regions: 9 dummies for region of hospital. Local area characteristics: deprivation deciles as measured by Multiple Deprivation Indices (9 dummies). Hospital characteristics: number of births in maternity unit and its square, mean, 25th and 75th percentile of maternal age, percentage white women, percentage nulliparous mothers, percentage multiple births, percentage cephalic presentation, percentage spontaneous labour, percentage caesarean, percentage low birth weight babies, percentage premature babies, plus combinations of the above (Robson Groups), indicators describing the type of midwifery care (9 dummies). Individual controls: male baby, firstborn baby, non-white ethnicity, distance to closest maternity unit, distance to closest maternity unit squared. Symbols: + significant at 10% level; * significant at 5% level; ** significant at 1% level.

Table 4: Weak identification tests for different versions of the instrument

	Separate BFI dummies	**Any BFI dummy**	**Distance to closest BFI unit**
Exclusive breastfeeding at 4 wks			
Cert. of Commit.	0.091**		
	(0.014)		
Full Accreditation	0.065*		
	(0.032)		
Any BFI status		0.090**	
		(0.014)	
Distance to closest BFI hospital			-0.005**
			(0.001)
Distance to closest BFI hospital squared			0.000**
			(0.000)
F-statistic	19.87	38.62	26.49
Critical values	[19.93]	[16.38]	[19.93]
Total R^2	0.080	0.080	0.079
Partial R^2	0.006	0.006	0.005
Exclusive breastfeeding at 8 wks			
Cert. of Commit.	0.080**		
	(0.014)		
Full Accreditation	0.065+		
	(0.037)		
Any BFI status		0.079**	
		(0.014)	
Distance to closest BFI hospital			-0.005**
			(0.001)
Distance to closest BFI hospital squared			0.000**
			(0.000)
F-statistic	16.11	31.01	31.04
Critical values	[19.93]	[16.38]	[19.93]
Total R^2	0.070	0.070	0.070
Partial R^2	0.005	0.005	0.005

Notes: Data from Millennium Cohort Study, UNICEF UK Baby Friendly, National Sentinel Caesarean Section Audit. All specifications control for all the variables included in columns 5-7 of Table 3. All specifications are estimated by linear regression and take into account sampling weights. Robust standard errors, clustered at the hospital level, in parentheses. First column shows results obtained when using separate dummies indicating whether the baby was born in a BFI accredited or certified hospital; the second column refers to a specification where breastfeeding rates are regressed on a single dummy variable indicating whether the birth took place in a hospital that was BFI accredited or certified; the third column reports a specification where breastfeeding rates are regressed on the distance in metres to the nearest BFI hospital and its square. Also shown: Kleibergen Paap Wald F-statistic and ciritcal values corresponding to the Stock and Yogo 10% maximal IV size statistic. Symbols: + significant at 10% level; * significant at 5% level; ** significant at 1% level.

Table 5a: OLS and 2SLS estimates of the effects of exclusive breastfeeding at 4 weeks on main cognitive child outcomes

		Bracken	BAS Naming Vocab.	BAS Picture Simil.	BAS Pattern Constr.	BAS Word Read.	Progress in Maths
Exclusive breastfeeding at 4 weeks							
3 years	OLS	0.263**	0.166**				
		(0.029)	(0.030)				
	2SLS	1.476**	1.080**				
		(0.408)	(0.375)				
	F-statistic	28.87	31.40				
	N	6,972	7,351				
5 years	OLS		0.164**	0.195**	0.140**		
			(0.025)	(0.024)	(0.029)		
	2SLS		0.699*	0.020	0.493		
			(0.321)	(0.451)	(0.458)		
	F-statistic		25.93	26.24	27.09		
	N		7,480	7,487	7,430		
7 years	OLS				0.197**	0.168**	0.165**
					(0.027)	(0.026)	(0.028)
	2SLS				0.713*	0.872**	1.084**
					(0.307)	(0.334)	(0.385)
	F-statistic				28.88	28.01	28.65
	N				6,849	6,761	6,867

Notes: Data from Millennium Cohort Study, UNICEF UK Baby Friendly, National Sentinel Caesarean Section Audit. All outcomes are expressed as z-scores, with mean zero and standard deviation one. All specifications control for all the variables included in columns 5-7 of Table 3. Coefficients estimated by OLS and 2SLS. All regression are weighted using sampling weights. Robust standard errors, clustered at the hospital level, in parentheses. Also shown: Kleibergen Paap Wald F-statistic. Symbols: + significant at 10% level; * significant at 5% level; ** significant at 1% level.

Table 5b: OLS and IV estimates of the effects of exclusive breastfeeding at 8 weeks on main cognitive child outcomes

		Bracken	BAS Naming Vocab.	BAS Picture Simil.	BAS Pattern Constr.	BAS Word Read.	Progress in Maths
Exclusive breastfeeding at 8 weeks							
3 years	OLS	0.227**	0.161**				
		(0.029)	(0.028)				
	2SLS	1.497**	1.113**				
		(0.431)	(0.391)				
	F-statistic	28.17	32.03				
	N	6,972	7,351				
5 years	OLS		0.168**	0.185**	0.146**		
			(0.026)	(0.028)	(0.031)		
	2SLS		0.709*	0.028	0.494		
			(0.320)	(0.460)	(0.452)		
	F-statistic		29.94	30.25	33.51		
	N		7,480	7,487	7,430		
7 years	OLS				0.184**	0.184**	0.165**
					(0.030)	(0.027)	(0.030)
	2SLS				0.744*	0.895**	1.117**
					(0.320)	(0.346)	(0.392)
	F-statistic				31.42	31.04	31.04
	N				6,849	6,761	6,867

Notes: Data from Millennium Cohort Study, UNICEF UK Baby Friendly, National Sentinel Caesarean Section Audit. All outcomes are expressed as z-scores, with mean zero and standard deviation one. All specifications control for all the variables included in columns 5-7 of Table 3. Coefficients estimated by OLS and 2SLS. All regression are weighted using sampling weights. Robust standard errors, clustered at the hospital level, in parentheses. Also shown: Kleibergen Paap Wald F-statistic. Symbols: $^+$ significant at 10% level; * significant at 5% level; ** significant at 1% level.

Table 6: OLS and 2SLS estimates of the effects of breastfeeding on other child outcomes

	EYFSP (age 5)	**Top score writing** (age 7)	**Top score reading** (age 7)	**Top score maths** (age 7)	**Birth weight** (grams)	**Gestation** (weeks)
Exclusive breastfeeding at 4 weeks						
OLS	0.193**	0.103**	0.125**	0.090**	31.904**	0.105**
	(0.026)	(0.018)	(0.018)	(0.016)	(10.587)	(0.031)
2SLS	1.068*	0.458**	0.229	0.240+	-252.799	-0.556
	(0.435)	(0.136)	(0.146)	(0.134)	(161.703)	(0.416)
F-statistic	17.11	20.56	20.09	21.19	26.48	26.48
Mean of dep. var.	0.000	0.338	0.489	0.431	3462.86	40.00
N	5,409	4,291	4,288	4,280	9,594	9,594
Exclusive breastfeeding at 8 weeks						
OLS	0.217**	0.106**	0.127**	0.096**	30.213*	0.123**
	(0.026)	(0.020)	(0.020)	(0.018)	(11.629)	(0.033)
2SLS	1.113**	0.504**	0.288+	0.248	-254.804	-0.553
	(0.429)	(0.154)	(0.163)	(0.156)	(164.980)	(0.420)
F-statistic	19.73	21.02	20.37	21.48	31.09	31.09
Mean of dep. var.	0.000	0.338	0.489	0.431	3462.86	40.00
N	5,409	4,291	4,288	4,280	9,594	9,594

Notes: Data from Millennium Cohort Study, UNICEF UK Baby Friendly, National Sentinel Caesarean Section Audit. Early Years Foundation Stage Profile (EYFSP), writing, reading and maths scores are teacher-assessed. EYFSP scores are expressed as z-scores, with mean zero and standard deviation one. A top score in reading, writing or maths (all binary indicators) is achieving level 3 or 4; level 2 is expected at age 7. All specifications control for all the variables included in columns 5-7 of Table 3. Coefficients estimated by OLS and 2SLS. All regression are weighted using sampling weights. Robust standard errors, clustered at the hospital level, in parentheses. Also shown: Kleibergen Paap Wald F-statistic. Symbols: + significant at 10% level; * significant at 5% level; ** significant at 1% level.

Table 7: Robustness checks

	BAS naming vocabulary at age 5				
	Any BFI hospital dummy as instrument	Excluding small or large hospitals	Excluding individuals living very close to a hospital	Excluding Greater London	Excluding least deprived areas
Exclusive breastfeeding at 4 weeks					
OLS	0.164^{**}	0.162^{**}	0.154^{**}	0.168^{**}	0.187^{**}
	(0.025)	(0.027)	(0.027)	(0.028)	(0.027)
2SLS	0.677^{+}	0.863^{**}	0.725^{*}	0.731^{*}	0.887^{*}
	(0.369)	(0.325)	(0.326)	(0.339)	(0.434)
F-statistic	38.44	28.43	22.54	22.77	16.64
N	7,480	6,685	6,722	6,504	6,415
Exclusive breastfeeding at 8 weeks					
OLS	0.168^{**}	0.174^{**}	0.148^{**}	0.160^{**}	0.211^{**}
	(0.026)	(0.028)	(0.027)	(0.029)	(0.027)
2SLS	0.771^{+}	0.902^{**}	0.726^{*}	0.713^{*}	1.043^{*}
	(0.412)	(0.335)	(0.329)	(0.342)	(0.465)
F-statistic	30.85	32.48	27.72	24.37	14.79
N	7,480	6,685	6,722	6,504	6,415

Notes: Data from Millennium Cohort Study, UNICEF UK Baby Friendly, National Sentinel Caesarean Section Audit. The BAS naming vocabulary score at age 5 is expressed as a z-score, with mean zero and standard deviation one. All specifications control for all the variables included in columns 5-7 of Table 3. Coefficients estimated by OLS and 2SLS. All regression are weighted using sampling weights. Robust standard errors, clustered at the hospital level, in parentheses. Also shown: Kleibergen Paap Wald F-statistic. Symbols: $^{+}$ significant at 10% level; * significant at 5% level; ** significant at 1% level.

Table 8: Heterogeneity

	BAS naming vocabulary at age 5					
	Less than A-levels	A-levels or above	Income <£20,800 per year	Income >£20,800 per year	Bweight below average	Bweight above average
Exclusive breastfeeding at 4 weeks						
OLS	0.118**	0.033	0.133**	0.118**	0.167**	0.156**
	(0.032)	(0.041)	(0.038)	(0.040)	(0.039)	(0.031)
2SLS	0.879+	-0.774	0.711+	-0.094	0.292	1.060*
	(0.479)	(0.838)	(0.401)	(0.692)	(0.430)	(0.338)
F-statistic	9.379	4.419	18.76	7.033	11.29	19.74
N	4,869	2,611	3,994	3,001	4,068	3,412
Exclusive breastfeeding at 8 weeks						
OLS	0.141**	0.024	0.155**	0.111**	0.187**	0.147**
	(0.036)	(0.038)	(0.043)	(0.035)	(0.041)	(0.032)
2SLS	0.770+	-1.167	0.816+	-0.099	0.324	1.163**
	(0.414)	(1.384)	(0.448)	(0.625)	(0.398)	(0.446)
F-statistic	15.01	1.629	21.63	6.610	19.44	12.08
N	4,869	2,611	3,994	3,001	4,068	3,412

Notes: Data from Millennium Cohort Study, UNICEF UK Baby Friendly, National Sentinel Caesarean Section Audit. The BAS naming vocabulary score at age 5 is expressed as a z-score, with mean zero and standard deviation one. All specifications control for all the variables included in columns 5-7 of Table 3. Coefficients estimated by OLS and 2SLS. All regression are weighted using sampling weights. Robust standard errors, clustered at the hospital level, in parentheses. Also shown: Kleibergen Paap Wald F-statistic. Symbols: + significant at 10% level; * significant at 5% level; ** significant at 1% level.

Table 9: Breastfeeding exclusively at 4 weeks and other child and maternal outcomes

Child health						
	Health problem		**Hospitalization**		**Overweight**	
	9 mths	**3 yrs**	**3 yrs**	**5 yrs**	**3 yrs**	**5 yrs**
OLS	-0.021*	-0.042**	-0.021+	-0.004	-0.028*	-0.022+
	(0.010)	(0.014)	(0.012)	(0.008)	(0.012)	(0.012)
2SLS	0.101	0.002	0.146	-0.029	-0.086	-0.058
	(0.266)	(0.120)	(0.095)	(0.094)	(0.103)	(0.113)
F-statistic	26.49	29.21	31.09	25.89	29.42	24.96
Mean dep. var.	0.789	0.593	0.171	0.116	0.233	0.214
N	9,592	7,510	7,755	7,586	7,176	7,510

Child emotional development						
	Total Difficulty (SDQ)			**Emotional Self-regulation**		
	3 yrs	**5 yrs**	**7 yrs**	**3 yrs**	**5 yrs**	**7 yrs**
OLS	-0.266**	-0.202**	-0.198**	-0.224**	-0.193**	-0.206**
	(0.027)	(0.030)	(0.032)	(0.028)	(0.030)	(0.033)
2SLS	-0.922**	-0.736*	-0.985**	-0.747**	-0.865*	-1.458**
	(0.257)	(0.316)	(0.244)	(0.278)	(0.415)	(0.313)
F-statistic	30.23	24.95	29.84	30.34	24.98	29.83
Mean dep. var.	0.000	0.000	0.000	0.000	0.000	0.000
N	7,260	7,305	6,761	7,422	7,359	6,806

Maternal outcomes						
	At work	**Return to work**	**Maternal well-being**			
	9 mths	**>4 mths**	**9 mths**	**3 yrs**	**5 yrs**	**7 yrs**
OLS	0.039**	-0.039*	-0.023*	-0.019*	-0.034*	-0.023*
	(0.011)	(0.015)	(0.009)	(0.009)	(0.008)	(0.010)
2SLS	-0.032	0.021	-0.318**	-0.110	-0.372**	-0.360**
	(0.151)	(0.228)	(0.100)	(0.115)	(0.103)	(0.109)
F-statistic	26.69	8.946	25.74	19.13	24.68	32.41
Mean dep. var.	0.449	0.619	0.135	0.104	0.113	0.121
N	9,581	4,323	9,224	6,773	7,109	6,593

Notes: Data from Millennium Cohort Study, UNICEF UK Baby Friendly, National Sentinel Caesarean Section Audit. Means of dependent variables shown. All outcomes variables are expressed as binary with the exception of the Total Difficulty score and the Emotional Self Regulation score which are expressed as z-scores, with mean zero and standard deviation one. All specifications control for all the variables included in columns 5-7 of Table 3. Coefficients estimated by OLS and 2SLS. The coefficients represent the effects of breastfeeding exclusively at 4 weeks on the indicated outcomes. All regression are weighted using sampling weights. Robust standard errors, clustered at the hospital level, in parentheses. Also shown: Kleibergen Paap Wald F-statistic. Symbols: + significant at 10% level; * significant at 5% level; ** significant at 1% level.